A Tutorial and Review on Inter-Layer FEC Coded Layered Video Streaming

Yongkai Huo*, Cornelius Hellge†, Thomas Wiegand†, *Fellow, IEEE,* and Lajos Hanzo*, *Fellow, IEEE*
*School of ECS, University of Southampton, UK.
†Fraunhofer HHI, Berlin, Germany.
Email: {yh3g09*, lh*}@ecs.soton.ac.uk, {cornelius.hellge†, thomas.wiegand†}@hhi.fraunhofer.de

***Abstract*—Layered video coding creates multiple layers of unequal importance, which enables us to progressively refine the reconstructed video quality. When the base layer (BL) is corrupted or lost during transmission, the enhancement layers (ELs) must be dropped, regardless whether they are perfectly decoded or not, which implies that the transmission power assigned to the ELs is wasted. For the sake of combating this problem, the class of inter-layer forward error correction (IL-FEC) solutions, also referred to as layer-aware FEC (LA-FEC[1]), has been proposed for layered video transmissions, which jointly encode the BL and the ELs, thereby protecting the BL using the ELs. This tutorial aims for inspiring further research on IL-FEC/LA-FEC techniques, with special emphasis on the family of soft-decoded bit-level IL-FEC schemes.**

I. Prologue

Sophisticated video processing is required for recording, production, or re-broadcasting of moving images, which has been changing our everyday life ever since the first clip was captured by a video tape recorder in 1951. Since then a number of video compression standards have been conceived for the sake of reducing the number of bits required, while retaining the fidelity of the visual information conveyed by a video clip. The achievable compression ratio has been substantially improved over the past years both as a benefit of advances in video signal processing and microelectronics. This was achieved at the cost of more complex algorithms, but fortunately the spectacular developments in low-power chip design were still capable of mitigating the power-dissipation imposed. We will briefly review the history of video compression standards in Section II-A of this treatise.

In Section II-A, we will see that these existing or emerging single-layer video compression standards are capable of achieving a high compression ratio in numerous applications scenarios. The first question that arises, is "why do we need layered video coding?" Typically, a single stream is generated by the classic single-layer video encoders, which as a whole will be forwarded to the distant receiver. However, this single layer video is unable to meet the complex multi-platform requirements of the multimedia era, where for example various receiver terminals, such as tablet computers, smart phones and large-screen televisions (TVs) may be used in various scenarios. In order to eliminate this problem, layered video coding became the center of research efforts and a number of standards emerged, such as the scalable video coding (SVC) scheme standardized in H.264. The state-of-the-art in layered video compression techniques will be portrayed in Section II-B, while the above-mentioned question will be answered in more detail in Section II-C.

Layered video coding generates a video stream containing multiple inter-dependent layers. Specifically, a less important layer may depend on a more important layer for decoding. Numerous techniques have been conceived for the sake of improving the attainable video quality by allocating different error protection, different power, etc to different layers. More details found in the literature will be reviewed in Section III. During the early days of layered video communication, for example different code-rates were invoked for the different layers. When the most important base layer is corrupted or lost, the less important layers also have to be dropped, regardless whether they are perfectly received or not, which implies that both the power and the bandwidth assigned to the less important layers was wasted.

The second question that arises, is "can we exploit the valuable resources allocated to the less important layers for recovering the more important layer?" Indeed, the answer is yes. Hence in this tutorial we design so-called inter-layer scheme. In this scheme, the information of the more important layer is implanted into the less important layer at the transmitter. Additionally, the information of the more important layer incorporated into the less important layers may be extracted and utilized for improving the error-resilience of the more important layer. A pair of different solutions will be detailed, namely the *packet level layer-aware (LA) forward error correction (FEC)* solution of Section IV and the *bit-level inter-layer (IL) FEC* solution of Section V. Both their differences and their relationship will be discussed in Section III.

Given that the information of the important layer is implanted into the less important layers, the third question that arises, "whether there is a way to measure how much extra information is gained from the implanted information?" Fortunately, the question may be answered with the aid of the powerful tool of extrinsic information transfer charts (EXIT). This tool will be employed in Section VI for analyzing the benefits of bit-level IL-FEC arrangements, followed by characterizing its performance in the context of both partitioned and scalable H.264/AVC video coding in Sections VII and

The financial support of 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 is gratefully acknowledged.

[1]The philosophy of LA-FEC and IL-FEC was used in [1], [2] for referring to packet-level and bit-level techniques, respectively. Consequently, we will employ LA- /IL-FEC for referring to packet- /bit-level schemes, respectively. The definitions of packet-level and bit-level notations will be given in Section III.

VIII, respectively. Our conclusions are provided in Section IX, where the pertinent design guidelines of layered video coding schemes will be summarized.

II. Introduction - From Uncompressed Video to Scalable Video Compression

Below, we firstly review the history of video compression standards followed by portraying the state-of-the-art in layered video compression techniques. Finally, we answer the question as to "why we need layered video coding?".

A. Video Compression

Uncompressed video clips consist of a sequence of frames captured from a real-world scene, which exhibit high intra- and inter-frame correlation [3], [4] among the pixels. A pair of uncompressed consecutive frames of the Football sequence are exemplified in Fig. 1a. Observe from Fig. 1a that there is substantial intra-frame correlation. Furthermore, the similarity between the two frames seen in Fig. 1a indicates the inter-frame correlation. In Fig. 1b, the transform domain representation of Fig. 1a is displayed after applying the discrete cosine transform (DCT), while the frame difference of Fig. 1a is displayed in Fig. 1c. Intuitively, the correlation residing in an uncompressed video sequence should be removed in order to represent the original video with the aid of less bits, yet without any substantial reduction of the perceived visual quality.

Again, video compression will reduce the storage required in a hard-drive for example, or the transmission bandwidth and the transmission power required for distributing the video. A simple frame-differencing based video codec architecture [5] is displayed in Fig. 2 for illustrating the basic video compression framework. Furthermore, a number of advanced video compression standards [5] have been designed during the past decades for the sake of achieving a high compression ratio, such as the H.120 [6], H.261 [7], H.263 [8], H.264/AVC [9], MPEG-2 [10] and MPEG-4 [5] schemes. The timeline of the video coding standards is shown in Fig. 3. The first video recorder was invented as early as in the 1950s.

1) A Simple Frame-Differencing Video Codec: Assuming that the Football video sequence is encoded by the encoder of Fig. 2a, the consecutive image frames f_n and f_{n-1} typically do not exhibit dramatic scene changes as exemplified in Fig. 1a. Hence, the consecutive frames are similar, a property that we refer to as being correlated. This implies that the current frame f_n can be approximated or predicted by the previous frame, which we express as $f_n \approx \hat{f}_n = f_{n-1}$, where $\hat{f}_n$ denotes the nth predicted frame. When the previous frame f_{n-1} is subtracted from the current one, namely f_n, this prediction typically results in a "line-drawing-like" difference frame, as shown in Fig. 1c. Most areas of this difference frame are "flat", having values close to zero, and the variance or second moment of it is significantly lower than that of the original frame. We have removed some of the predictable components of the video frame.

(a) Two consecutive frames

(b) Transform domain, DCT

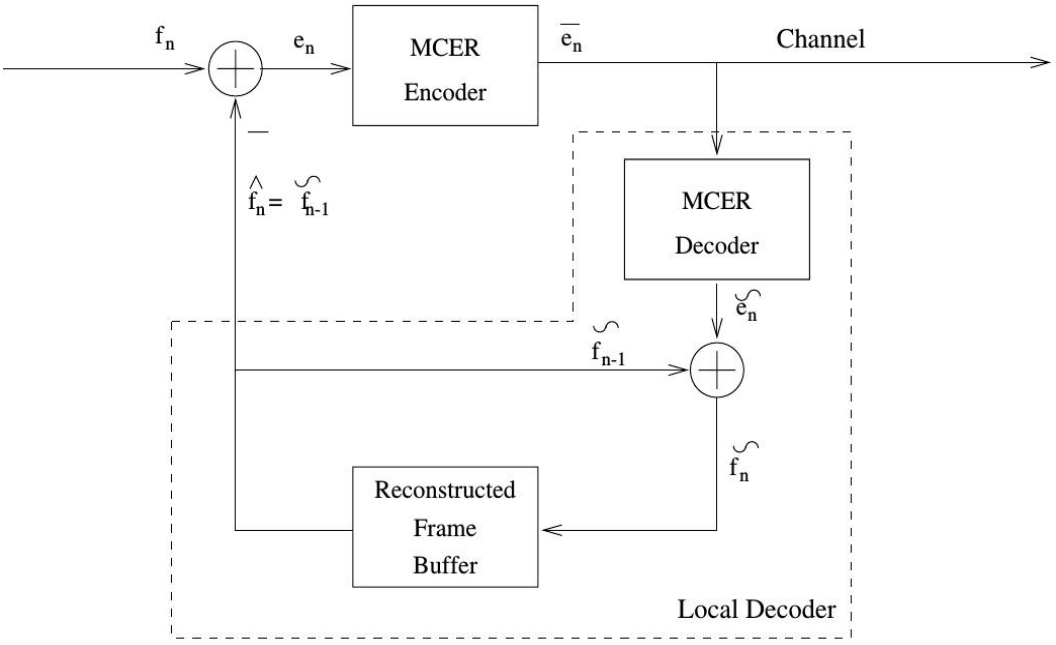
(c) Frame difference

Figure 1: Two consecutive frames of the Football sequence.

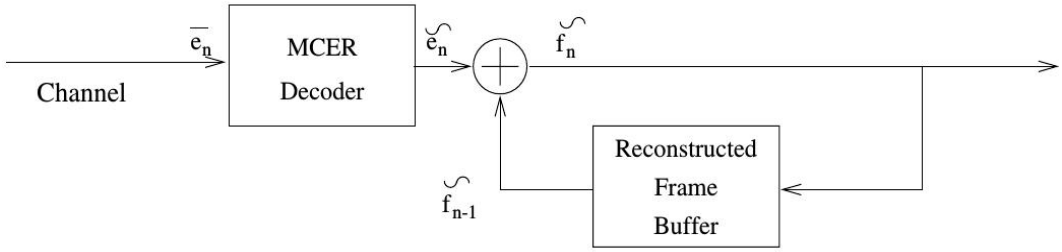
This reduced-variance difference signal of Fig. 2a, namely $e_n = f_n - f_{n-1}$, is often referred to as Motion Compensated Error Residual (MCER) since the associated frame-differencing effectively attempts to compensate for the motion of the objects between consecutive video frames, yielding a reduced-variance MCER. Thus, e_n requires a reduced coding rate, that is, a lower number of bits than f_n in order to represent it with a certain distortion. The MCER $e_n = f_n - f_{n-1}$ can then be encoded as $\bar{e}_n$, at a given distortion using a variety of techniques [5]. The quantized or encoded MCER signal $\bar{e}_n$ is then transmitted over the communications channel. In order to reproduce the original image $f_n = e_n + f_{n-1}$ exactly, knowledge of e_n would be necessary, but only its quantized version, $\bar{e}_n$ is available at the decoder of Fig. 2b, which is contaminated by the quantization distortion introduced by the MCER encoder, inflicting the reconstruction error $\Delta e_n = e_n - \hat{e}_n$. Since the previous undistorted image frame f_{n-1} is not available at the decoder, image reconstruction has to take place using their available approximate values, namely, $\bar{e}_n$ and $\tilde{f}_{n-1}$, giving the reconstructed image frame as follows:

$$\tilde{f}_n = \tilde{e}_n + \tilde{f}_{n-1}, \tag{1}$$

where in Fig. 2a we found that the locally decoded MCER residual $\tilde{e}_n$ is an equal-valued, noiseless equivalent representation of $\bar{e}_n$. The above operations are portrayed in Fig. 2a, where the current video frame f_n is predicted by $\hat{f}_n = \tilde{f}_{n-1}$,



(a) Encoder



(b) Decoder

Figure 2: Basic video codec schematic using frame-differencing [5], where the corresponding video encoder and video decoder are displayed in (a) and (b) respectively.

which is an estimate based on the previous reconstructed frame $\tilde{f}_{n-1}$. Observe that ($\hat{}$) indicates the predicted value and ($\sim$) the reconstructed value, which is contaminated by some coding distortion. Note furthermore that the encoder in Fig. 2a contains a so-called local decoder, which is identical to the remote decoder of Fig. 2b. This measure ensures that the decoder of Fig. 2b uses the same reconstructed frame $\tilde{f}_{n-1}$ in order to reconstruct the image, as the one used by the encoder of Fig. 2a to generate the MCER. Note, however, that in case of transmission errors in $\tilde{e}_n$, the local and the remote reconstructed frames become different, which we often refer to as the encoder and decoder being "misaligned". This phenomenon leads to transmission error propagation through the reconstructed frame buffer, unless counter-measures are employed. To be more explicit, instead of using the previous original frame, the so-called locally decoded frame is used in the motion compensation, where the phrase "locally decoded" implies decoding it at the encoder (i.e., where it was encoded). This local decoding yields an exact replica of the video frame at the distant decoder's output. This local decoding operation is necessary, because the previous original frame is not available at the distant decoder, and so without the local decoding operation the distant decoder would have to use the reconstructed version of the previous frame in its attempt to reconstruct the current frame.

2) H.120 (1984-1988): The International Telecommunication Union (ITU) standardized the recommendation H.120 [6], which is the first digital video coding standard. The techniques employed in H.120 include differential pulse code modulation (DPCM), scalar quantization of pixels and variable-length coding (VLC) of the transform domain coefficients. Although H.120 video failed to achieve adequate quality for practical video compression, it provided important knowledge for its successors.

3) H.261 (1988-1993): The ITU H.261 standard [7] pioneered the era of practical digital video compression techniques during the 1990s, which was designed for transmitting 352×288-pixel Common Intermediate Format [7] (CIF) and for 176×144-pixel Quarter Common Intermediate Format [7] (QCIF) video clips over Integrated Services Digital Networks (ISDN). This standard employed a hybrid video coding scheme, which formed the basis of all state-of-the-art video coding standards.

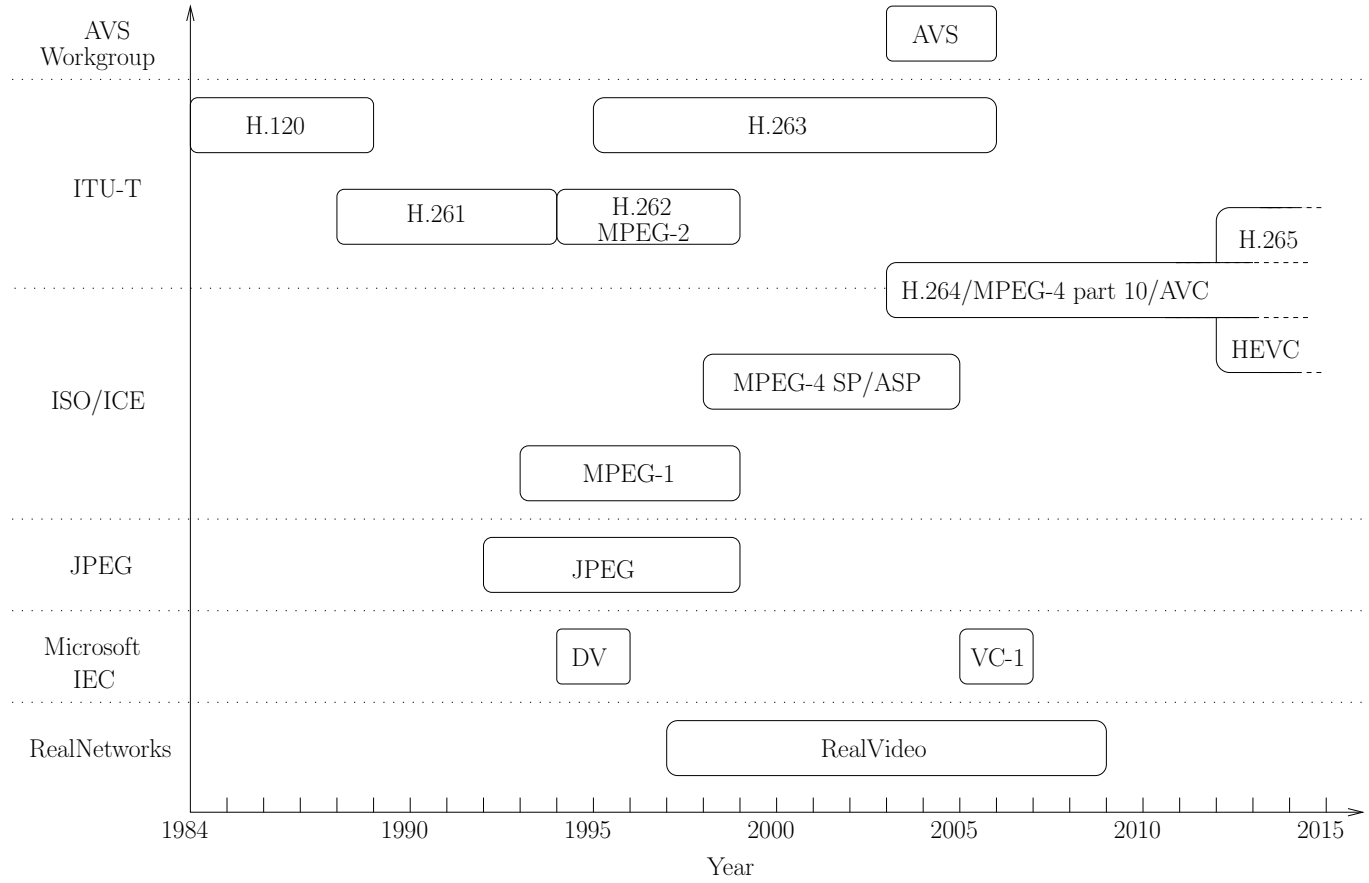


Figure 3: Development timeline of the video compression standards. The timing box indicates the major development period of the corresponding standard.

4) Joint Photographic Experts Group (1992-1998+): Motion-JPEG employs the image compression standard of the Joint Photographic Experts Group [11] (JPEG), which was standardized in 1992 for digital still images. Based on JPEG, Motion-JPEG treats a video clip as a sequence of independent images, which are encoded by JPEG separately. Since the temporal correlation is not removed, Motion-JPEG results in a higher bitrate at a lower computational complexity than that of its motion-compensated counterparts. Although Motion-JPEG may be used for digital cameras and video processing systems, it has not been standardized by any organization.

5) MPEG-1 (1993-1998+): Then the Moving Picture Experts Group (MPEG) proposed the MPEG-1 standard in 1992 [12] for CIF- or 352×240-pixel videos. MPEG-1 may be used by almost all Personal Computers (PCs), Video Compact Disc (VCD) players and Digital Versatile Disc (DVD) players. However, MPEG-1 only supports progressively scanned images, which cannot be used for the interlaced video frames of the National Television System Committee's (NTSC) standard or for the so-called Phase Alternating Line (PAL) video formats [13]. The lack of this compatibility correspondingly prompted the development of the MPEG-2/H.262 standard [10].

6) MPEG-2/H.262 (1994-1998+): In 1993, the International Organization for Standardization (ISO) and ITU jointly developed the MPEG-2/H.262 [10] standard for (720×576)- or (720×480)-pixel resolutions, and for high-definition (HD) video with a pixel resolution of 1920×1080. MPEG-2 is widely used as the format of digital TeleVision (TV) signals, which are broadcast by terrestrial, cable and direct satellite TV systems.

7) DV (1994-1995+): The Digital Video (DV) coding specification [14] was standardized by the International Electrotechnical Commission (IEC) in 1994, mainly targeting camera recorders, which encodes a video clip on a frame-by-frame basis, i.e. by dispensing with motion compensation. The DV efficiency is comparable to that of the Intra-frame coding mode of MPEG-2/H.262 and it is better than that of Motion-JPEG.

8) H.263 (1995-2005): The H.263 specification was standardized in 1995 by the ITU Telecommunication Standardization Sector [8], [15], [16] (ITU-T), which targeted video conferencing at low bitrates for mobile wireless communications and it is superior to all prior standards in terms of its compression efficiency. The H.263v2 standard, namely H.263+, was completed in 1998, which is the informal acronym for the second edition of the ITU-T H.263 standard. In this edition, the capabilities of the H.263 scheme were enhanced by adding several annexes for detailing, how to achieve a substantially improved encoding efficiency. Later, the definition of H.263v3, also known as H.263++, added three further annexes in 2000.

9) RealVideo (1997-2008+): RealVideo [17], [18] is a successful proprietary video compression format developed by the RealNetworks company, which was first released in 1997. RealVideo is supported by numerous computing platforms, including Windows, Mac, Linux, Solaris and several mobile phones.

10) MPEG-4 SP/ASP (1998-2004+): MPEG-4 standardization [5] was initiated in 1995 and has been continually enhanced by a number of new profiles, including wavelet-based still image coding, scalable coding, 3D images etc. The MPEG-4 video standard strikes adjustable compression quality versus bitrate trade-off, where the so-called Simple Profile (SP) is very similar to H.263, while the Advanced Simple Profile (ASP) further increased the compression efficiency attained.

11) WMV9/VC-1/SMPTE 421M (2005-2006+): The society of Motion Picture and Television Engineers (SMPTE) 421M [19] developed a scheme, which is, also known as Windows Media Video version 9 (WMV9) or VC-1. This was initially developed as a proprietary video format by Microsoft, but it was then released as a SMPTE video codec standard in 2006. VC-1 was designed as an alternative to the latest ITU-T and MPEG video codec H.264/MPEG-4 AVC. It was shown that the VC-1 compression efficiency is lower than that of H.264/AVC, while imposing a reduced computational complexity.

12) H.264/MPEG-4 part 10/AVC (2003-2014+): The H.264/AVC standard [9], also known as MPEG-4 part 10 or Advanced Video Coding (AVC), was completed in May 2003, albeit its research continued by adding more extensions. H.264/MPEG-4 part 10/AVC was jointly developed by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC JTC1 MPEG, which became one of the most commonly used formats of video recording, compression and distribution. The design goal was to halve the bitrate required by the previous video standards, while retaining the same video quality. A major recent extension of H.264/AVC was the Scalable Video Coding [20] (SVC) scheme completed in 2007, which is specified in Annex G, allowing the construction of bitstreams that contain sub-bitstreams of the H.264/AVC standard. Another major extension of H.264/AVC was the Multiview Video Coding (MVC) scheme completed in 2009 [21], which is specified in Annex H, enabling the construction of bitstreams that represent more than one view of a video scene. The MVC extension contains two profiles, which are the Multiview High Profile [9] representing an arbitrary number of views and the Stereo High Profile [9] for stereoscopic video.

13) Audio and Video Coding Standard (2003-2005): The Advanced Video Standard (AVS) [22] was initiated by the government of China for replacing MPEG-2, which has been standardized in 2005.

14) HEVC/MPEG-H Part 2/H.265 (2012-2014+): High Efficiency Video Coding (HEVC) [23] is undergoing development as a successor to H.264/AVC, which is being jointly developed by the ISO/IEC MPEG and ITU-T VCEG as ISO/IEC 23008-2 MPEG-H Part 2 and ITU-T H.265. HEVC aims for halving the bitrate of the H.264/AVC standard at the same video quality and it aims for supporting Ultra High Definition (UHD) videos at a resolution of 8192×4320. The emerging H.265 scheme will continue to support scalability by the scalable HEVC (SHVC) profile, which enables the transmitter to meet multiple users' preferences.

B. Layered Video Coding Standards

Layered video compression [9], [20], [26], [27] encodes a video sequence into multiple layers, which enables us to progressively refine the reconstructed video quality at the receiver. Generally, the most important layer is referred to as the base layer (BL) and the less important layers are termed as enhancement layers (ELs), which rely on the BL. Furthermore, an EL may be further relied upon by less important ELs. Again, when the BL or an EL is lost or corrupted during its transmission, the dependent layers cannot be utilized by the decoder and must be dropped. A layered video scheme is displayed in Fig. 4, where the video sequence captured from the scene is encoded into four layers by the layered video encoder, namely $L_0 \sim L_3$, where layer L_i $(0 < i \leq 3)$ depends on layer L_{i-1} for decoding, while layer L_i improves the video quality of layer L_{i-1}. In other words, layer L_0 is the BL and layers $L_1 \sim L_3$ are ELs depending on the BL. Furthermore, as shown in Fig. 4, the ELs L_2 and L_3 rely on the EL L_1. In other words, if layer L_1 is corrupted, then layers L_2 and L_3 are dropped by the decoder. A number of layered video coding techniques have been investigated and/or standardized, which will be introduced below.

1) Partition Mode of H.264: A number of layered video coding schemes [28] have been developed and some of them are adopted by recent video coding standards, for example the scalable video coding (SVC) [20] and data partitioned mode (DP) [9], [29], [30]. In the data partitioning mode, the data streams representing different semantic importance are categorized into a maximum of three bitstreams/partitions [31] per video slice, namely type A, type B and type C partition. The header information, such as macroblock (MB) types, quantization parameters and motion vectors are carried by the type A partition. The type B partition is also referred to as the intra-partition, which contains intra-frame-coded information, including the coded block patterns (CBPs) and intra-coded coefficients. The type B partition is capable of prohibiting error propagation in the scenario, when the reference frame of the current frame is corrupted. In contrast to the type B partition, the type C partition is the inter-partition, which carries the inter-CBPs and the inter-frame coded coefficients. The type C partition has to rely on the reference frame for reconstructing the current picture. Hence, if the reference picture is corrupted, errors may be propagated to the current frame. Amongst these three partitions, the type A partition may be deemed to be the most important one, which may be treated as the BL. Correspondingly, the type B and C partitions may be interpreted as a pair of enhancement layers, since they are dependent on the type A partition for decoding. Albeit the information in partition B and C cannot be used in the absence of partition A, partition B and partition C can be used independently of each other, given the availability of partition A. The dependency of the layers in the partitioned mode of

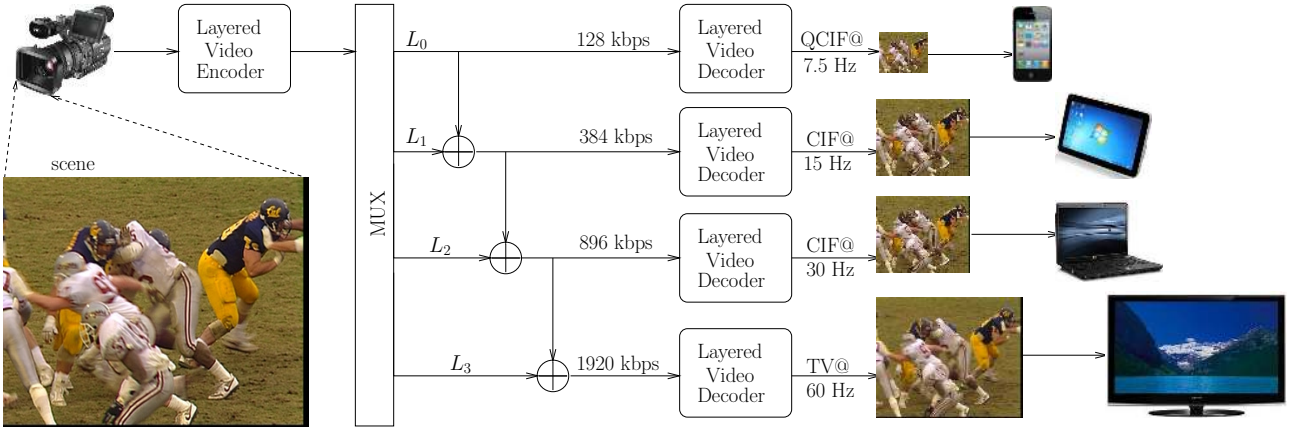


Figure 4: Architecture of a layered video scheme [24], [25], where the video quality is refined gradually.

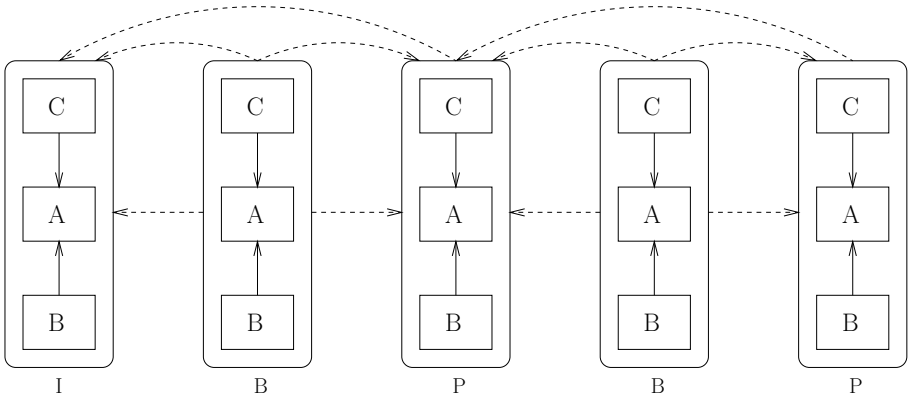


Figure 5: Dependencies associated with the GOP=2 partition mode of the H.264/AVC video stream, where A, B and C represent the partition A, B and C, respectively. "B$\rightarrow$A" indicates B *depends on* A, while "B$\dashrightarrow$A" indicates B *is predicted from* A.

the H.264/AVC video codec is exemplified in Fig. 5[2], where the group of pictures (GOP) parameter is GOP=2.

2) Scalable Video Coding: The subject of scalable video coding [20], [32] has been an active research field for over two decades. This terminology is also used in the Annex G extension of the H.264/AVC video compression standard [9]. Indeed, SVC is capable of generating several bitstreams that may be decoded at a similar quality and compression ratio to that of the existing H.264/AVC codec. When for example low-cost, low-quality streaming is required by the users, some of the ELs may be removed from the compressed video stream, which facilitates flexible bitrate-control based on the specific preferences of the users. A H.264/AVC scalable video stream contains a sequence of network abstraction layer units (NALUs) [9], which consist of a header and a payload. The header contains the information about the type of NALU and its function in the video reconstruction process, while the payload carries the compressed signals of a video frame. The parameters dependency_id (DID), temporal_id (TID) and quality_id (QID) contained in the NALU header describe the scalability feature of the bitstream. Specifically, DID, TID and QID represent Coarse Grain Scalability (CGS), Temporal Scalability (TS) and Medium Grain Scalability (MGS) [20], respectively. The CGS feature facilitates the coarse adaption of video properties, such as the spatial resolution of the video, reconfiguring from QCIF to CIF, where the video can be encoded into a set of coarse enhanced sub-streams referred to as dependency-layers. The DID parameter represents the dependency-layer the current NALU belongs to. The decoding of a NALU with $DID > 0$ depends on the NALUs associated with $(DID-1)$, but with the same TID and QID values. Based on this dependency rule, the video quality may be readily reduced by removing the NALUs with a DID larger than a specific DID parameter. Similar dependency rules exist for the temporal scalability and MGS features. The dependency of the layers in the SVC stream is exemplified in Fig. 6.

3) Multiview Video Coding: Recently, the Joint Video Team (JVT) proposed MVC as an amendment to the H.264/AVC standard [9]. Apart from the classic techniques employed in single-view coding, multi-view video coding invokes the so-called inter-view correction technique by jointly processing the different views for the sake of reducing the bitrate. Hence, the first encoded view may be termed as the BL, while the remaining views may be treated as the ELs. The dependency

[2]Here the phrase x "depends on" y implies that the layer x will be discarded by the video decoder, if layer y is lost. The phrase x "is predicted from" y means that the layer x may still be usefully utilized by the video decoder, if layer y is lost. Hence the relationship "depends on" is stronger than "is predicted from".

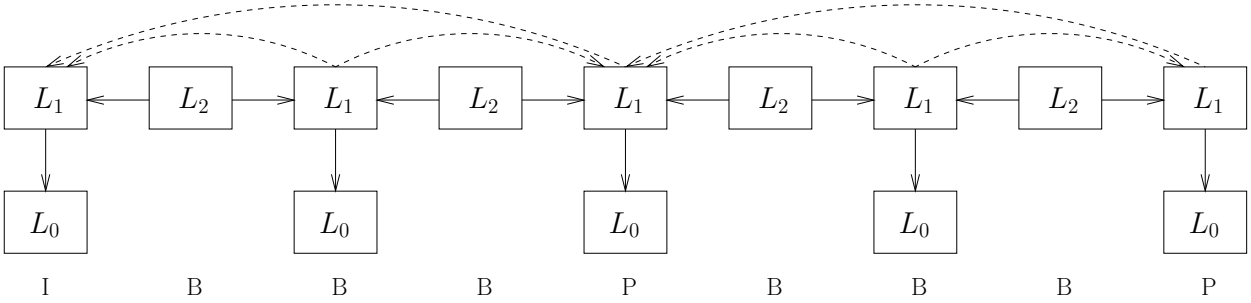


Figure 6: Dependencies with an GOP=4 SVC stream, where L_0, L_1 and L_2 represent the BL, spatial EL and temporal EL, respectively. "$\rightarrow$" indicates "depends on", while "$\dashrightarrow$" indicates "is predicted from".

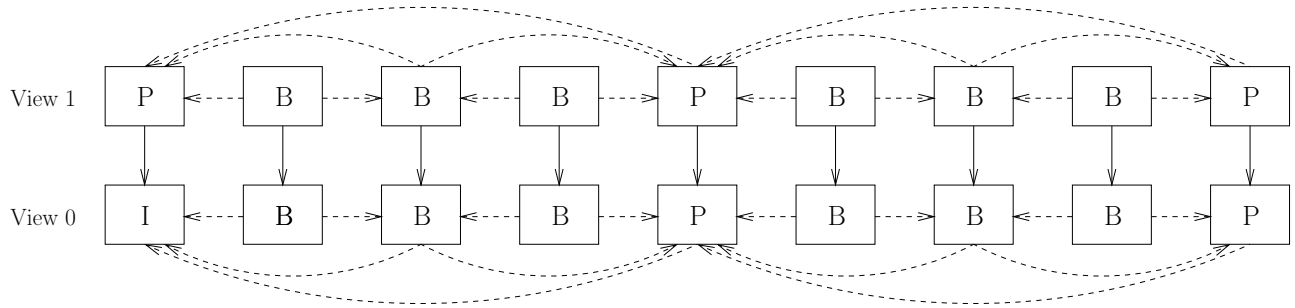


Figure 7: Dependencies with an GOP=4 MVC encoded stereoscopic video stream. "$\rightarrow$" indicates "depends on", while "$\dashrightarrow$" indicates "is predicted from".

of the layers in the MVC stream is exemplified in Fig. 7.

4) Others: Set-partitioning in hierarchical trees (SPIHT) [33], [34] was originally proposed as an image compression algorithm, which encodes the most important wavelet transform coefficients first and allows an increasingly refined reproduction of the original image. A multiview profile (MVP) [26] was developed by the moving picture expert group (MPEG)'s [5] video coding standard, where the left view and right view were encoded into a BL and an EL, respectively. Again, the emerging H.265 scheme will continue to include a scalability profile.

C. Why Layered Video Coding

The MPEG-2 based digital video systems were designed for broadcasting services over satellite and terrestrial channels, as well as for storage. Generally, these applications require fixed quality video representations [20]. However, these traditional video representations fail to meet the requirements of the various transmission scenarios in the era of the Internet and of mobile networks [35], [36], which have to support heterogeneous devices, operating under unreliable network conditions, bandwidth fluctuations etc. Layered video coding, such as SVC, emerged in order to support operation in these new transmission scenarios.

1) Heterogeneous Devices: Nowadays, contemporary video transmission and storage scenarios require diverse receiving devices with more flexible communication qualities [37], as exemplified by the mobile phones, tablet PCs, laptops and TV sets of Fig. 4. Satisfying these diverse requirements is facilitated by layered video compression, for video-streaming, teleconferencing, surveillance, broadcast and storage for diverse devices.

Given only the layer L_0 having a bitrate of 128 kbits per second (kbps), the corresponding layered video decoder of Fig. 4 reconstructs the video with a resolution of QCIF at 7.5 frame per second (FPS), while a CIF based video sequence at 30 FPS can be reconstructed with the aid of layers L_0, L_1 and L_2, which require bitrates of 128 kbps, 256 kbps and 512 kbps, respectively. If the TV screen of Fig. 4 is utilized by the user, all four layers $L_0 \sim L_3$ may also be streamed for achieving the highest video quality. In practice, the different video streaming scenarios of Fig. 4 require different bandwidth and hence achieve different visual quality. The users may rely on different video screens, such as those of mobile phones, tablet PCs, PC and TV screen, as seen in Fig. 4 for example.

2) Unreliable Networks: Internet-based communication relies on the Transmission Control Protocol (TCP)-Internet Protocol (IP), which aims for providing best-effort services for the application layer. However, TCP-IP is unable to guarantee the reliability of Internet links, since unpredictable network congestions, breakdown of routers, jitter, cache overflow etc. [38], [39] may occur. In comparison to wired links, wireless channels are more hostile due to both small-scale and large-scale fading [40]. In the scenario of transmitting layered video over unreliable networks [41], the BL and ELs may be assigned different protection strengths [2], channels [42], or different packet-dropping priority [43], [44] etc, for the sake of combating the network-induced impairments. For example, retransmissions may be used for the BL to limit the burden imposed on congested networks [43].

3) Bandwidth Fluctuations: In wired networks, bandwidth fluctuations are mainly caused by the effects of network con-

gestion. In wireless links, the bandwidth fluctuates for several reasons [35]. Firstly, the bandwidth may vary dramatically when a mobile terminal switches between different networks, e.g., from a wireless local area network (LAN) to a wireless wide area network (WAN). Secondly, the throughput of a wireless channel may be reduced due to interference imposed both by reflecting physical objects and by other wireless devices. Thirdly, the bandwidth of a wireless link may fluctuate due to the changing distance between the base station and the mobile terminal.

In traditional transmission of single layer video, the receiver may experience long delays due to the limited bandwidth, which is unacceptable for real-time video streaming. Hence layered video coding may be utilized for the sake of combating the bandwidth fluctuations, which allows us to dynamically adapt the bandwidth requirements [45]. We consider the scenario of Fig. 4 as an example. When the bandwidth is low, we may transmit L_0 and L_1 of the bitstream, resulting in a bandwidth requirement of 384 kbps, while the receiver can still reconstruct a CIF-sized video at 15 FPS. When the bandwidth becomes higher, L_0~L_3 of the bitstream may be transmitted, resulting in a bandwidth requirement of 1920 kbps. As a result, the receiver can reconstruct a HD video at 60 FPS.

III. Layered Video Communications

From the perspective of the source codec, there are two ways of correcting or concealing the transmission errors. Popular class of solutions relies on the so-called error-concealment/error-resilient techniques. Numerous error resilient techniques have been investigated [46]–[49], such as intra-MB refreshing, motion vector recovery, frame-copy, distributed source coding etc. Generally, these techniques aim for reducing the channel-effects imposed on the received video is, thereby improving the reconstructed video quality. For example, the authors of [47] aimed for reducing the error propagation effects using the so-called "peg" frames instead of the traditional intra-coded frames. Indeed, it is intuitively benificial to perform error concealment, since the source codecs have the most intricate knowledge of the source stream. However, the quality of video cannot be guaranteed in a vulnerable network environment, even when the error effects may be reduced significantly. The other class of correcting/concealing errors includes the techniques that are performed after/before the source encoding/decoding stage, which include unequal error protection [50], inter-layer FEC [2], priority-encoding based transmission [51] etc. The benefit of the these methods is that a high reconstructed video quality may be achieved in a vulnerable error-prone network environment for example by reducing the FEC coding rate. However, the worst-case network conditions have to be considered, when allocating error protection coding for achieving near-error-free real-time video, which seems wasteful of resources. In this treatise, we mainly consider the latter class of techniques.

Unequal error protection (UEP)was firstly proposed by Masnick and Wolf in [50], which allocates stronger FEC to the more important data, while dedicating weaker FEC to the less important video parameters. Since then numerous UEP techniques have proposed. A novel UEP modulation concept was investigated in [52] for the specific scenarios, where channel coding cannot be employed. Hence UEP was achieved by allocating different transmission power to individual bits according to their bit error sensitivity albeit in practice this remains a challenge. Additionally, the UEP capabilities of convolutional codes (CC) were studied in [53]. Furthermore, as a benefit of the outstanding performance of low-density parity-check (LDPC) codes, a number of UEP design methodologies [54]–[57] have been investigated using LDPC codes. The so-called UEP density evolution (UDE) technique of [54], [57] was proposed for transmission of video streams over binary erasure channels (BEC). The authors of [55] proposed a new family of UEP codes, based on LDPC component codes, where the component codes are decoded iteratively in multiple stages, while the order of decoding and the choice of the LDPC component codes jointly determine the level of error protection. A practical UEP scheme using LDPC codes was proposed in [56], where the high-significance bits were more strongly protected than low-significance bits.

As mentioned in Section II-B, when the BL is corrupted or lost due to channel impairments, the ELs must also be dropped by the video decoder, even if they are perfectly received. Moreover, the less important layers have lower priority and hence may be dropped in the transmission scenario of network congestion or buffer overflow [58]. Hence, it is intuitive to perform UEP for both the BL and the ELs for the sake of improving the system's performance. However, most of the above UEP studies considered artificially generated signals of unequal significance, rather than realistic video signals. Naturally, the significance differentiation of practical video signals is more challenging. In compressed video streams, as in layered video coding, different bits may have different significance. Therefore, UEP became the most important technique in the context of layered video communication invoked for achieving an improved reconstructed video quality. Generally, UEP may be performed by carefully and judiciously allocating resources such as the coding rate, the transmit power, modulation mode, etc. Nonetheless, a number of contributions have been made in the field of layered video communications. Generally, these scalable video streaming techniques [1], [2], [30], [41], [42], [59]–[103] may be classified into the following four categories:

- Transceivers based on UEP schemes: as in [42], [68], [72], [86], [90], [95]–[97], allocate an unequal amount of transmit power to the different layers. These schemes are employed in the physical layer of the Open Systems Interconnection (OSI) model.
- FEC based UEP Schemes
  - Packet-level Schemes: as in [1], [59], [60], [62], [63], [67], [69], [71], [74], [78], [85], [87]–[89], [92], [93], [98], [99], [101], mitigate the packet-loss events, as exemplified by the packets lost in Internet-routers [39]. These schemes are typically hard-decoded at the higher layers.
  - Bit-level Schemes: as in [2], [30], [65], [66], [77], [79], [100], [102], are devoted to eliminating bit-

errors in wireless scenarios [35], [104]. These schemes may be soft- or hard-decoded at the lower layers of the OSI stack.

- Cross-layer operation aided schemes: as in [41], [61], [70], [73], [75], [76], [81], [91], [94], [103], are typically invoked for optimizing the scalable video streaming systems by considering multiple signal processing stages, such as the source-compression, FEC-encoding, modulation etc. These schemes tend to collaborate across multiple layers of the OSI stack.

Below, we will introduce these four categories in the Sections ranging from III-A to III-D.

A. Transceivers Based UEP

The major contributions on UEP aided layered video streaming techniques are summarized in Table I. In [68], [72], [80], [105], UEP schemes conceived for video transmission using hierarchical quadrature amplitude modulation (HQAM) were investigated, which considered the unequal importance of the different layers in SVC, as well as the unequal importance of both the intra-coded frames (I-frame) and of the predicted frames (P-frame). Specifically, the video bits of different importance were mapped to the different-integrity bits of the modulation constellation points of HQAM. An adaptive channel selection (ACS) based layered video transmission scheme was proposed for transmission over multiple input multiple output (MIMO) schemes in [42], which periodically switched each bit stream among multiple antennas. Specifically, the ordering of each sub-channel's signal-to-noise ratio (SNR) was exploited as partial channel quality information (CQI) at the receiver. The bitstream was carefully mapped by exploiting the SNR-based CQI of the channels. Essentially, the higher-priority layer's bitstream has to be mapped to higher-SNR channel by the proposed algorithm. A HQAM based UEP scheme was proposed in [86] for H.264/AVC video transmission over frequency selective fading channels. For the sake of preventing the more important data to be mapped onto sub-carriers subjected to deep fading, an orthogonal frequency-division multiplexing (OFDM) sub-carrier classification strategy relying on a pair of SNR thresholds was proposed. Cooperative multicasting of scalable video was investigated in [97]. The so-called opportunistic cooperative multicast (OppCM) and coded cooperative multicast (CodedCM) were proposed, where opportunistic listening, conditional demodulation and multi-resolution modulation were utilized to enhancing the system's performance.

A cooperative MIMO framework was proposed in [90] for scalable video transmission, which relies on a sophisticated power control strategy invoked for controlling the on/off mode of the relays and the specific power allocation among them. In [95], UEP was achieved by striking a tradeoff between the achievable transmission integrity and data rates, which were controlled both by the FEC and the MIMO mode selection for the sake of minimizing the average distortion. A scalable resource allocation framework was proposed in [96] for streaming scalable videos over multiuser MIMO-OFDM networks, where the attainable multidimensional diversity gains were investigated. Specifically, a scalable resource-allocation framework was proposed for achieving different quality of service objectives for different scalable video layers. A beneficial scheme was designed [96] for guaranteeing that each user was entitled to maintain a specific MAXMIN fairness for ensuring that their BL video packets are indeed received.

Generally speaking, these transceiver-based UEP techniques are mainly employed in the physical layer, as shown in Fig. 8, since they rely on adapting controlling the modulation mode or the channel selection for the sake of improving the attainable performance.

B. Packet-Level FEC

The major contributions on packet-level FEC based layered video streaming techniques are summarized in Table II. In [106], the source of unequal importance was firstly transformed to and encoded with the aid of multiple descriptions using Reed-Solomon codes. Then the FEC coding rate assignment of different layers was solved by the classic Lagrange Multiplier method for the sake of robust transmission. In [59], the authors proposed a sophisticated framework for optimizing the rate-distortion relationship [109] in the context of video communications over error-prone packet-switched networks, which was based on the principle of layered coding relying on transport prioritization. Su *et al.* in [62] aimed for finding an optimal FEC assignment scheme for scalable video transmission over bursty channels invoking packet-loss rate feedback. An iterative algorithm was derived for calculating a new packet-loss probability function conditioned on the past loss rates. Diverse UEP algorithms were designed in [60], [63], [64], [110] for transmitting 3D SPIHT coded images and videos over networks imposing packet-loss events, where Reed-Solomon (RS) codes were utilized as FEC codes. An UEP scheme was conceived for object-based video communications in [67] for achieving the best attainable system performance under specific bitrate and delay constraints in an error-prone network environment. In [69], UEP was allocated to the different frames (I- or P-frame) of a GOP, and in each frame, unequal protection was allocated to the progressive bitstream of scalable video for the sake of providing a graceful degradation of the video quality, as the packet loss rate increased. A genetic algorithm (GA) was utilized to for efficiently identifying the UEP patterns, which was inefficient with the aid of conventional methods owing to the excessive search-space. A packetization scheme was proposed in [71], which efficiently combated packet loss events by combining UEP, retransmission and GOP-level interleaving. Intra-GOP rate allocation was invoked for minimizing the distortion of individual GOPs, while inter-GOP rate allocation was proposed for reducing the video quality fluctuations by adaptively allocating bandwidth according to both the video signal characteristics and to the buffer-fullness. An adaptive UEP scheme was proposed in [74] for robust and scalable wireless video streaming. By jointly exploiting the temporal inter-frame dependency and the quality dependency between the scalable layers, the proposed scheme adaptively assigned unequal-protection FEC codes to the packets of each layer and

Table I: Major contributions on transceivers based UEP aided layered video streaming.

Year	Author(s)	Contribution
2006	Chang *et al.* [68]	proposed HQAM based UEP for H.264/AVC coded video, which considered the different importance of the I-frames and P-frames.
	Ghandi and Ghanbari [72]	benchmarked SVC against the partition mode of H.264/AVC video codec, where UEP is achieved by employing HQAM.
2007	Song and Chen [42]	proposed adaptive channel selection (ACS) based layered video streaming over MIMO system, which periodically switched each bit stream among multiple antennas and allocated the higher-priority layer's bitstream to higher-SNR channels.
2010	Li *et al.* [86]	proposed an OFDM sub-carrier classification strategy for a HQAM based UEP scheme, thereby avoiding the more important data to be mapped onto sub-carriers in deep fading.
	Xiao *et al.* [90]	proposed a cooperative MIMO framework for scalable video communications, which employs a sophisticated power control strategy for controlling the on/off mode of relays and the specific power allocation among them.
2013	Kim *et al.* [95]	considered the tradeoff between the achievable transmission integrity and data rates, which were controlled by the FEC and the MIMO mode selection to minimize the average distortion.
	Li *et al.* [96]	proposed a scalable resource allocation framework for streaming scalable videos over multiuser MIMO-OFDM networks, where the achievable multidimensional diversity gains were investigated.
	Wang and Liao [97]	proposed opportunistic cooperative multicast (OppCM) and coded cooperative multicast (CodedCM) for scalable video communications.

adjusted the transmission rate by dynamically selecting the number of layers to transmit according to both the available network bandwidth and to the packet loss rate. A novel UEP method using RS codes was proposed in [78] 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 RS coding rates were assigned based on LW-EZEP for minimizing the expected video distortion. A two-dimensional (2D) layered multiple description coding [107] arrangement was proposed for scalable video transmission over unreliable networks, which allocates multiple description based sub-bitstreams of a scalable 2D bitstream to two different network paths exhibiting unequal packet loss rates. Furthermore, the video distortion was minimized, conditioned both on the given total rate budget and on the packet loss probabilities.

In recent years, cross-layer operation aided scalable video streaming designed for error-prone channels was investigated in [87], 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. The authors of [88] studied an UEP scheme using Luby Transform (LT) codes [108], [111] for recovering the video packets lost owing to network congestions. The multistream UEP (M-UEP) concept was proposed in [92], which allocated separate streams to separate sets of packets for maintaining their independence. The concept of permuted systematic RS codes defined in [92] was employed for beneficially dispensing the message symbols across the packets, which interleaved the source symbols with the redundancy symbols in order to form the codewords. M-UEP improved the efficiency of classic UEP by ensuring that all received source symbols are decoded. The priority encoding transmission (PET) [51] protection philosophy was introduced in [98] for streaming scalable video streams over erasure channels, where a small number of retransmissions were allowed. In principle, the choice of the most appropriate protection depended not only on the importance of each stream element, but also on the expected channel behavior. By formulating a collection of hypotheses concerning a specific stream's own behavior in future transmissions, the limited-retransmission aided PET (LR-PET) regime was capable of effectively constructing channel codes spanning multiple transmission slots and thus offered a better protection than the original PET of [51]. Based on packet-level transmission distortion modeling, in [99] the significance of each video packet was estimated in terms of the reconstructed video quality, which defined the priority level of each packet. UEP was then allocated to the video packets according to both priority levels as well as to the prevalent channel conditions. The proposed RD-based UEP resource allocation problem of [99] was formulated as a constrained nonlinear optimization problem. Then an algorithm based on Particle Swarm Optimization (PSO) was developed for solving the optimal resource allocation problem. In [101], UEP was conceived for packets containing scene-transition frames, which have to be better protected for the sake of attaining an improved video quality. A sophisticated FEC code allocation strategy was adopted based on the minimization of the end-to-end distortion, assuming that error concealment was adopted at the decoder. Two different FEC allocation strategies were proposed [101] for the Block of Packets (BOP) structure, namely an iterative modified hill climbing approach and a reduced-complexity heuristic approach.

In the above UEP schemes, 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,

Table II: Major contributions on packet-level FEC based layered video streaming.

Year	Author(s)	Contribution
1999	Puri *et al.* [106]	transformed and encoded scalable sources into multiple descriptions using the family of Reed-Solomon codes for robust transmission.
2001	Gallant *et al.* [59]	proposed a framework for optimizing the rate-distortion relationship of video communications over error-prone packet-switched networks based on the principle of layered coding relying on with transport prioritization.
2002	Thornton *et al.* [60]	designed UEP algorithms for transmitting 3D-SPIHT coded videos over networks subject to packet-loss events, where packet-level FEC codes were employed.
2003	Su *et al.* [62]	found an optimal FEC assignment scheme for scalable video transmission over channels inflicting bursty errors and relying on loss rate feedback.
	Kim *et al.* [63]	proposed a bit-plane-wise UEP algorithm for 3D-SPIHT coded progressive bitstream transmission over lossy networks.
2005	Wang *et al.* [67]	conceived an UEP scheme for object-based video communications in order to achieve the best attainable system performance under specific bitrate- and delay-constraints in an error-prone network.
2006	Fang and Chau [69]	performed UEP of the different frames within a GOP and of the different layers of each frame.
	Gan *et al.* [71]	combated packet loss events by combining UEP, retransmission and GOP-level interleaving. Intra-GOP and inter-GOP rate allocation were also investigated.
2007	Shi *et al.* [74]	proposed an adaptive UEP scheme for robust scalable wireless video streaming, which adaptively assigned different rate FEC codes to each layer and adjusted the transmission rate by dynamically selecting the number of layers to transmit.
2008	Ha and Yim [78]	proposed a technique referred to as layer-weighted expected zone of error propagation (LW-EZEP) for quantifying the error propagation effects imposed by packet loss events.
2009	Xiang *et al.* [107]	allocated multiple description sub-bitstreams of a scalable 2D bitstream to two different network paths exhibiting unequal packet loss rates and minimized the end-to-end distortion.
	Sejdinović *et al.* [82]	proposed the so-called packet-level expanded window fountain (EWF) codes for UEP.
	Vukobratović *et al.* [85]	applied the EWF codes for scalable video multicast over networks inflicting packet loss events, where ELs conveyed parity information protecting the more important BL.
2010	Maani and Katsaggelos [87]	proposed cross-layer operation aided scalable video streaming, which estimated the expected video distortion according to the prevalent channel conditions.
	Ahmad *et al.* [88]	developed Luby Transform (LT) [108] coded UEP for combating the packet loss events imposed by the network.
	Dumitrescu *et al.* [92]	proposed the multistream UEP (M-UEP) philosophy, which utilized permuted systematic Reed-Solomon (RS) codes for enhancing the distribution of message symbols amongst the packets.
	Nguyen *et al.* [89]	designed hierarchical network coding (HNC), where the less important bits are used for protecting the more important bits.
2011	Halloush and Radha [93]	proposed the so-called Multi-Generation Mixing (MGM) technique, which allows us to recover packets using data encoded into other packets.
	Hellge *et al.* [1]	proposed the layer-aware FEC (LA-FEC) philosophy using a Raptor codec for video transmission over the BEC.
2013	Xiong *et al.* [98]	advocated a so-called priority encoding transmission (PET) [51] protection for streaming scalable video over erasure channels, where a small number of retransmissions were allowed.
	Zhang *et al.* [99]	estimated the importance of each video packet according to their impact on the reconstructed video quality. UEP was then allocated to different video packets according to their priority levels and the dynamic channel conditions.
2014	Midya *et al.* [101]	identified the more important packets containing scene-transition frames and invoked UEP for the sake of improved video quality.

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 [85] applied the packet-level expanded window fountain (EWF) codes of [82] for scalable video multicast over networks inflicting packet loss events, where the ELs conveyed parity information protecting the more important BL. By contrast, hierarchical network coding (HNC) was proposed in [89], where the encoded ELs carry some of the information of the BL. Hence, when only a low number of coded packets is received, the most important data can be recovered with a high probability. As a further advance, Multi-Generation Mixing (MGM) was developed in [93] for combating the deleterious effects of packet-loss events. When an insufficient number of packets is received for a specific chunk of packets, it is still possible to recover the chunk using the data embedded into other chunks. The packet-level layer-aware FEC (LA-FEC) philosophy using a hard-decoded Raptor code was designed for scalable video transmission over the BECs in [1], [112]. 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. Similar techniques were also investigated in [54], [113], [114]. The common characteristic within the contributions of [1], [85], [89], [93] is that they all facilitate the enhanced protection of the BL using the information of the ELs. In the rest of the paper, we refer to these schemes as LA-FEC.

Generally speaking, these packet-level FEC schemes are typically employed in the higher layers of the OSI protocol stack, as shown in Fig. 8, since they are designed for combating the packet loss events.

C. Bit-Level FEC

The major contributions on bit-level FEC based layered video streaming techniques are summarized in Table III. These bit-level schemes [2], [30], [65], [66], [77], [79], [100], [102], [115], [116] tended to employ physical layer FEC codes [117] and performed soft decoding [118] for wireless video communications.

The authors of [65] minimized the mean distortion by non-uniformly distributing the redundancy imposed by the turbo code between the successive video frames, where the H.263 video codec was employed. In [66], an UEP scheme using a turbo transceiver was optimized for wireless video telephony. LDPC code based UEP was investigated in [57]. The UEP performance of data-partitioned [9] H.264/AVC video streaming systems using recursive systematic convolutional (RSC) codes was evaluated in [30], while turbo coded modulation [119] based UEP was investigated in [77], where both the cutoff rates and the channel capacity of each of the UEP levels was determined. The authors of [79] considered the unequal importance of both the video-frames in a GOP and the significance of the diverse MBs 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 [79], which limits the attainable system performance. The authors of [100] show that the side information (SI) values within different positions of the Wyner-Ziv (WZ) frames have different error probability. Hence UEP of these non-uniformly distributed SI values was employed for the sake of reducing the required bitrate in the application of distributed video coding (DVC) [120], [121].

Similar to the packet-level schemes of [1], [85], [89], [93], where the information of the ELs was utilized for protecting the BL, the authors of [25] developed bit-level inter-layer coded FEC (IL-FEC) arrangements for layered video telephony over wireless fading channels in [2], [115], [116] relying on soft-decoded RSC, turbo and self-concatenated convolutional codes, respectively, 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 [2] was also combined with the UEP philosophy for the sake of further improving the attainable system performance. In [2], 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. In [102], the authors proposed 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, they found 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. In a nutshell, the authors of [102] focused their effects on the optimization of bit-level IL-FEC encoded scalable video communications over wireless fading channels. The UEP based low-density parity-check codes (UEP-LDPC) of [54] may also be used for bit-level protection, which may rely on both soft-decoding and hard-decoding.

These bit-level FEC schemes may be further divided into soft decoded and hard-decoded categories. The soft-decoded bit-level FEC schemes are typically employed in the physical layer, since they are capable of combating the effects of the wireless channel. By contrast, as shown in Fig. 8, the hard-decoded schemes may be invoked in the data link or network layer for mitigating the bit-erasure effects.

D. Cross Layer

The major contributions on cross-layer operation aided layered video streaming techniques are summarized in Table IV. The authors of [61] proposed an adaptive cross-layer protection strategy for enhancing the robustness of scalable

Table III: Major contributions on bit-level FEC based layered video streaming.

Year	Author(s)	Contribution
2004	Marx and Farah [65]	minimized the mean distortion by non-uniformly distributing the redundancy imposed by a turbo code between the successive video frames.
2005	Ng *et al.* [66]	optimized an UEP scheme using a turbo transceiver for wireless video telephony.
2007	Rahnavard *et al.* [57]	investigated an UEP scheme using partially regular LDPC codes.
2008	Aydinlik and Salehi [77]	investigated UEP based turbo coded modulation, where both the channel capacity and the cutoff rates of the UEP levels were determined.
	Chang *et al.* [79]	considered the unequal importance of the frames in a , as well as that of the macroblocks in a video frame.
2011	Nasruminallah and Hanzo [30]	evaluated the UEP performance of data-partitioned H.264/AVC video streaming systems using RSC codes.
2013	Huo *et al.* [2]	developed a bit-level inter-layer coded FEC (IL-FEC) scheme for layered wireless video steaming relying on a soft-decoded FEC code, where the systematic bits of the BL are implanted into the ELs at the transmitter.
	Micallef *et al.* [100]	conceived UEP for side information (SI) values at different position of the Wyner-Ziv (WZ) frames for the sake of reducing the bitrate in DVC.
2014	Huo *et al.* [102]	proposed 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.

Table IV: Major contributions on cross-layer operation based layered video streaming.

Year	Author(s)	Contribution
2003	Schaar *et al.* [61]	proposed a cross-layer protection for robust scalable video transmission by striking tradeoff amongst the attainable throughput, reliability and delay.
2004	Zhang *et al.* [41]	judiciously allocated the resource between the source and channel encoders based on either the minimum-distortion or minimum-power consumption criterion.
2006	Qu *et al.* [73]	adaptively adjusted the rate of the video encoder and the channel encoder to maximize the video quality delivered based upon both application-layer video motion estimates and link-layer channel estimates.
	Barmada *et al.* [70]	combined turbo coding and HQAM to provide a high protection for the BL under the constraint of the affordable channel coding redundancy.
2007	Ha *et al.* [76]	optimized the source and channel coding rates for minimizing the overall distortion.
	Huusko *et al.* [75]	proposed a cross-layer operation method and a protocol architecture for the sake of transmitting the required control information and for optimizing the overall multimedia transmission quality over wireless and wired IP networks.
2009	Shan *et al.* [81]	amalgamated a two-stage FEC scheme with an enhanced link-layer protocol conceived for multimedia transmission over wireless LANs. Both packet-level FEC and bit-level FEC were employed.
	Jubran *et al.* [83]	estimated the video distortion for specific channel conditions, which was then exploited for optimally selecting both the application layer parameters and the physical layer parameters.
2010	Zhang *et al.* [91]	investigated the efficient streaming of scalable video from a base station to multiple clients over a shared fading wireless network by considering both the application layer information and the wireless channel conditions.
2012	Khalek *et al.* [94]	proposed an APP/MAC/PHY cross-layer architecture for optimizing the perceptual quality of delay-constrained scalable video transmission.
2014	Cicalo and Tralli [103]	designed a cross-layer optimization framework for scalable video delivery over OFDMA wireless networks, which jointly addressed both rate adaptation and resource allocation.

video transmission by striking tradeoff amongst the attainable throughput, reliability and delay by taking into account the near-instantaneous channel conditions and application requirements. By carefully considering both the wireless channel conditions and the SVC characteristics, the authors of [41] judiciously allocated the resources between the source and channel encoders based either on the minimum-distortion or the minimum-power consumption criterion. By exploiting the knowledge of the source and the channel conditions, the authors of [73] adaptively adjusted the operating parameters of both the video source encoder and of the FEC channel encoder for maximizing the video quality delivered based upon both the application-layer video motion estimates and the link-layer channel estimates. The authors of [70] combined turbo coding and HQAM to provide a high protection for the BL under the constraint of the maximum affordable channel coding redundancy. Both the source and channel coding rates were optimized in [76] for the sake of minimizing the overall distortion. An efficient cross-layer operation aided protocol architecture was proposed in [75] for the sake of transmitting the required control information and for optimizing the multimedia transmission quality both over wireless and wired Internet Protocol (IP) networks. A two-stage FEC scheme was investigated in [81] in conjunction with an enhanced link-layer protocol specifically designed for multimedia transmission over wireless LANs. At the application layer, the stage-one packet-level FEC scheme protected the packets against packet losses due to congestion and route disruption. The stage-two bit-level FEC was then added to both application packets and stage-one FEC coded packets to recover bit errors imposed by the link layer. Then at the link layer, a combined header cyclic redundancy check (CRC)/FEC scheme was used to enhance the attainable protection and to act in unison with the two-stage FEC scheme. The proposed scheme thus provided joint protection across the entire protocol stack. The authors of [83] estimated the distortion of the received video and proposed different error concealment schemes for diverse channel conditions. This estimated end-to-end video distortion was then utilized for optimally selecting the application layer parameters, such as the quantization parameter (QP), the GOP size and the physical layer parameters, such as the coding rate, the symbol constellation in a constrained-bandwidth framework.

The authors of [91] investigated the streaming of scalable video from a base station to multiple clients over a shared fading wireless network by considering both the available application layer information as well as the wireless channel conditions. An APP/MAC/PHY cross-layer architecture was proposed in [94] that optimizes the perceptual quality of delay-constrained scalable video transmission. Specifically, an online Quality of Service-to-Quality of Experience (QoS-to-QoE) mapping technique was employed for quantifying the effects of packet loss events endured by each video layer using the Acknowledgement (ACK) history and perceptual metrics. At the PHY layer, the authors developed a link adaptation technique relying on the QoS-to-QoE mapping for providing perceptually-optimized UEP. At the APP layer, the source rate was adapted by carefully selecting the specific set of temporal- and quality-layers to be transmitted based on a combination of the prevalent channel statistics, on the source rates and on the playback buffer-fullness. A cross-layer optimization framework was proposed in [103] for scalable video delivery over orthogonal frequency division multiple access (OFDMA) wireless networks, which jointly addressed rate adaptation and resource allocation with the aim of maximizing the sum of the achievable rates, while minimizing the distortion difference among multiple video streams. These cross-layer operation based schemes may over-arch the entire OSI stack, as shown in Fig. 8.

E. Summary

Below, Section III-E1 summarizes the contributions reviewed above, followed by the illustration of the main focus of this tutorial.

1) Resources-Solutions-Objective: The main objective of layered video streaming schemes is to minimize the video distortion, which is indicated by the innermost kernel of Fig. 8. Furthermore, the available resources and the four solution categories are also displayed in Fig. 8. Generally, a specific solution may change the configuration or parameters of several resources for the sake of approaching the minimum video distortion. The available resources are listed as follows:

- Carrier: represents an electro-magnetic wave of fixed amplitude and frequency that is modulated in amplitude, frequency, or phase in order to carry a modulating signal with the aid of radio-frequency (RF) transmission. Multiple carriers may exist in a wireless communication system, which may experience different fading effects. Hence a less faded carrier may be selected for conveying the more important BL as exemplified in [86].
- Bandwidth: In wireless communications, the bandwidth represents the frequency range occupied by a carrier wave. For wireless transmission of the BL a wider bandwidth may be allocated for the sake of improving its quality. In wired networks, the bandwidth predetermines the achievable data transfer rate. In congested networks, adaptive bandwidth allocation may be beneficial for layered video streaming, as exemplified in [71].
- Channel: Layered video transmission may benefit from the diversity gain of MIMO systems, where different transmitter and receiver antenna pairs experience different channel fading effects. Hence it is intuitive to optimize the channel assignment algorithm for transmitting different layers, as exemplified in [42].
- Power: In wireless streaming of layered video, the employment of adaptive power control may be beneficial, which increases/reduces the transmit power depending on the importance of the data, as exemplified in [90]. Furthermore, the hierarchical constellation point mapping of [72] may be interpreted as unequal power allocation.
- FEC Redundancy: FEC encodes a packet by imposing redundant bits for the sake of providing an error correction capability. Generally, more redundancy results in lower FEC coding rates and stronger protection. Hence different importance layers may be encoded differently

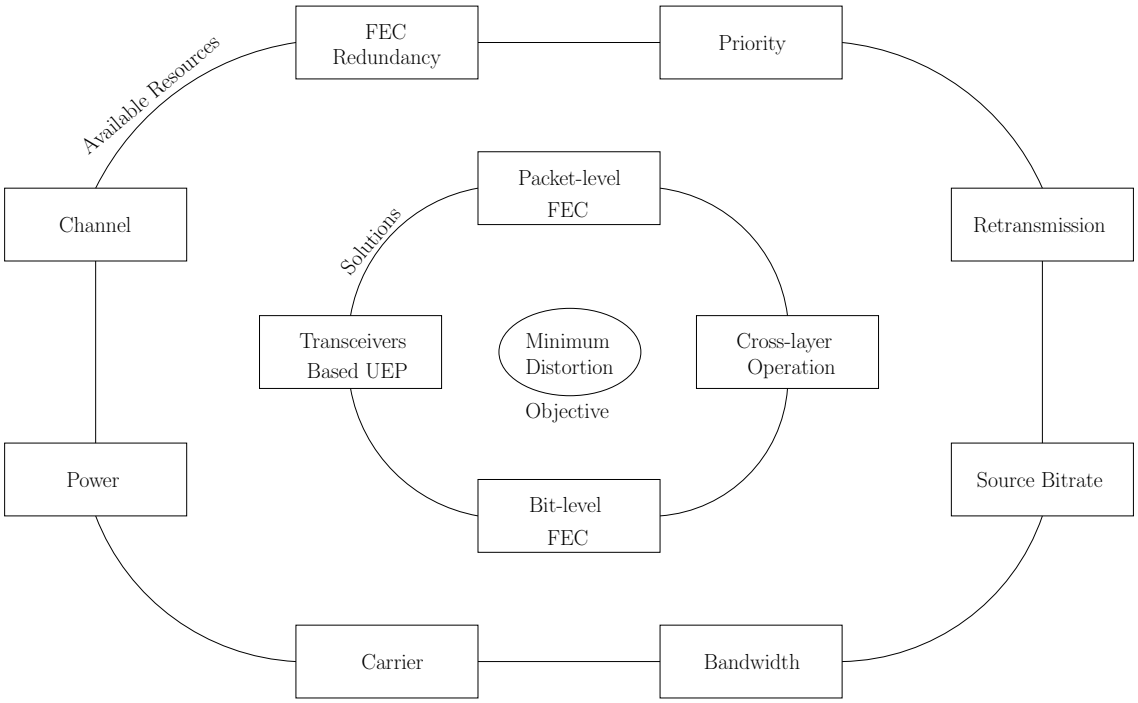


Figure 8: The objective, solutions and available resources in existing layered video streaming schemes.

for the sake of improving the overall video quality, as exemplified in [2].

- Retransmission: Retransmissions are widely utilized in congested networks for improving the QoS, provided that the delay imposed may be acceptable in non-interactive broadcast-scenarios. Layers of different importance may be retransmitted for the sake of improving the performance, as exemplified in [71].
- Priority: The priority classes affect both the routing, as well as the congestion control of a packet. Hence in vulnerable networks it is intuitive to assign a higher priority for the BL than for the ELs, as exemplified in [43], [44].
- Source Bitrate: In vulnerable and throughput-limited networks, streaming of a high-bitrate source bitstream may result in unacceptable reconstructed video quality at the receiver owing to the preponderance of errors. Hence we may have to dynamically adapt the bitrate allocation. For example, the source bitrate may be reduced for the sake of increasing the FEC coding rate within a fixed total bandwidth, as exemplified in [41].

2) Our Focus: According to the state-of-the-art review provided in the Sections ranging from III-A to III-D, the information of the ELs was exploited in the LA-FEC and IL-FEC based UEP schemes of [1], [2], [85], [89], [93], [102] for the sake of protecting the BL, as seen in Fig. 9. Hence, they were capable of substantially improving the system's performance. For the sake of inspiring further research on LA- / IL-FEC solutions, we will review the family of packet-level LA-FEC and bit-level IL-FEC solutions in the rest of this tutorial, as indicated in Fig. 9 in bold fonts. Moreover, there is a paucity of research on the optimization of soft-decoded bit-level IL-

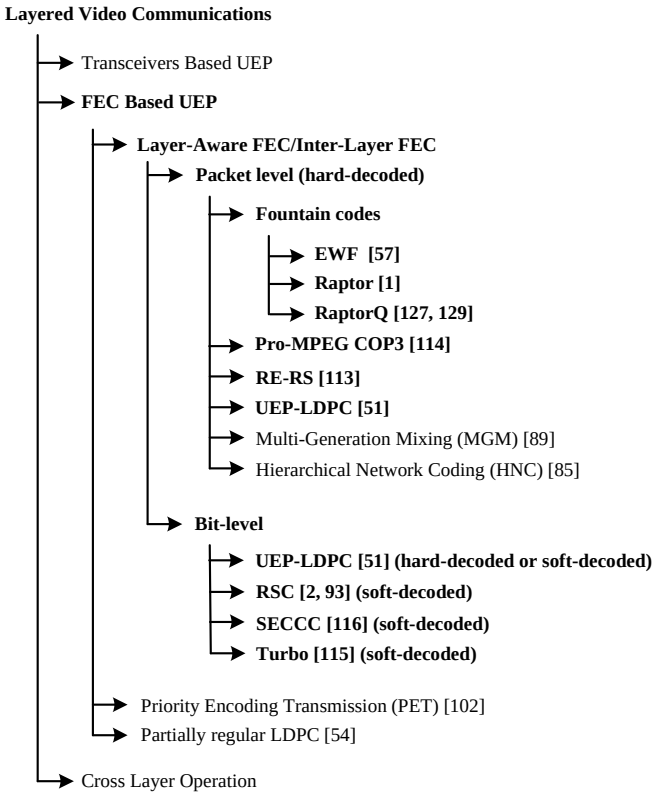


Figure 9: The schemes of [1], [2], [51], [54], [57], [85], [89], [93], [102], [113]–[116], [122] employ the information of the ELs for protecting the BL. The techniques considered in this treatise are shown in bold.

FEC based schemes. Since the soft-decoded FEC codes are more beneficial than their hard-decoded counterparts, we will mainly focus our attention on the class of soft-decoded IL-FEC schemes.

IV. Packet-Level Protection Using: Layer-Aware FEC

Layer-Aware FEC schemes are characterized by encoding BL and EL separately while allowing a combined decoding of BL and EL FEC data on the receiver side. This is achieved by incorporating the BL source data in the FEC generation of the EL so that EL data also protects BL data. The key for an efficient design of LA-FEC is the integration in today's state-of-the-art FEC codes while not deteriorating the original coding performance and functionalities of those highly optimized FEC codes. This section reviews the integration of LA-FEC in today's packet-level FEC and both, packet- and bit-level, algorithms.

A. Layer-Aware FEC Using Expanded Window Fountain (EWF) Codes

Fountain codes [108], [117], [123] provide a universal solution for combating packet loss events, when multicasting data over unreliable networks [124], since they are capable of generating an arbitrary number of encoded packets given the input source packets. LT codes [108] were the first practical fountain codes. LT codes are equal error protection (EEP) FEC codes, since they treat all source symbols as equally important. Hence as a further development, fountain codes capable of providing UEP have emerged in [82], [85], [125]. The family of EWF codes was proposed in [82] as a class of UEP fountain codes, which may be viewed as a generalized version of standard LT codes.

1) LT Encoding: The LT decoding process is detailed in [108]. Suffice to say that LT codes are capable of may recovering the transmitted source packets with a high probability, when the number of received packets is sufficiently higher than the number of input packets. The LT encoding process is listed as follows:

1) Divide the source sequence S into K blocks of equal length.
2) A random degree d is generated according to a specific degree distribution, where d is defined as the number of source-packets combined by the modulo-2 operation to generate an output packet, for example the so-called robust soliton distribution was proposed in [108], which determines the attainable performance of the LT codes.
3) For generating a new LT-coded packet, d of the K blocks are randomly selected. Then the binary exclusive OR (XOR) operation is performed on these d blocks, generating the new LT-encoded packet.
4) Repeat the Steps 2 and 3, until a sufficiently high number of packets was generated for frame S.

2) EWF Codes: The EWF encoding process is illustrated in Fig. 10, where we consider a total of k consecutive source symbols. As shown in Fig. 10, the r expanded windows

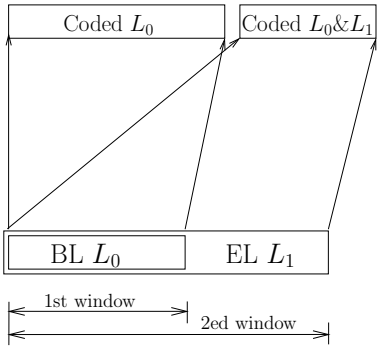


Figure 11: Example of expanded window fountain codes invoked for scalable video streaming.

defined over the source symbols generate r classes of different-rate protection. Specifically, the symbols falling in the index range of $1 \sim k_1$,$\cdots$, $(k_{i-1}+1) \sim k_i$, $\cdots$,$(k_{r-1}+1) \sim k_r$ define the 1^{st},$\cdots$, i^{th}, $\cdots$,r^{th} windows, which represent the importance order of $1^{st} > \cdots > i^{th} > \cdots > r^{th}$. A EWF code may be defined by the set:

$$\mathcal{F}(\Pi, \Gamma, \Omega^{(1)}, \cdots, \Omega^{(j)}), \tag{2}$$

where the protection-classes are defined by the generating polynomial

$$\Pi(x) = \sum_{i=1}^{r} \frac{s_i}{k} x^i, \tag{3}$$

with $s_i = k_i - k_{i-1}$ indicating the number of symbols within the i^{th} protection class. The encoding process of this EWF code is formulated as follows [82], [85]:

1) Randomly choose a window from the window-selection distribution of $\Gamma(x) = \sum_{i=1}^{r} \Gamma_i x^i$, where Γ_i is the probability of picking the i^{th} window.
2) Encode the symbols of the i^{th} window using the LT code $\Omega^{(i)}(x) = \sum_{l=1}^{i} \Omega_l^{(i)} x^i$, generating the output symbols of the i^{th} protection-class as exemplified in Fig. 10.
3) Repeat the above process, until a sufficient high number of symbols was generated.

3) Layered Video Multicast Using EWF Codes: Based on the above discussions, we note that the EWF code defined by the set $\mathcal{F}(\Pi, \Gamma, \Omega^{(1)}, \cdots, \Omega^{(j)})$ constitutes a fountain code, which assigns an output symbol to the i^{th} window using the probability distribution Γ_i. Then this protection class is encoded using a standard LT code based on the distribution $\Omega^{(i)}(x)$, which treats the source symbols within the the i^{th} protection class as of equal importance. Hence the EWF code may be considered as a specific class of UEP LT codes. The classic belief propagation (BP) [126] algorithm or alternatively low-complexity XOR operations [82], [108] may then be applied for decoding both the conventional LT codes and the EWF codes.

From Section IV-A2, we observe that the i^{th} window is protected by the $(i+1)^{st}$,$\cdots$,r^{th} windows. This property enables UEP for the different layers of scalable video. More specifically, let us consider the scenario of two layers, namely

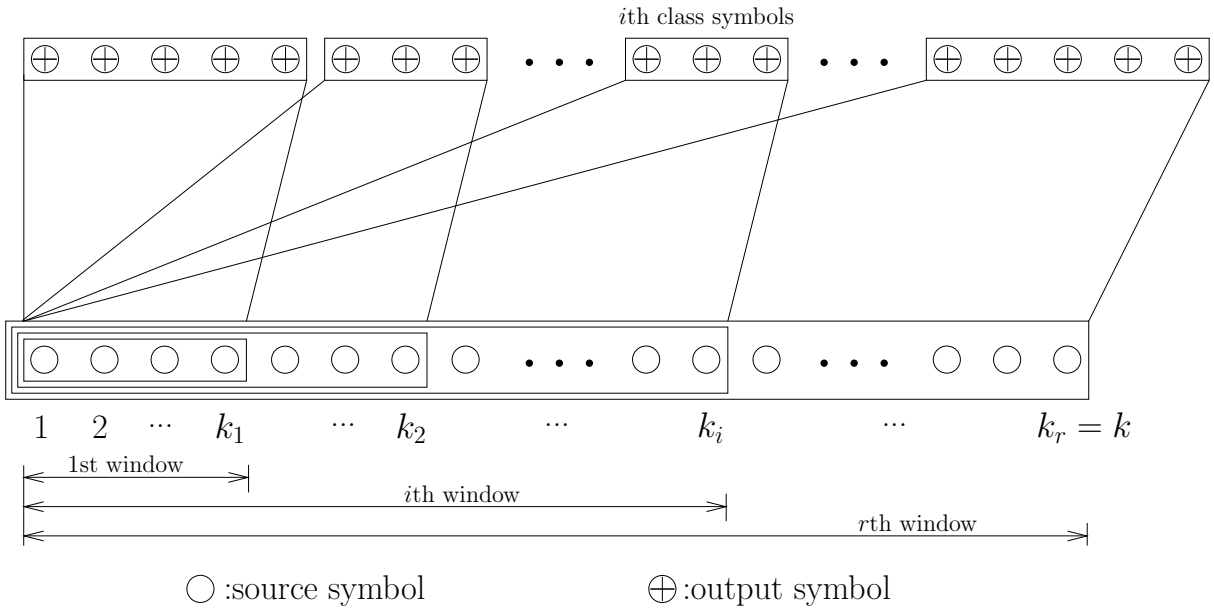


Figure 10: LT code with Expanded window fountain codes [82].

BL L_0 and an EL L_1. We may construct the EWF codes with two window as displayed in Fig. 11. The first window will be allocated to the BL L_0, while the second window will contain both BL L_0 and EL L_1. Hence, the BL L_0 may be protected by the encoded results of both 1^{st} window and 2^{ed} window, as illustrated in Fig. 11. Since the encoding symbols of an LT code are computed by linear combinations of the source symbols, the encoding results of both windows can also be used in an additive way.

B. Layer-Aware FEC Using Raptor Codes

Raptor codes were proposed in [127] as improved LT codes, which imposed linearly increasing encoding and decoding complexity as a function of the number of source packets. Raptor codes are furthermore specified under the auspices of the IETF. As a first version, RFC5053 [128] was ratified, which was then followed by an evolved version in RFC6330 [129]. The integration of the LA-FEC concept into RFC5053 and RFC6330 is detailed both in [1] and in [122]. The basic principle is similar to the EWF approach of Section IV-A, but requires additional modifications due to the specific encoding process of the Raptor FEC.

1) Raptor Coding: The basic idea behind Raptor codes relies on precoding the input symbols before applying to them an LT code. Alternatively, Raptor codes are typically constructed by concatenating from two codes, namely the outer code/precode and the inner LT code, where the inner LT code encodes the precoded symbols its input.

Again, Raptor codes may be viewed as an evolution of the LT coding philosophy. The disadvantage of the LT code is that for facilitating decoding, all original uncoded source symbols have to be covered by the received symbols. This leads to a relatively high decoding overhead, since the decoder has to wait in some situations until the reception of a 'repair' symbol, which covers a specific missing source symbol. Raptor codes eliminate this impediment of the LT codes by concatenating the LT code with a precoding operation, which relieves the LT code from the need to decode all source symbols and leads to a significantly lower coding overhead.

Fig. 12 illustrates the coding process of the Raptor code as specified in RFC5053 [128] and RFC6330 [129]. As mentioned above, the encoding process consists of two encoding steps, namely the precoding and the LT coding. In the first step, the precoded symbols are generated by the precoding matrix, which contains the first k rows of the LT encoding matrix of the second encoding step. The precoded symbols are then used as input to the second encoding stage, which invokes an LT encoding process for generating a 'fountain' of encoded symbols. Since the precoding matrix contains the first k rows of the LT coding matrix, the final LT encoding step is an inverse process and the first k symbols of the encoded symbol stream are identical to the k source symbols. This makes the code systematic, which is important for real-world applications for the sake of avoiding the need for FEC decoding under error-free conditions, because in this case simply the systematic information part is retained as the decoded codeword.

2) Layer-Aware Raptor Codes: In order to preserve the original coding performance and functionality of the Raptor code, its integration with the LA process has to satisfy the following requirements:

1) Preserve the original systematic structure of the BL's and EL's symbols, which guarantees their separability;
2) Facilitate the superimposed, enhanced-quality additive decoding of the BL's and EL's symbols;
3) Preserve the original error-resilience of both the BL and EL.

In the following, we consider the example of the EWF code of Fig. 11 with the aid of two SVC layers. The encoded symbols L_0 of the 1st window are generated with the aid of the

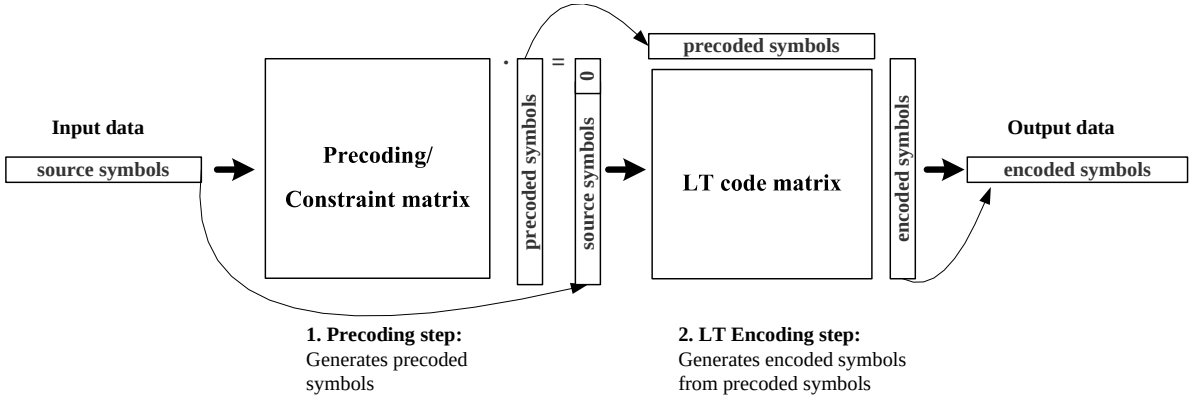


Figure 12: Raptor encoding process as specified in RFC5053 [128] and RFC6330 [1], [129].

standard Raptor encoding process discussed in Section IV-B1. Similarly, following the EWF approach of Fig. 11, the encoded symbols L_1 of the 2nd window are also generated with the aid of another standard Raptor code that covers all source symbols of the 2nd window. However, there are two major problems, when using the standard Raptor encoding process, which may be circumvented by the scheme of Fig. 13, as detailed below:

1) Generally, the EL's encoding process is non-systematic, since the EL is only involved in the encoding process of the 2ed window, which contains the symbols of both the BL and EL.
2) The BL's and EL's encoded symbols cannot be readily decoded to provide an enhanced video quality, since the BL's precoded symbols differ in the original BL coding process and in the EL coding process.

Fortunately, the 2nd problem can be solved by incorporating the BL's precoded symbols in the 2nd encoding step, similar to the EWF process, but using a concatenation of the LT coding matrices of both the BL and of the EL. This ensures that the BL's precoded symbols become part of the EL's precoded symbols. Furthermore, this approach also preserves the original degree distribution of the LT code matrix of both the BL as well as of the EL and thereby preserves the original coding performance. Serendipitously, the 1st problem can also be solved, which is achieved by integrating the LT matrix of the second encoding step into the precoding step and thereby ensures that the first k_1 encoded symbols become identical to the k_1 original source symbols. This coding process is graphically illustrated in Fig. 13.

The above-mentioned modifications of the Raptor code specification satisfy both the 1st and 2nd requirements. Additionally, the results seen in Table V show that the above-mentioned approach is also capable of satisfying the 3rd requirement. The results are based on the LA RaptorQ[3] scheme of the RFC6330 standard as specified in [122]. The performance of Raptor codes is typically characterized in terms of the number of symbols required for successful decoding, i.e. by the parity-overhead. In the specific Raptor code's context, the BL can be decoded iff at least k_0 BL symbols have already been received. By contrast, for the LA-Raptor code, the BL can either be decoded in isolation, provided that more than k_0 BL symbols have been received or alternatively by, using the already available BL symbols in combination with the EL's symbols. In the latter case, both layers can be decoded iff at least $(k_0 + k_1)$ symbols have been received. For the LA-Raptor scheme, we assume having the same size for the BL and EL in conjunction with k_0=k_1. Table V shows the failure probability of the Raptor decoding process for a given source block of k source symbols in conjunction with different number of received symbols. Specifically, k_0 and $(k_0 + 1)$ symbols are received for the BL, while $(k_0 + k_1)$ and $(k_0 + k_1 + 1)$ for the EL.

The results of Table V demonstrate that the above-mentioned LA-Raptor modifications achieve a comparable decoding performance to that of the original Raptor code. The reception of a single additional symbol is sufficient in both codes for attaining a near-unity probability of successful decoding.

C. Layer-Aware FEC Using Randomized Expanded Window Based RS Codes (RE-RS)

The integration of the LA-FEC approach with Reed-Solomon (RS) codes is described in [113]. The authors conceived the randomized expanded window based RS coding (RE-RS) philosophy for delay sensitive real-time video streaming. In this context, the RS coding is expanded over a number of previously received pictures in order to recover the reference pictures in the case of transmission problems, albeit this is achieved at the cost of a moderate single-codeword FEC delay. The general approach is illustrated in Fig. 14. Again, in order to limit the delay introduced by the FEC to a single video frame, the FEC source block is expanded only to video frames received in the past, thereby each video frame may be decoded and displayed at the appropriate time instant.

The authors designed the RE-RS encoding scheme according to the LA-FEC approach, constructing the RE-RS in a way so that all the RS-decoded information of the current video frame and that of all the previous video frames can be conveniently combined at the decoder and that the

[3]RaptorQ code [129] is a variant of Raptor codes [128], which offers a better coding performance, i.e. its reduced symbol overhead for guaranteeing successful decoding and supports larger source symbol block sizes.

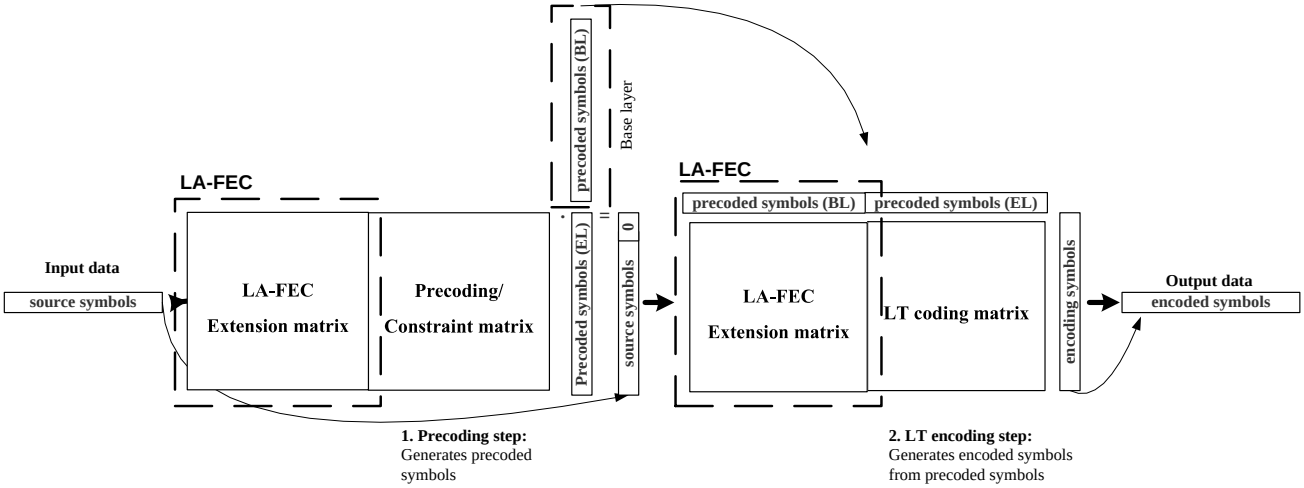


Figure 13: Modified Layer-Aware Raptor coding process [1].

Table V: Decoding failure probabability of both the Raptor and of the LA-Raptor scheme of the RFC6330 standard [122].

	Raptor		RaptorLA	
Received	k_0	k_0+1	k_0+k_1	k_0+k_1+1
k=50	4.68×10^{-3}	3.82×10^{-16}	4.26×10^{-3}	1.42×10^{-16}
k=100	4.47×10^{-3}	2.71×10^{-16}	3.95×10^{-3}	1.43×10^{-16}
k=200	4.98×10^{-3}	3.77×10^{-16}	2.91×10^{-3}	3.12×10^{-16}
k=400	5.19×10^{-3}	1.07×10^{-16}	1.98×10^{-3}	2.31×10^{-16}
k=800	3.97×10^{-3}	2.75×10^{-16}	2.59×10^{-3}	3.48×10^{-16}
k=1600	2.82×10^{-3}	1.95×10^{-16}	3.86×10^{-3}	3.52×10^{-16}

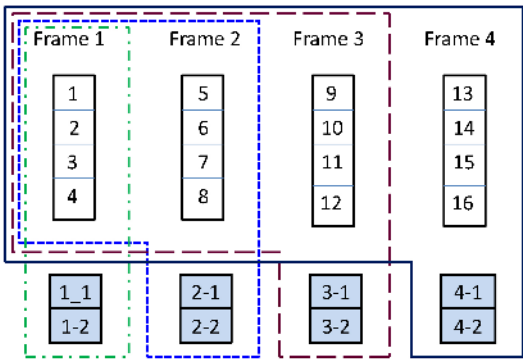


Figure 14: Example of the RE-RS scheme, where each transmission frame contains 4 original m-bit video strings, each representing an m-bit RS-symbol, and 2 redundant RS-symbols [113].

coding performance of the combined RS code was preserved compared to a conventional RS coded scheme.

1) Reed-Solomon Codes: RS codes are widely used in numerous operational applications. An $RS(N, K)$ code takes K symbols as its input and generates a codeword having a total length of N symbols. They are maximum distance separable (MDS) codes, which have the beneficial characteristic that receiving *any* K out of the N RS-encoded symbols is sufficient for the successful decoding of the K *original* source symbols. Let us assume for example using a systematic RS codeword of $C = (c_1, c_2, ..., c_N)$ having $N = 2^m - 1$ encoded symbols and m bits per RS-encoded symbol, where $2t$ symbols are lost. A decoder will attempt to recover the $2t = (N - K)$ packets by solving the $2t$ parity-check equations of

$$CH^t = 0, \tag{4}$$

where H is the parity-check matrix, formulated as follows:

$$H = \begin{bmatrix} 1 & \alpha & ... & \alpha^{2^m-3} & \alpha^{2^m-2} \\ 1 & \alpha^2 & ... & (\alpha^2)^{2^m-3} & (\alpha^2)^{2^m-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \alpha^{N-K} & ... & (\alpha^{N-K})^{2^m-3} & (\alpha^{N-K})^{2^m-2} \end{bmatrix}, \tag{5}$$

with α being the so-called primitive element of the Galois Field GF(2^m). Eq. (4) can be solved, when we have $2t \leq N - K$, since the rank of the matrix H is $(N - K)$.

To elaborate a little further, t physically represents the number of m-bit symbols that can be corrected by an $(N, K, 2t)$ code. Explicitly, t equations are solved for finding the positions of the erroneous m-bit symbols, and t equations for finding the specific error-magnitudes of the t erroneous m-bit symbols. By contrast, in the scenario considered the position of the unrecovered symbols is known, hence the $2t$ equations can be directly solved for finding the missing $2t = (N-K)$ symbols.

2) RE-RS Scheme: When considering the example seen in Fig. 14, the basic approach of the RE-RS code is to use a standard RS code, such as an RS(16,14) code defined over the

Galois Field GF(2^4) and to encode the first frame with the aid of a shortened RS(6, 4) code, where the 10 remaining unused symbols are zero-valued padding symbols, which are removed after encoding. The second video frame of Fig. 14 is encoded with the aid of the same RS(16, 14) code but using only 6 zero-valued padding symbols, hence yielding an RS(10, 8) code and so forth. As a benefit, the different RS-encoded data can be combined at the receiver, since they belong to the same RS codeword. However, the authors of [113] demonstrated that with the aid of the standard RS code, a combination of the RS-encoded data of the different video frames would fail to achieve the performance of an RS code. This is illustrated by a simple example, which considers the first two frames of Fig. 14. The underlying assumption is that the source symbol 1, 2, and 3 of the first video frame and 5 of the second video frame are lost, where the total number of lost symbols is given by $2t = 4$. According to Eq. (4), the parity-check equations of the first frame can be formulated as follows:

$$\begin{aligned} X_1 + \alpha X_2 + \alpha^2 X_3 + C_1 &= 0 \\ X_1 + (\alpha^2) X_2 + (\alpha^2) X_3 + C_2 &= 0, \end{aligned} \tag{6}$$

where X_1, X_2, and X_3 denote the three lost m-bit RS-symbols, while C_1 and C_2 are constant values, which are determined by the RS-symbols received for the first transmission frame. It is obvious that this pair of equations is insufficient for finding the three unknowns. However, as an explicit benefit of the RE-RS approach, we can now combine the RS-decoded data of the two transmission frames, which leads now to the following four parity-check equations:

$$\begin{aligned} X_1 + \alpha X_2 + \alpha^2 X_3 + C_1 &= 0 \\ X_1 + (\alpha^2) X_2 + (\alpha^2) X_3 + C_2 &= 0 \\ X_1 + \alpha X_2 + (\alpha^2) X_3 + (\alpha^4) X_5 + C_3 &= 0 \\ X_1 + (\alpha^2) X_2 + (\alpha^2)^2 X_3 + (\alpha^2)^4 X_5 + C_4 &= 0, \end{aligned} \tag{7}$$

where X_5 denotes the lost RS-symbol of the second transmission frame, while C_3 and C_4 are constant values determined by the RS-coded transmission received for the second transmission frame. This equation has to be solved over the Galois field GF(2^m). However, since the coefficients of X_1, X_2 and X_3 in the first and the third equation are the same, the rank of the combined matrix becomes 3, which is insufficient for determining four variables. Hence upon using the standard RS code, the four lost symbols cannot be recovered.

The authors of [113] solved this problem of combining two different RS codes by a randomized reordering of the m-bit video strings representing the RS-symbols of a transmission frame before the actual RS encoding step. This approach ensures with a high probability that the coefficients of the parity-check equations derived for different transmission frames become independent of each other. This is shown with the aid of an example in [113], where the randomly reordered positions of the lost RS-coded symbols {1,2,3} become {6,3,11}. By the same token, for the second RS-encoded window, the positions {1,2,3,5} become {7,1,4,12}. In this case, the combined RS parity-check equations for the second frame become:

$$\begin{aligned} \alpha^5 X_1 + \alpha^2 X_2 + \alpha^1 0 X_3 + C_1' &= 0 \\ (\alpha^2)^5 X_1 + (\alpha^2)^2 X_2 + (\alpha^2)^1 0 X_3 + C_2' &= 0 \\ \alpha^6 X_1 + X_2 + (\alpha^3) X_3 + (\alpha^1 1) X_5 + C_3' &= 0 \\ (\alpha^2)^6 X_1 + X_2 + (\alpha^2)^3 X_3 + (\alpha^2)^1 1 X_5 + C_4' &= 0. \end{aligned} \tag{8}$$

In the example considered, the rank of the coefficient matrix defined over the Galois Field GF(2^4) of the RS(16,14) code becomes 4, which makes it a full rank matrix and hence the equations can be readily solved. It should be noted that the randomization approach relies on an intelligent algorithms that minimizes the probability of having the same coefficients in the different windows.

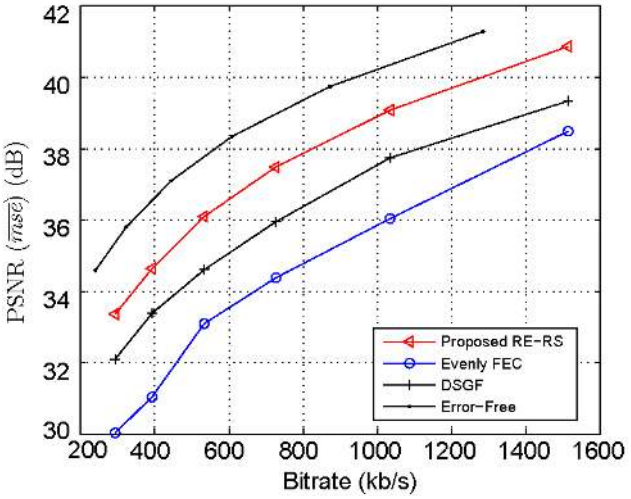
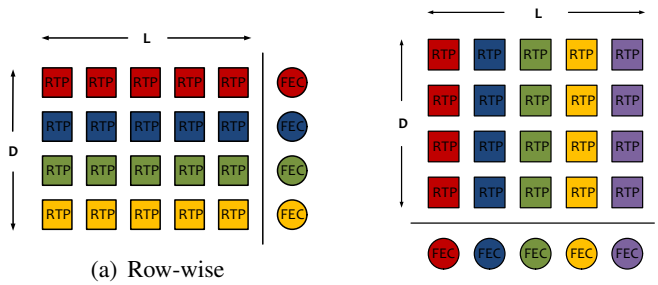


Figure 15: RE-RS: Average PSNR versus bitrate curves; i.i.d. average packet loss rate is 5% and the redundant overhead was 20% for the Foreman sequence [113].

The performance of the proposed scheme is characterized in Fig. 15. Here the authors of [113] compared their RE-RS scheme to an evenly distributed FEC (Evenly-FEC), to the so-called dynamic sub-GOP FEC (DSGF) scheme [130] and to an error-free curve. An H.264/AVC video stream and an 10-bit per symbol RS code defined over GF(2^{10}) is used. The figure shows the peak signal-to-noise ratio (PSNR) of the decoded video clip averaged over 200 trials. For the three approaches, the same quantization parameter (QP) is used for encoding the video sequences, and the same number of redundant symbols was inserted for ensuring a fair comparison. An average packet loss ratio of 5% was used. The results of Fig. 15 show the clear benefit of the RE-RS approach compared to DSGF and to the evenly distributed FEC scenario.

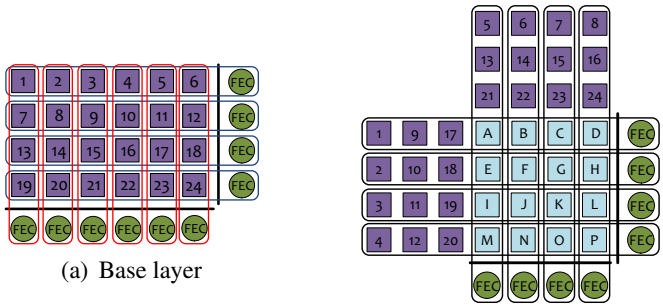
D. Layer-Aware FEC Using the Pro-MPEG COP3 Code

The authors of [114] demonstrated the integration of the LA-FEC technique with the Pro-MPEG COP3 code of [131], which is widely used in IP-based video delivery networks. Although this code does not achieve the error correction performance of the other codes described earlier in this



(a) Row-wise (b) Column-wise

Figure 16: The standard Pro-MPEG COP3 code's parity generation [114].



(a) Base layer (b) Enhancement + base layer

Figure 17: Layer-Aware extension of the Pro-MPEG COP3 code [114].

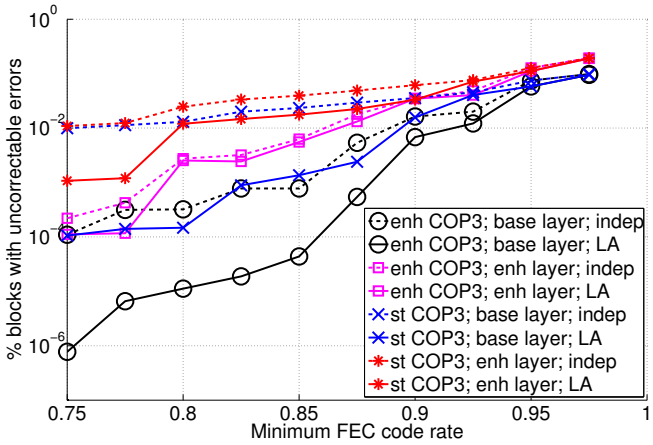


Figure 18: Performance results for Layer-Aware extension of Pro-MPEG COP3 code [114].

section, it is particularly beneficial for employment in high-throughput networks owing to its extremely low computational and memory demand, while providing a sufficiently high error correction capability.

1) Pro-MPEG COP3 Code: The Pro-MPEG COP3 code is based on a simple XOR-aided combination of the video source symbols. Before encoding, the source symbols are arranged in a matrix having D rows and L columns, as shown in Fig. 16. The FEC-encoded redundancy can then be generated row-wise (a), which is suited for correcting independent packet loss events, or column-wise (b), which is intended for correcting bursts of consecutive errors. The L and D parameters are used for adjusting the code rate between the column-wise and row-wise FEC codes.

2) Layer-Aware Pro-MPEG COP3 Code: For the design of the Layer-Aware Pro-MPEG COP3 code, the simple concatenation of the BL and EL parity matrices shown for the EWF code in Fig. 10 does not provide a suitable solution. This is because in the situation that a row of the matrix cannot be corrected by the single BL-protection FEC symbol available, the single FEC symbol provided for the EL will have the same problem and hence a combined decoding becomes impossible. To circumvent this and to provide further benefits, when the BL and EL are jointly decoded, the authors of [114] proposed a reordering of the BL source bits for the EL FEC data generation, as shown in Fig. 17.

The scheme illustrated in Fig. 17 operates as follows. The BL source data of Fig. 17 (a) is protected by its own FEC code, following the standard Pro-MPEG COP3 procedure [131]. For the EL source data of Fig. 17 (b), firstly the EL parity matrix is generated by using the standard Pro-MPEG COP3 code. In a second step, the BL data is uniformly distributed across the two-dimensions of the EL with the aid of their XOR connection, whilst avoiding XORing the same source symbols, as in the BL FEC generation. The performance difference of a variety of schemes is shown in Fig. 18. Specifically, the curves with the "enh" label refer to another enhancement of the Pro-MPEG COP3 code proposed by the authors of [114], which increases the attainable protection capability by incorporating additional FEC protection across the diagonal of the parity matrix. The curves with the "st" label refer to the standard Pro-MPEG COP3 code of [131]. For both cases, the figure shows a comparison of the BL and EL block error rate using both an independent encoding and the LA-FEC extension of this section.

E. Layer-Aware FEC Using UEP-LDPC Code

A number of application layer FEC codes were mentioned above. By contrast, **LDPC** codes are widely used as physical layer codes in broadcast systems as a benefit of their capacity approaching performance. In [54], the authors proposed a UEP scheme based on LDPC codes (UEP-LDPC), which relies on an EWF-like approach used for LT codes in Section IV-A. Note that this UEP-LDPC code may be soft-decoded or hard-decoded for employment in the physical layer or application layer, respectively.

1) LDPC Code: LDPC codes belong to the family of linear error correcting codes defined by a sparse parity-check matrix H, where the legitimate codeword c has to satisfy the parity-check equations of $Hc^T = 0$. The parity-check matrix H can be represented by a Tanner graph [132][4], constituted by the

[4]Tanner graphs were proposed by Michael Tanner for constructing long error correcting codes from multiple shorter ones using recursive techniques. Tanner graphs contain two types of nodes, namely check nodes and variable nodes [133]. For linear block codes, such as LDPC, the check and variable nodes represent the rows and columns of the parity-check matrix H, respectively. An edge connects the check node i to the variable node j, if the entry (i, j) of matrix H is nonzero.

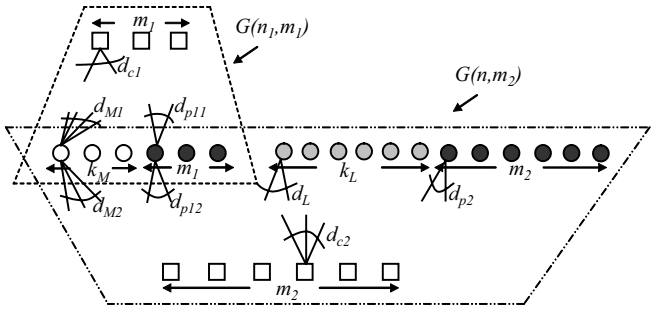


Figure 19: The Tanner graph of the proposed ensemble [54].

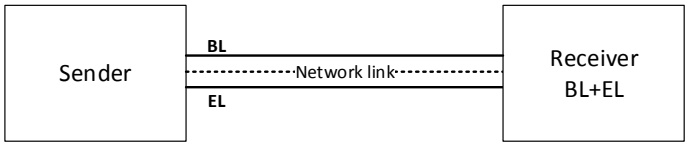


Figure 20: Single-link scenario.

inter-connected set of variable nodes and check nodes with the aid of a set of edges.

2) UEP-LDPC Code: The authors of [54] proposed to encode a pair of different-importance classes using a technique similar to the sliding window approach of the EWF codes of Section IV-A [82], [85]. This approach is characterized by two overlapping Tanner graphs, as seen in Fig. 19. The first Tanner graph was conceived for protecting the $n_1 = (k_1+m_1)$ variable nodes, representing for example the BL data. By contrast, the second Tanner graph protects $(n_1 + n_2)$ variable nodes, which correspond to the conceived BL and EL data.

When using this approach, the two UEP importance classes, namely the BL and EL, have their source-sensitivity-matched protection, which can be jointly decoded with the aid of the appropriately combined parity-check matrix given by:

$$H = \begin{bmatrix} H_1 & 0 \\ H_{21} & H_{22} \end{bmatrix}. \quad (9)$$

F. An Application Scenario and Performance Analysis

All codes considered in this section rely on the basic principle of the Layer-Aware FEC [1], which facilitates the joint decoding of the BL and EL, whilst preserving the original coding performance of the layer-agnostic schemes. The performance benefits of the LA-FEC manifest themselves when they are invoked for transmission over realistic channels. In [134], the author analyzed the performance of LA-FEC vs. standard FEC schemes in different application scenarios. The first one is shown in Fig. 20, where both layers are transmitted over the same link. Fig. 21 illustrates the second scenario, where the upper receiver requests only the BL stream, while the receiver seen at the bottom requires both the BL and the EL. The BL and EL are transmitted over different links of the network, where one of them may be a cellular network, the other one a WiFi network.

For simplicity, in our forthcoming discourse an ideal LA-FEC code is assumed, which requires $r \geq k_0$ received symbols for decoding the BL in its own right, while $r \geq k_0 + k_1$ symbols are necessitated for decoding both the BL and EL, when the BL cannot be recovered in isolation. Note that this ideal performance is close to that of the LA-Raptor code's performance shown in Table V or to that of the RE-RS code characterized in Section IV-C. The FEC scheme was specifically designed in order to avoid increasing the delay introduced by the system, which was that of the video codec imposed by the frame-duration.

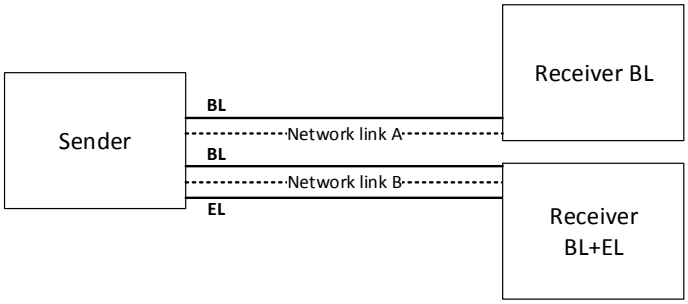


Figure 21: Layered multicast for two different receivers.

The results were recorded both for a packet erasure channel (PEC) and for a bursty erasure channel simulated by a Gilbert-Elliot (GE) model, where the GE model is configured according to typical IPTV scenarios [135] associated with an average error-burst length of $L = 100$ms and an average packet erasure ratio of p_{er}=0.05. The PEC channel has the same packet erasure ratio of p_{er}=0.05.

We embarked on testing diverse distributions of the parity information across the two layers and decoding operations were curtailed as soon as the target decoding performance of both layers was achieved. Note that the successful decoding of the EL depends on the successful decoding of the BL. Two different bit rate ratios (BRR) of the BL and EL were tested, namely $1:3$ and $3:1$, where $1:3$ corresponds to the typical spatial scalability of the SVC bitstream, while $3:1$ corresponds to the temporal scalability of an H.264/AVC bitstream. The amount of the overall required protection is measured in terms of the normalized overhead, where for example a value of 1 indicates an unprotected media stream and a value of 1.1 indicates that a 10% higher bit-rate is required for protection in order to achieve the 99% successful decoding probability. The resultant overheads are summarized in Tables VI and VII for both conventional FEC and for the LA-FEC code discussed above. Tables VI and VII also portray the optimal code rate (CR) allocation for the different layers of the different settings, where the code rate is calculated as CR=K/N with K being the number of source symbols and N the number of FEC encoded symbols. A total of 10000 test runs were performed.

1) Single-link Scenario: In single-link broadcast scenarios the target performance assumed corresponds to guaranteeing a 99% successful decoding probability of the BL. By contrast, for the EL the service provider may relax this constraint to a 95% successful decoding probability. Therefore we considered different successful decoding probabilities for the EL. The corresponding results are shown in Table VI.

The results recorded in Table VI for the single link scenario

Table VI: Single link: The overhead and the related FEC code rate (CR) allocation required for achieving diverse decoding probability for the BL and EL for transmission over the GE channel in conjunction with a BRR of 1:3 and 3:1 between BL and EL.

	Decoding probability		
EL	99%	95%	91%
	Overhead		
ST 1:3	1.68	1.62	1.55
LA 1:3	1.62	1.57	1.48
ST 3:1	1.67	1.65	1.60
LA 3:1	1.63	1.60	1.58
	FEC code rate BL/EL		
FEC 1:3	0.58/0.61	0.60/0.63	0.58/0.70
LA 1:3	0.89/0.57	0.62/0.62	0.61/0.70
ST 3:1	0.69/0.45	0.70/0.46	0.71/0.50
FEC 3:1	0.89/0.31	0.80/0.40	0.70/0.50

of Fig. 20 show the reduced overhead of the LA-EC compared to that of the layer-agnostic standard FEC (ST-FEC) for all decoding probabilities and bit rate distributions. When considering a 91% successful decoding probability, the gain observed for the LA-FEC is in the range between 20% and 70% in conjunction with the BRRs of 3:1 and 1:3. By observing the optimal code rate distribution we infer the general trend for the LA-FEC, which suggests that a stronger protection has to be invoked for the EL. By contrast, for layer-agnostic ST-FEC, typically a stronger protection of the BL is beneficial for the BRR of 1:3. For 3:1, the stronger protection has to be applied to the EL.

2) Layered multicast Scenario: For the layered multicast scenario of Fig. 21, it is assumed that the users expect to have a high-quality service for all layers. In this context, our objective is to find the minimum amount of protection required for achieving a successful decoding probability of 99% for each layer. Note that in the multicast scenario considered, the "Receiver BL" only receives the BL data and has to achieve the target successful decoding probability by relying on the BL data only. By contrast, the "Receiver BL+EL" scenario receives both layers, while relying on at least the same amount of protection as that required for the BL. The corresponding results recorded for the layered multicast scenario are shown in Table VII.

Specifically, the results of Table VII show that the LA-FEC provides gains compared to the layer-agnostic ST-FEC protection in terms of a 5% lower overhead for transmission over the PEC and an 18% lower overhead for the GE channel, which corresponds to an approximately 22% lower protection required for both channel types and for both BRRs. Concerning the optimal CR allocation, the results show that the ST-FEC invoking stronger protection of the BL shows the best performance, albeit in case of the PEC channel associated with the BRR of 3:1, the smaller EL requires a stronger protection than the BL in order to achieve the target decoding probability of 99%. For LA-FEC, typically the arrangement applying a stronger protection to the EL exhibits the best performance.

Table VII: Layered multicast: The overhead and the related FEC code rate (CR) allocation required for achieving a 99% decoding probability for the BL and EL for transmission over the PEC and GE channels in conjunction with a BRR of 1:3 and 3:1 between BL and EL.

	PEC		GE	
	ST-FEC	LA-FEC	ST-FEC	LA-FEC
Overhead (1:3)	1.23	1.18	1.80	1.62
CR of BL	0.77	0.88	0.55	0.65
CR of EL	0.84	0.84	0.56	0.61
Overhead (3:1)	1.23	1.18	1.80	1.62
CR of BL	0.84	0.88	0.48	0.66
CR of EL	0.79	0.77	0.66	0.55

G. Summary

The packet-level LA solution was detailed above, which may be applied to a number of FEC codes, such as LDPC, Raptor codes etc. Furthermore, the simulation results of Table VII show that the LA-FEC solution outperforms the traditional FEC protection schemes by significantly reducing both the overhead and the protection required for achieving a specific decoding probability.

V. Bit-Level Protection Using: Inter-Layer FEC

Diverse bit-level IL-FEC schemes were investigated in [2], [115], [116] relying on RSC, turbo and SECCC codes. In this section, we will review the general architecture of the inter-layer FEC scheme proposed in [2], [102], [115], which was conceived for layered video transmission. Section V-A details the preliminaries, including the MI[5] flow between the variable node decoder (VND) and check node decoder (CND), as seen in Fig. 22. Then we continue by detailing the transmitter side of Fig. 22, which relying on coding rates of r_0 and r_1. The online "Code Rate Optimization" block will be illustrated in Section V-D, while the inter-layer H.264/AVC decoding techniques will be illustrated in Section V-C, with special emphasis on how the VND and the CND exchange their inter-layer redundancy for improving the overall performance of the system. Finally, Sections V-E and V-F discuss the overheads imposed by the IL technique.

As seen in Fig. 22, L_0 is the BL and L_1 is an EL, where the EL will be utilized for protecting the BL L_0. Based on [2], [102], in this section, we assume that both the layers L_0 and L_1 contain n bits for the sake of convenience of explanation. However, this algorithm may be readily extended to the more general scenarios, where more layers carrying an unequal number of bits may be considered, as detailed in [2], [102]. Moreover, the blocks "Mod." and "Demod." of Fig. 22

[5]The MI is a metric invoked for representing the reliability of a signal sequence. Generally speaking, a higher MI indicates a lower BER value of the measured signal sequence, while lower BER normally indicates a lower PLR.

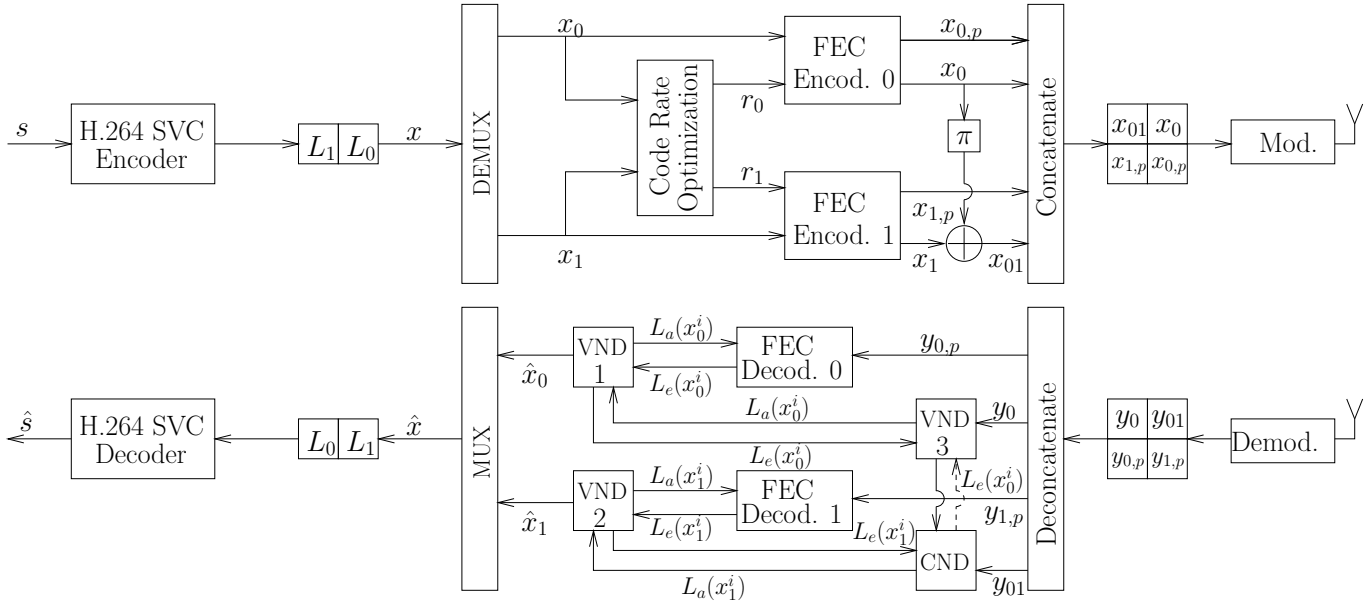


Figure 22: Bit-level IL-FEC architecture of H.264/AVC SVC coded video.

represent a generalized transceiver architecture, but in practice a diverse suite of transceiver structures may be employed, as discussed in [2], [102], [115], [116].

A. Preliminaries

The structures of the VND and CND [136] in Fig. 22 are further detailed in Fig. 23. Both the VND and CND blocks may accept a maximum of three soft information inputs and generate a maximum of three soft information outputs with the goal of iteratively exploiting all IL dependencies amongst the FEC coded layers L_0 and L_1. The behavior of the VND and CND will be characterized based on the assumption that u_1, u_2 and $u_3 = u_1 \oplus u_2$ are random binary variables.

1) VND: The action of the VND of Fig. 23a sums two LLR inputs for generating a more reliable LLR output, which may be formulated as $L_{o_3}(u_1) = L_{i_1}(u_1) + L_{i_2}(u_1)$.

As seen in Fig. 23a, the operation of the VND may be characterized as $L_{o_3}(u_1) = L_{i_1}(u_1) + L_{i_2}(u_1)$. Assuming that the inputs $L_{i_1}(u_1)$ and $L_{i_2}(u_1)$ of the VND have the MI values of $I_{i_1}[u_1; L_{i_1}(u_1)]$ and $I_{i_2}[u_1; L_{i_2}(u_1)]$ respectively, the MI value of the output $L_{o_3}(u_1)$ may be expressed as [137]

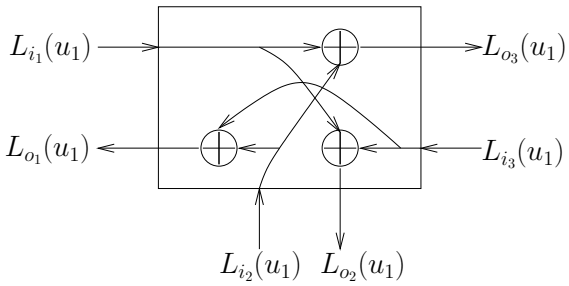
$$I_{o_3}[u_1; L_{o_3}(u_1)] = J\left(\sqrt{J^{-1}(I_{i_1}[u_1; L_{i_1}(u_1)])^2 + J^{-1}(I_{i_2}[u_1; L_{i_2}(u_1)])^2}\right), \tag{10}$$

where $J(\cdot)$ and $J^{-1}(\cdot)$ are defined in [136].

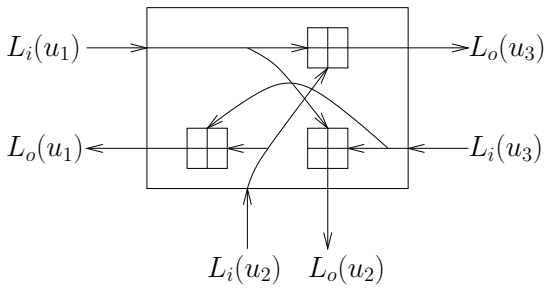
2) CND: The boxplus operation of $L(u_3 = u_1 \oplus u_2) = L(u_1) \boxplus L(u_2)$ [138] may be utilized for deriving the confidence of the bit u_3, given that the confidence of the bits u_1 and u_2 is known. Specifically, the boxplus operation $\boxplus$ is defined as follows [139]

$$\begin{aligned} L(u_1) \boxplus L(u_2) &= \log \frac{1 + e^{L(u_1)} e^{L(u_2)}}{e^{L(u_1)} + e^{L(u_2)}} \\ &= \text{sign}[L(u_1)] \cdot \text{sign}[L(u_2)] \cdot \min[|L(u_1)|, |L(u_2)|] \\ &+ \log\left[1 + e^{-|L(u_1)+L(u_2)|}\right] - \log\left[1 + e^{-|L(u_1)-L(u_2)|}\right]. \end{aligned} \tag{11}$$

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,



(a) VND



(b) CND

Figure 23: The structure of VND and CND, where $\oplus$ and $\boxplus$ indicate the addition and boxplus operation, respectively. $L_i(\cdot)$ and $L_o(\cdot)$ indicate the input and output LLR, respectively.

the MI value $I_o\left[u_3; L(u_3)\right]$ of u_3 may be readily derived with the aid of LUTs, as detailed in [102].

B. Transmitter Model

At the transmitter of Fig. 22, the video source signal s is compressed using the data partitioning mode of the H.264/AVC encoder, generating the bitstream x carrying the layers L_0 and L_1. Then the output bitstream is de-multiplexed into two bitstreams by the *DEMUX* block of Fig. 22, namely into the binary streams x_0 and x_1, representing the layers L_0 and L_1, respectively. Then 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. 22, 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 without changing the parity bits of EL $x_{1,p}$. Specifically, the implantation process results in the check bits $x_{01}^i = x_0^i \oplus x_1^i$. 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, which contain the information of the layers L_0 and L_1. Additionally, the interleaver π is 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 and then transmitted to the receiver.

C. Receiver Model

At the receiver of Fig. 22, the demodulation is performed [140], generating the LLRs, which contain the systematic information y_0, y_{01} and the parity information $y_{0,p}$ and $y_{0,p}$, for the L_0 and L_1 layers, respectively. Following the demodulator, the IL-FEC decoder of Fig. 22 is invoked for exchanging extrinsic information across the three layers. The IL aided FEC decoding process is illustrated by the flow-chart of Fig. 24. 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 BL L_0 will be input to the "VND1-VND3-CND-VND2" block of Fig. 24 for extracting the *a-priori* LLRs $L_a(x_1^i)$[6] of EL L_1, which is carried out by following the processing of the LLRs in the VND1, VND 3, CND and VND 2 components of Fig. 23. Specifically, the "VND1-VND3-CND-VND2" block of Fig. 24 performs the following operations step-by-step:

1) VND 1 generates the information of BL L_0 for VND 3. The inputs of the VND 1 block are constituted by the soft information $L_e\left(x_0^i\right)$ generated by the FEC decoder 0 and the soft information $L_a\left(x_0^i\right)$ generated by the VND 3 block. The output of the VND 1 block is the soft information of L_0. The output can be readily derived as detailed in Fig. 23a. The extrinsic LLR $L_e\left(x_0^i\right)$ generated by the FEC decoder 0 is input to the VND 1 block of Fig. 22, which extracts the extrinsic LLR information $L_e\left(x_0^i\right)$ and forwards it to the VND 3 block of Fig. 22. Since VND 1 has two input branches, it simply duplicates the soft information $L_e\left(x_1^i\right)$.
2) VND 3 generates the information of BL L_0 for CND. The inputs to VND 3 block are constituted of the soft information $L_e\left(x_0^i\right)$ generated by the VND 1 block and the channel information y_0. The output of the VND 3 block is the soft information of L_0. The output can be readily derived as detailed in Fig. 23a. The VND 3 block of Fig. 22 extracts the extrinsic LLR information $L_e\left(x_0^i\right)$ and forwards it to the CND block of Fig. 22.
3) CND generates the information of layer L_1 for VND 2. The inputs of the CND block are the soft check information y_{01} received from the channel, the soft information $L_e\left(x_0^i\right)$ of BL L_0 generated by VND 3 and the soft information $L_e\left(x_1^i\right)$ of EL L_1 generated by the VND 2 of Fig. 22. The output of CND is the

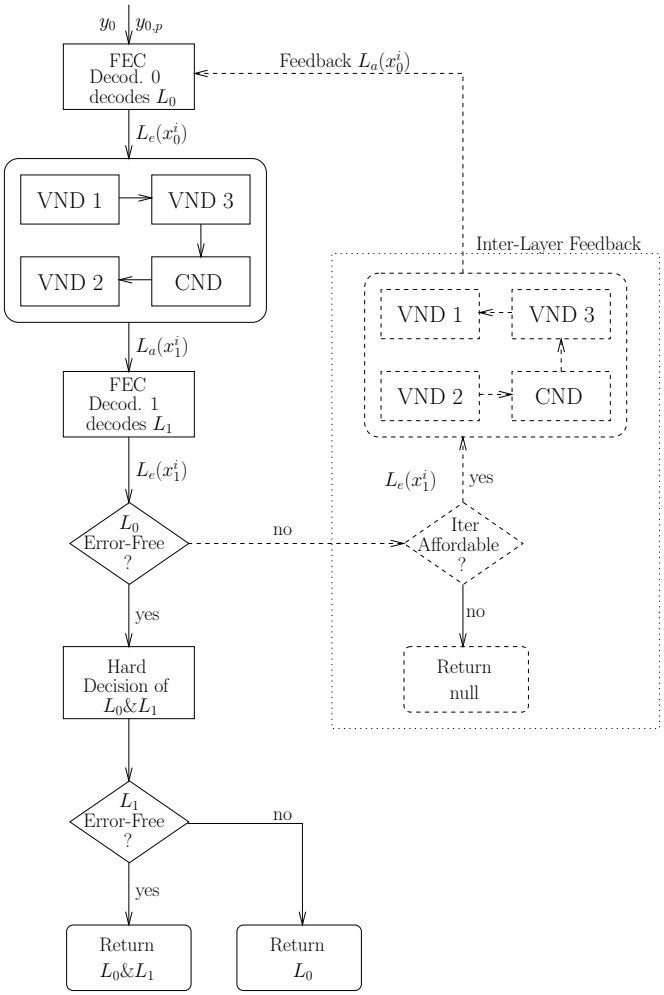


Figure 24: The flow chart for inter-layer aided FEC decoding of BL L_0 and EL L_1.

[6]As usual, the subscripts "a" and "e" in L_a and L_e stand for the a-priori information and extrinsic information [141], respectively.

soft information of EL B $L_a\left(x_1^i\right)$. The outputs can be readily derived as detailed in Fig. 23b. The LLR information $L_e\left(x_0^i\right)$ and the received check information y_{01} is input to the CND block of Fig. 22 for extracting the LLR information of the systematic bit x_1^i, namely the soft input $L_a\left(x_1^i\right)$ of VND 2.

4) VND 2 generates the information of EL L_1 for FEC decoder 1. The inputs to the VND 2 block are the soft information $L_a\left(x_1^i\right)$ gleaned from CND and the soft information $L_e\left(x_1^i\right)$ generated by FEC decoder 1. The output of VND 2 is the soft information of layer L_1. The LLR information $L_a\left(x_1^i\right)$ extracted by the CND is input to the VND 2 block of Fig. 22, which extracts the LLR information $L_a\left(x_1^i\right)$ input to the FEC decoder 1 of Fig. 22.

Then, the FEC decoder 1 of Fig. 24 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. 22. Afterwards, the classic CRC is invoked for detecting, whether the recovered BL L_0 is error-free or not, as shown in Fig. 24. This check results in two possible decoding processes, as shown in Fig. 24 and described as follows:

1) With Inter-Layer Feedback: When the bits x_0 of the BL are not successfully decoded, the iterative IL technique will be activated for exploiting the extrinsic information of BL L_0 fed back from the FEC decoder 1. In this case, both the solid lines and the dashed lines shown in the decoder of Fig.s 22 and 24 will be activated. More explicitly, the "VND2-CND-VND3-VND1" block of Fig. 24 will be utilized for extracting the extra LLR information $L_e(x_0^i)$ for BL L_0 based on both the extrinsic LLR $L_e(x_1^i)$ and the soft check information y_{01}. Generally, the "VND2-CND-VND3-VND1" block of Fig. 24 represents a process similar to that of the "VND1-VND3-CND-VND2" block of Fig. 24. After this stage, improved *a-priori* information is generated for the BL L_0, which concludes the current IL decoding iteration. Afterwards, the receiver will return to the beginning of the flow chart shown in Fig. 24. The iterative IL decoding process continues, until the affordable number of iterations is exhausted or the BL L_0 is perfectly recovered, as shown in Fig. 24.

2) Without Inter-Layer Feedback: When the BL L_0 is successfully recovered, the layers L_0 and L_1 will be estimated by the hard decision block of Fig. 24. Afterwards, the receiver may discard layer L_1, depending on whether it is deemed to be error-free or not by the CRC check. In this case, only the solid lines of Fig.s 22 and 24 will be activated.

Moreover, after decoding BL L_0, the recovered error-free hard bits x_0 may be represented using infinite LLR values, indicating the hard bits $0/1$, respectively. Then, the CND process invoked for generating the LLR $L\left(x_1^i\right)$ shown in Fig. 22 may be derived as follows using the boxplus operation

$$\begin{aligned} & L\left(x_1^i\right) \\ & = L(x_0^i) \boxplus L(x_{01}^i) \\ & = \operatorname{sign}\left[L(x_{\mathrm{o}}^i)\right] \cdot \operatorname{sign}\left[L(x_{01}^i)\right] \cdot \min\left[\infty, |L(x_{\mathrm{o}1}^i)|\right] \\ & + \log\left(1+e^{-\infty}\right) - \log\left(1+e^{-\infty}\right) \\ & = \operatorname{sign}\left(\tilde{x}_0^i\right) \cdot L\left(x_{01}^i\right), \end{aligned} \tag{12}$$

where $\tilde{x}_0^i$ is the modulated version of the bit x_0^i and the LLR input $L\left(x_{01}^i\right)$ is obtained by soft demodulating the received signal y_{01}.

Note that since the process of recovering y_1 from y_{01} expressed by Eq. (12) is essentially an LLR sign-flipping operation, it does not affect the absolute value of the LLR information of x_1. This implies that in this scenario the IL technique is equivalent to the traditional UEP techniques, where layers L_0 and L_1 are encoded and decoded independently. Moreover, since BL L_0 is decoded independently without feedback from EL L_1, the two layers are only decoded once, without any extra complexity imposed on the receiver. Additionally, in practical applications, BL L_0 may be reconstructed immediately when it is received, without waiting for the arrival of the EL L_1.

In both of the above cases, if the decoded bit sequence $\hat{x}_0$ of the BL is corrupted after the IL-FEC decoding stage of Fig. 22, it will be dropped together with the EL $\hat{x}_1$. Otherwise they will all be forwarded to the H.264/AVC decoder of Fig. 22 for reconstructing the video signal $\hat{s}$.

D. Distortion Minimization

In Sections V-B and V-C, we detailed the IL-FEC coded system of [2] conceived for layered video transmission, when the code rates r_0 and r_1 of Fig. 22 are given. However, in practical scenarios, different coding rates r_0 and r_1 are required for achieving the best system performance. This code-rate optimization process is accomplished by the "Code Rate Optimization" block of Fig. 22. Below, we introduce the techniques invoked by the "Code Rate Optimization" block for finding the optimized coding rates for the channel-coded bitstreams "on-the-fly" at the transmitter for the sake of optimizing the IL-FEC coded system's performance. Below, we consider an RSC codec and a binary phase shift keying (BPSK) transceiver for describing the proposed solution. However, the proposed techniques are not limited to the RSC codec and indeed, further research is required in the context of diverse other components. Firstly, the following notations are defined, which will aid the analysis:

- $d(L_0)$: video distortion, namely the PSNR reduction, induced by the corruption of the BL L_0, which is measured using the PSNR;
- $d(L_1)$: the video 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. 22;
- r_i: the coding rate of layer L_i, $0 \leq i \leq 1$;
- $p\left(L_0\right)$: the packet error ratio (PER) of layer L_0;
- $p\left(L_1\right)$: the PER of layer L_1, when the layer L_0 is correctly decoded.

According to the IL-FEC decoding process detailed in Section V, both the FEC decoders 0 and 1 of Fig. 22 affect the PER $p(L_0)$, where FEC decoder 1 feeds back *a-priori* information of BL L_0 through the VND and CND blocks of Fig. 22. 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

$$p(L_0) = f_0\left(SNR, |L_0|, |L_1|, r_0, r_1\right). \tag{13}$$

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

$$p(L_1) = f_1\left(SNR, |L_1|, r_1\right). \tag{14}$$

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 [78]. Then the expected PSNR reduction 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 reduction 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 $E\left[d(L_0,L_1)\right]$, may be estimated as

$$E\left[d(L_0,L_1)\right] = p(L_0) \cdot d(L_0) + \left[1 - p(L_0)\right] \cdot p(L_1) \cdot d(L_1). \tag{15}$$

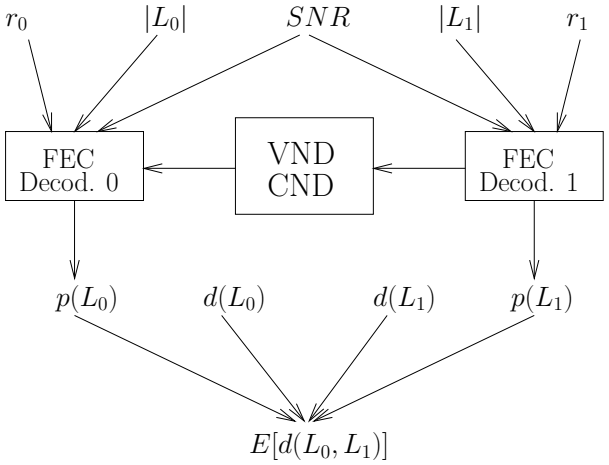


Figure 25: Dependency chains invoked for calculating the expected distortion formulated in Eq. (15).

Based on the above discussions, the expected distortion $E\left[d(L_0,L_1)\right]$ may be calculated by following the dependency chains displayed in Fig. 25, which relies on the parameters SNR, $|L_0|$, $d\left(L_0\right)$, r_0, $|L_1|$, $d\left(L_1\right)$ 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 $d\left(L_0\right)$ and $d\left(L_1\right)$ are gleaned from experiments [78]. Hence the coding rates r_0, r_1 determine the distortion $E\left[d(L_0,L_1)\right]$ expected for the layers L_0, L_1. *Therefore, the objective of the "Code Rate Optimization" block seen in Fig.* 22 *is to find the specific rates* r_0 *and* r_1 *capable of minimizing the expected distortion* $E\left[d(L_0,L_1)\right]$ *of Eq. (15), which may be expressed as*

$$\arg\min_{r_0,r_1}\left\{E\left[d(L_0,L_1)\right]\right\}, \tag{16}$$

subject to the condition of

$$\frac{|L_0|}{r_0} + \frac{|L_1|}{r_1} = \frac{|L_0| + |L_1|}{R}. \tag{17}$$

As illustrated in Fig. 25, Eq. (15) relies on the PER definitions of Eqs. (13) and (14), which cannot be theoretically solved. Below, the solution of Eqs. (13) and (14) will be detailed with the assistance of LUTs and the MI, where LUTs are employed for characterizing the system components of Fig. 22 and MI is employed for numerically quantifying the reliability of the LLRs. Specifically, Section V-D1 introduces the LUTs constructed for solving Eqs. (13) and (14), followed by the PER estimation of the BL and EL in Sections V-D2 and V-D3, respectively. Finally, both the estimated PERs $p(\cdot)$ and the video distortions $d(\cdot)$ will be exploited for the sake of determining the optimized coding rates in Section V-D4.

1) Lookup Tables: The receiver of Fig. 22 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. 25. 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.

In this context, an RSC-coded BPSK-modulated system will be considered for transmissions over a non-dispersive uncorrelated Rayleigh fading channel. Specifically, the authors of [102] modelled BPSK based transmissions over a Rayleigh fading channel as a function of the channel SNR and generated the demodulator's output LLRs, where LLRs are quantified by the MI value, as in [142]. Furthermore, since the MI invoked for quantifying the reliability of the soft information can be applied for Gaussian distributed LLRs [142] generated by arbitrary transceivers, the procedure may be deemed generically applicable, provided that the transceivers generate near-Gaussian distributed LLRs. By contrast, the RSC codec was 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 were quantified by the MI value. The following LUTs are created for modeling the transceiver and the RSC codec:

- $T_b\left(SNR\right)$: The MI value of the LLRs output by the BPSK demodulator for variable channel SNRs. 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 V-F.
- $T_e(SNR, I_s, r)$: The MI value of the extrinsic LLR output of the RSC decoder recorded for variable channel SNRs, 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 V-F.
- $T_p(SNR, I_s, r)$: The PER value associated with the LLRs output by the RSC decoder at diverse channel SNRs, where I_s represents the MI between the *a-priori* LLRs of the systematic information and the correspond-

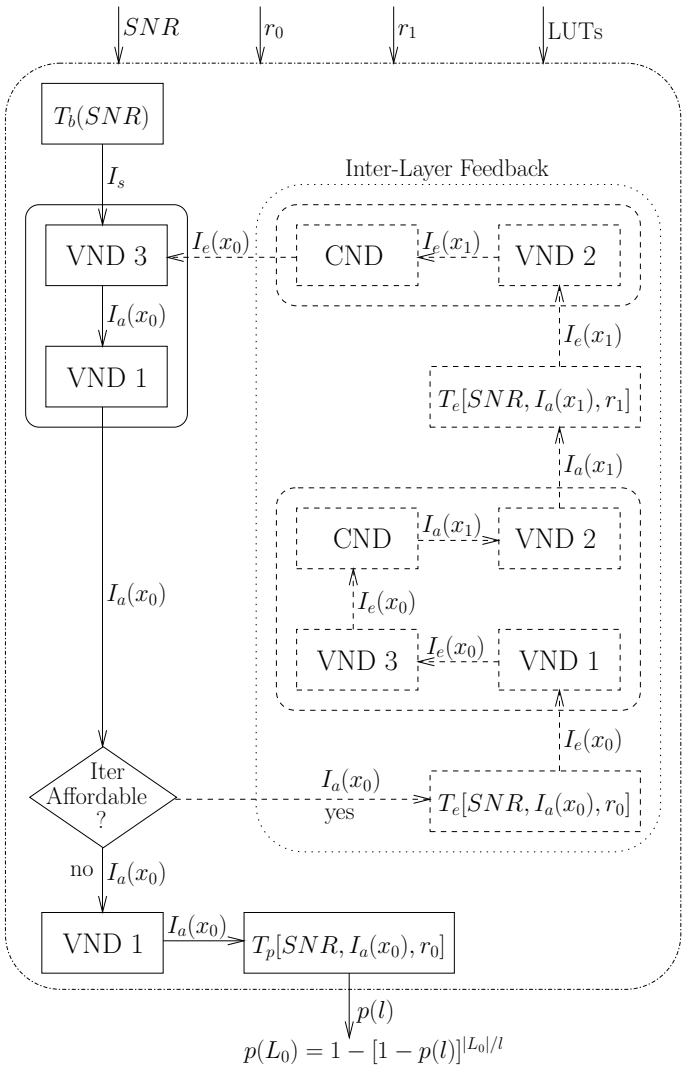
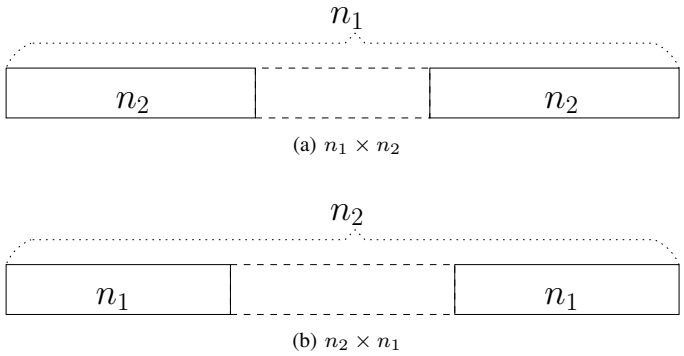


Figure 26: Flow-chart of the PER estimation for the IL-FEC coded BL L_0, where $p(L_0)$ will be employed by the "Code Rate Optimization" block of Fig. 22 for resolving the objective function of Eq. (16).

ing 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) PER Estimation of BL: As shown in Fig. 25, 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. 23. Below we introduce the solution conceived for estimating the PER of Eq. (14) with the aid of the LUTs defined in Section V-D1.

The PER estimation flow-chart of the IL-FEC coded BL is illustrated in Fig. 26, which follows the decoding process of Fig. 22. 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 $T_e\left[SNR, T_b(SNR), r_0\right]$. Then the "VND1-VND3-CND-VND2" process of Fig. 26 is capable



(a) $n_1 \times n_2$

(b) $n_2 \times n_1$

Figure 27: 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.

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

$$I_e(x_1) = T_e\left[SNR, T_b(SNR), r_1\right].$$

Then, following the "VND2-CND-VND3-VND1" process of Fig. 26, 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 $T_p\left[SNR, I_a(x_0), r_0\right]$, which is also shown in Fig. 26. Below, we now detail the method of deriving the PER $p(L_0)$ of the BL L_0 from $p(l) = T_p\left[SNR, I_a(x_0), r_0\right]$.

The burst error distribution of RSC codec has been investigated in [143], 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. 27. Specifically, it may be divided into n_1 packets, 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. 27a the PER $p\left(n_1 \cdot n_2\right)$ may be estimated as

$$p\left(n_1 \cdot n_2\right) = 1 - \left[1 - p(n_2)\right]^{n_1}, \tag{18}$$

where $p(n_2)$ is the PER of the n_1 packets of Fig. 27a. Similarly, we have $p\left(n_1 \cdot n_2\right) = 1 - \left[1 - p(n_1)\right]^{n_2}$ based on Fig. 27b. Then, for arbitrary numerical values of n_1, n_2 we have

$$p\left(n_1\right) = 1 - \left[1 - p(n_2)\right]^{n_1/n_2}. \tag{19}$$

Upon assuming that n_1, n_2 of Eq. (19) are given by $|L_0|$ and l, respectively, the PER $p(L_0)$ of the BL L_0 may be estimated as

$$\begin{aligned} p(L_0) &= f_0\left(SNR, |L_0|, |L_1|, r_0, r_1\right) \\ &= 1 - \left[1 - T_p\left[SNR, I_a(x_0), r_0\right]\right]^{|L_0|/l}, \end{aligned} \tag{20}$$

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

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

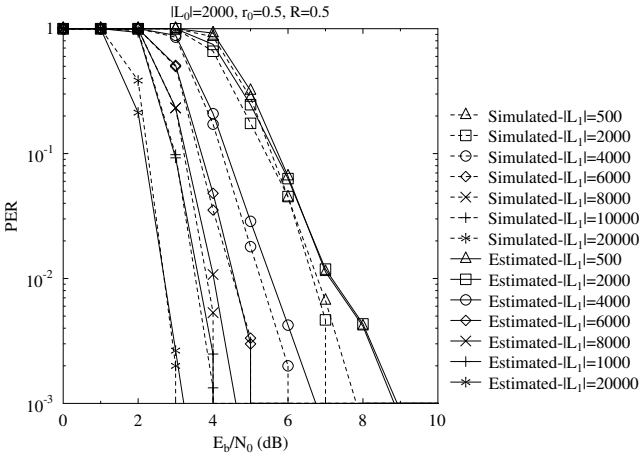


Figure 28: Comparison of the simulated and estimated PER of the BL of Fig. 22, where the parameters $|L_0| = 2000$, $r_0 = 0.5$, $R = 0.5$ are employed.

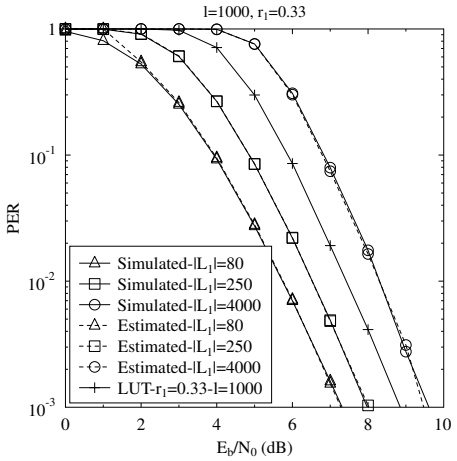


Figure 29: Comparison of the simulated and estimated PER for the EL of Fig. 22, where the parameters $l = 1000$, $r_1 = 0.33$ are employed..

For visualizing the accuracy of the 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. 28 for different SNR values. As observed in Fig. 28, 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 the 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.

3) PER Estimation of EL: When the BL L_0 is correctly decoded, infinite LLRs will be input to the CND 3 of Fig. 22 [2], 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. (20), the PER $p(L_1)$ of the EL L_1 may be finally estimated as

$$\begin{aligned} p(L_1) &= f_1\left(SNR, |L_1|, r_1\right) \\ &= 1 - \left[1 - T_p(SNR, I_s, r_1)\right]^{|L_1|/l}, \end{aligned} \tag{21}$$

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

For visualizing the accuracy of the EL PER estimation, $r_1 = 0.33$ is assumed, while the RSC codec is configured by the generator polynomials of $[1011, 1101, 1111]$. Moreover, $|L_1|$ values ranging from 80 to 4000 are considered, 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. 29. Observe in Fig. 29 that the estimation error is much smaller than that in Fig. 28 due to the fact that the EL is protected by an RSC codec without using IL techniques.

4) Optimized Coding Rates: Based on Sections V-D2 and V-D3, 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 $E\left[d(L_0, L_1)\right]$ imposed by the wireless transmission using rates of r_0 and r_1 may be readily calculated using Eq. (15). In order to find the rates r_0, r_1 minimizing the distortion $E\left[d(L_0, L_1)\right]$ in Eq. (16), a full search of the coding rate space may be performed. Note that r_1 is determined by Eq. (17) for a given r_0, hence the full search is performed in a one-dimensional space to find the optimal r_0.

E. Overheads at the Transmitter

All the optimization operations detailed in Section V-D 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 [2]. 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 V-D4. 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, these overheads will be analyzed in order to characterize the system in more depth.

1) Estimation of $d(\cdot)$: $d(L_i)$ is estimated in a similar manner to the procedure of [78], 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 [79], [83], [87], [101], [144]–[146] may be applied in the system.

2) Generation of LUTs: Three LUTs were generated in the 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 straightforward 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 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 V-A, the MI flows of the VND and CND impose a low computational complexity. Furthermore, in the 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 $[0 : 0.01 : 1]$[8] 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. (17) 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 [147]–[149]. Specifically, the adaptive particle swarm optimization (APSO) technique of [149] 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.

F. Overheads at the Receiver

Below, we discuss the overheads imposed by the IL techniques at the receiver.

1) Complexity: As detailed in Section V-C, the signal-flows are based on low-complexity operations compared to the FEC decoding. When the BL L_0 can be recovered in its own right, only sign-flipping is necessitated for extracting the systematic LLR information of the EL L_1. In short, only a modest extra complexity is imposed, as it will be detailed in Section VII-C.

2) Delay: This technique is implemented using the H.264/AVC video codec, where each video frame may be encoded into a number of layers. Since the IL encoding and decoding process is performed within each frame, no extra delay is imposed by the IL technique. At the receiver, the BL L_0 may be firstly decoded without waiting for the arrival of the EL L_1. In video communications, each video frame can only be reconstructed with the aid of the decoded layers L_0 and L_1. Hence no frame-delay is imposed by the IL techniques.

3) FEC-redundancy: The BL L_0 does not rely on the ELs for its decoding operations and the systematic LLR information of the EL L_1 can be extracted from the received check information y_{01} without any loss, provided that the BL is perfectly decoded. Furthermore, since the transmitted bit sequence x_{01} has the same length as that of the bit sequence x_1, the IL-FEC scheme does not impose any extra FEC redundancy.

[8]The value of I_s ranges from 0 to 1 with a constant step of 0.01.

G. Summary

The family of bit-level IL-FEC coded video schemes was discussed in detail, where the systematic bits of the BL were incorporated into the systematic bits of the ELs using an XOR operation. At the receiver, a sophisticated IL-FEC decoding technique was activated for the sake of attaining an improved system performance.

Furthermore, an optimization technique was detailed in the context of IL-FEC coded layered video transmission over wireless channels, where soft-decoding aided FEC is utilized. The PSNR degradation was minimized by performing a full search of all legitimate coding rate arrangements. This optimization solution may be readily invoked for arbitrary transceivers generating near-Gaussian distributed LLRs, as well as for arbitrary non-iteratively decoded channel codes and for arbitrary layered video transmission arrangements.

VI. EXIT Chart Analysis of IL-FEC Coded Systems

EXIT charts [137], [142] are widely used in the state-of-art for analyzing the performance of FEC codes or systems. In this section, EXIT charts will be utilized for illustrating why the bit-level IL-FEC coded systems [2], [102] are capable of improving the performance of the system. Moreover, an RSC and a turbo code [150] will be employed in this section for providing insights into the benefits of IL-FEC techniques.

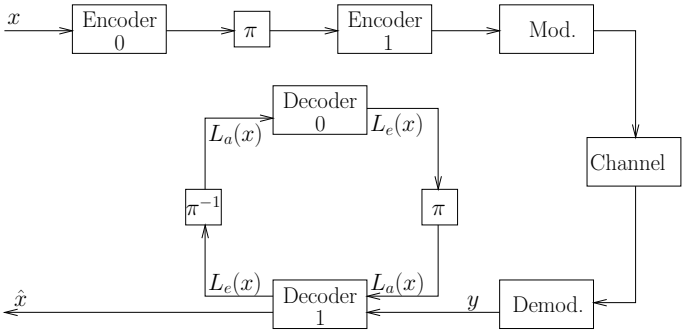


Figure 30: Example of an iterative decoding scheme.

Table VIII: Parameters employed in the systems, where "AA" indicates antenna array.

System Parameters	Value	System Parameters	Value
Codec 0	SBC [2, 6]	Coding rate 0	3/4
Codec 1	RSC[1011, 1101, 1111]	Coding rate 1	1/3
Channel	Narrowband Rayleigh Fading Channel	Overall Coding Rate	1/4
		Modulation	QPSK

The calculation of the MI between the log-likelihood ratio (LLR) and the related information bits was detailed in [142], where the LLRs must be near-Gaussian distributed signals. The EXIT chart can be generated for arbitrary iterative decoding aided receivers. A general iterative decoding scheme

is exemplified in Fig. 30, where the original bitstream x is encoded by the serially concatenated Encoder 0 and Encoder 1. At the receiver, Decoder 0 and Decoder 1 iteratively exchange extrinsic information during the decoding process. More specifically, we may employ a short block code (SBC) [151] as Encoder 0 and an RSC code as Encoder 1 for exemplifying the EXIT chart. In order to augment those concept more explicitly, we carried out the following investigation. The FEC encoded source bits were then mapped to a quadrature phase-shift keying (QPSK) modulator and then they were transmitted over Rayleigh fading channels. All the parameters are listed in Table VIII. The BER curves of different iterations and the related EXIT chart are displayed in Fig. 31. Observe from Fig. 31 that the BER tends to become vanishing low upon increasing the number of iterations at E_b/N_0=2 dB, while the related trajectory is converging to the top right corner at the (1,1) point of the EXIT chart.

A. Scenarios

Table X: Coding rates of RSC and turbo codec error protection arrangements in the system of Fig. 22 for the BL L_0 and the EL L_1. The code-rates were adjusted by variable-rate puncturers.

Error Protection Arrangements	Code Rates		
	L_0	L_1	Average
EEP	0.5	0.5	0.5
UEP1	0.54	0.46	0.5
UEP2	0.47	0.53	0.5

For the sake of simplifying the analysis, we assume that there are two layers: a BL L_0 and an EL L_1. Furthermore, we employed a 1/3 RSC having the generator polynomials $[1011, 1101, 1111]$[9] and a 1/3 turbo codec consisting of two identical RSC codec with the generator of $[111, 101]$. The system parameters used in the simulations are summarized in Table IX. In the ensuing analysis, two layers are considered, where the BL is protected by the IL-FEC codec. Hence, we consider the decoding convergence behavior of the BL. For the sake of analyzing the IL-FEC codec, different error protection arrangements are considered, as shown in Table X.

B. Mutual Information Analysis Using a RSC Codec

Fig. 32 plots the extrinsic MI at the output of the RSC decoder for different E_b/N_0 values for all the codes in Table X. Observe from Fig. 32 that the schemes employing the iterative inter-layer technique always acquire a higher MI value than those dispensing with the IL-FEC technique. For example, the RSC-EEP scheme and RSC-EEP-IL scheme generate 0.91 and 0.975[10] extrinsic information at -8 dB. This improvement is attained by the IL-FEC scheme due to the fact that extra MI is fed back to the BL from the EL. Similar trends can be observed for the system pairs RSC-UEP1 and RSC-UEP1-IL, RSC-UEP2 and RSC-UEP2-IL, as displayed in Fig. 32.

[9]The first polynomial indicates the feedback parameter, while the rest represent the feed-forward parameters. The code rates were adjusted by variable-rate puncturers.

[10]Larger amount of extrinsic information indicates a lower BER [152].

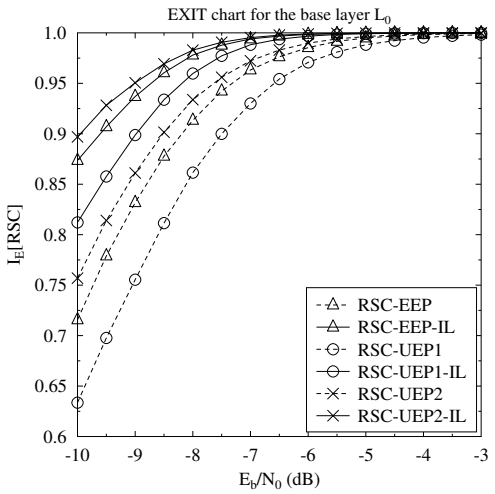


Figure 32: Extrinsic information generated by the RSC decoders for all error protection arrangements of Table X. The schematic of Fig. 22 was used.

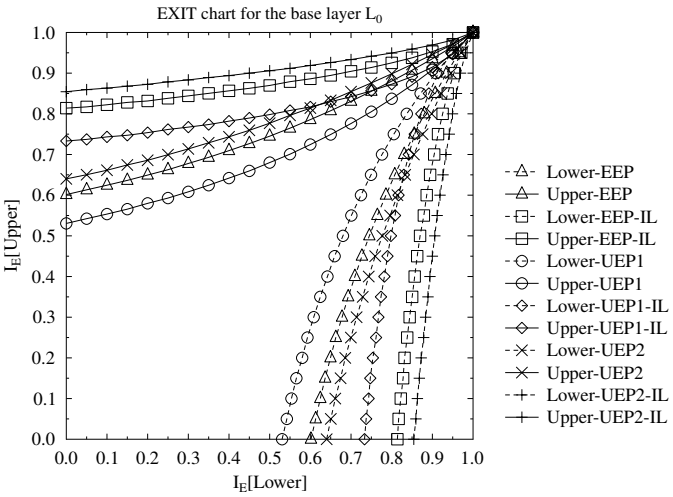


Figure 33: Comparison of the upper RSC and lower RSC EXIT curves at $E_b/N_0 = -8.5$ dB for all error protection arrangements of Table X. The schematic of Fig. 22 was used.

C. EXIT Chart Analysis Using Turbo Codec

Fig. 33 plots the EXIT chart for the turbo coded systems of Fig. 22 using the coding arrangements of Table X. Observe from Table X, the error correction capability of the BL L_0 increases in the order of UEP1, EEP and UEP2, while the width of the open EXIT tunnel increases in the same order, as observed from Fig. 33. Observe in Fig. 33 that at $E_b/N_0 = -8.5$ dB, the IL-turbo coded system has a wider open EXIT tunnel than the system dispensing with the IL-turbo. More explicitly, if we consider the UEP2 aided system, then it is clear from Fig. 33 that the IL-turbo coded system has a wider open EXIT tunnel. In other words, the IL-turbo coded

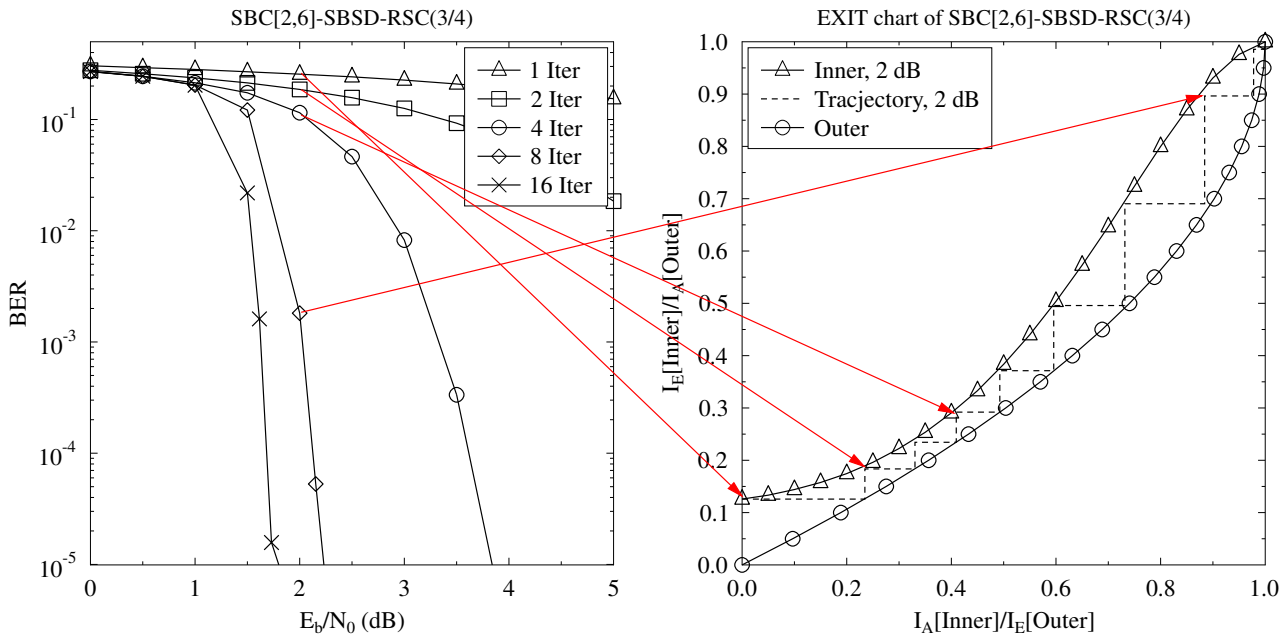


Figure 31: BER curves and the related EXIT chart for transmission over an uncorrelated Rayleigh channel using the parameters of Table VIII and the schematic of Fig. 30.

Table IX: Parameters employed for characterization of the system in Fig. 22,where "AA" indicates antenna array.

System Parameters	Value	System Parameters	Value
FEC 1	RSC-$[1011, 1101, 1111]$	Number of Tx antennas	4
FEC 2	turbo-$[111, 101]$	Elements Per AA	4
Modulation	QPSK	Number of Rx antennas	4
Channel	Narrowband Rayleigh Fading Channel	Overall Coding Rate	1/2

system requires a lower E_b/N_0 than its counterpart dispensing with IL-turbo in order to attain an open tunnel. This implies that the IL-turbo system is capable of attaining a better BER performance for the BL than its counterpart dispensing with IL-turbo coding. The reason for attaining a wider EXIT tunnel by the scheme is due to the fact that extra MI is fed back to the BL from the EL.

An EXIT trajectory comparison of the EEP-turbo system and of the EEP-IL-turbo system is displayed in Fig. 34, which is based on Monte-Carlo simulations. Observe from Fig. 34 that the EEP-IL-turbo system has a wider open tunnel than the EEP-turbo system, as discussed in the previous paragraph. The stair-case-shaped decoding trajectory of the EEP-IL-turbo system reaches the point $(0.93, 0.93)$, while that of the EEP-turbo system is curtailed around $(0.83, 0.83)$ point. Hence the EEP-IL system has a better convergence behavior than the EEP technique, which results in a better BER performance [140]. Observe in Fig. 34 that, although there is an open EXIT tunnel between the curves "Lower-EEP" and "Upper-EEP", the trajectory fails to converge to the$(1, 1)$ point of perfect convergence to a vanishingly low BER due to the fact that we employ short interleavers. The length of the interleaver is constrained in real-time video streaming application for the sake of delay control. Therefore, it can be inferred from Figs. 33 and 34 that employing the IL-FEC coding results in a better

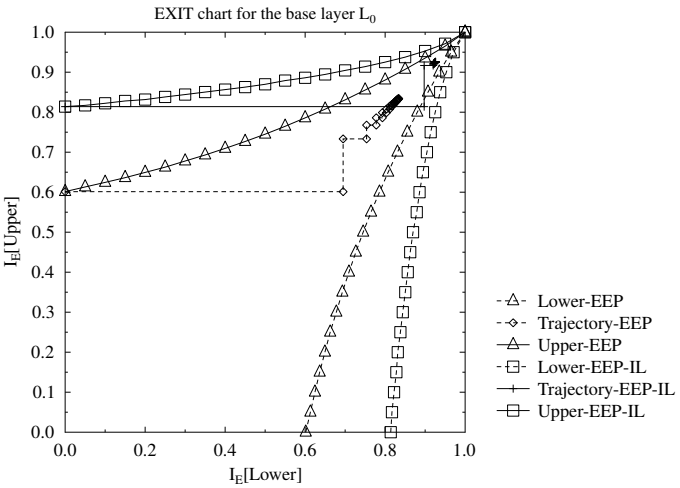


Figure 34: Trajectory comparison of the upper RSC and lower RSC EXIT curves at $E_b/N_0 = -8.5$ dB for the systems employing EEP. The schematic of Fig. 22 was used.

BER performance, which is demonstrated by the wider open EXIT tunnel shown in Figs. 33 and 34.

D. Summary

EXIT charts were utilized for illustrating why bit-level IL-FEC coded systems are capable of improving the attainable system performance. As displayed in Fig. 33, the IL-FEC schemes exhibit a wider open EXIT tunnel compared to the remaining schemes. This is due to the fact that the ELs feed back extra MI concerning the BL, which may be utilized by the decoder of the BL for the sake of correcting the errors.

VII. Results of IL-FEC Using Partitioned H.264: Empirical Coding Rates

Let us continue by benchmarking the IL-FEC system against the traditional UEP aided FEC system using a RSC codec. Additionally, in order to meet the challenging performance requirements of bandwidth-constrained environments, a MIMO system, namely the sophisticated Layered Steered Space-Time Codes (LSSTCLSSTC) of [140], [153] are employed in the transceivers. The LSSTCs are capable of providing both a diversity gain for the sake of achieving a high BER performance in mobile environments as well as attaining a multiplexing gain in order to maintain a high data rate. The system parameters are listed in Table IX.

Two 30-frame video sequences, namely the Foreman and Football clips, represented in (352×288)-pixel CIF and 4:2:0 YUV format were encoded using the JM/AVC 15.1 H.264/AVC reference video codec operated in its data partitioning aided mode. The video scanning rates expressed in FPS were 30 and 15 for the Foreman and Football sequences, respectively. The motion-copy[11], based error concealment tool built into H.264/AVC reference codec was employed for the sake of combating the effects of channel impairments. Moreover, the H.264/AVC encoder was configured to generate fixed-byte[12] slices, as defined in [9]. Both of the 30-frame video sequences were encoded into an I-frame, followed by 29 P-frames. The bi-directionally predicted frame (B-frame) was disabled due to the fact that it relies on both previous and future frames for decoding, which may introduce more error propagation as well as additional delay. All the above configurations jointly result in a bitrate of 655 kbps and an error-free PSNR of 38.4 dB for the Foreman sequence. On the other hand, the coded Football bitstream has a bitrate of 1522 kbps and an error-free PSNR of 37.6 dB. We employed the Foreman and Football sequences in order to show the suitability of the scheme for the transmission of both low-motion and high-motion video. The parameters of the employed sequences are shown in Table XI.

[11]When the information of a MB is lost, the motion vector of this MB may be copied or estimated from its adjacent MBs or previously decoded reference frames. Then, the MB may be reconstructed using the estimated motion vector.

[12]In this mode, the H.264/AVC codec will endeavor to encode a frame into multiple slices, each having a fixed number of bytes.

Table XI: The parameters of the video sequences employed for the characterization of the system in Fig. 22.

	Football	Foreman
Representation	YUV 4:2:0	YUV 4:2:0
Format	CIF	CIF
Bits Per Pixel	8	8
FPS	15	30
Number of Frames	30	30
Video Codec	H.264 DP	H.264 DP
Bitrate	1522 kbps	655 kbps
Error-Free PSNR	37.6 dB	38.4 dB
Error Concealment	Motion-Copy	Motion-Copy

The H.264-compressed bitstream was FEC encoded and transmitted on a NALU [9] basis, which is the smallest element to be used by the source decoder. At the receiver, each error-infested NALU must be dropped by the video decoder, if errors are detected by the CRC check. All experiments were repeated 100 times for the sake of generating smooth performance curves.

Below, we will firstly describe the error-protection arrangements in Section VII-A. Then we will characterize the attainable BER versus channel SNR performance and PSNR versus channel SNR performance employing a lower-complexity RSC codec in Section VII-B. Finally, in Section VII-C we will quantify the system's computational complexity by counting the number of decoding operations executed.

A. Error Protection Arrangements

Table XII: Coding rates of different error protection arrangements used in the system of Fig. 22 for the Football/Foreman sequence. The code-rates were adjusted by variable-rate puncturers.

Error Protection Arrangements	Code Rates			
	Type A	Type B	Type C	Average
EEP	0.5/0.5	0.5/0.5	0.5/0.5	0.5/0.5
UEP1	0.35/0.40	0.57/0.65	0.57/0.65	0.5/0.5
UEP2	0.45/0.55	0.52/0.46	0.52/0.46	0.5/0.5
UEP3	0.65/0.60	0.47/0.43	0.47/0.43	0.5/0.5
UEP4	0.75/0.70	0.45/0.39	0.45/0.39	0.5/0.5
UEP5	0.85/0.80	0.44/0.37	0.44/0.37	0.5/0.5
UEP6	0.95/0.90	0.43/0.35	0.43/0.35	0.5/0.5

In the simulations, we employ the overall coding rate[13] of 1/2 for both EEP and UEP schemes. For each compressed bitstream, all NALUs were scanned to calculate the total number of bits for the A, B, and C partitions. Let us assume that the A, B and C partitions have a total N_a, N_b and N_c bits, respectively. The A, B, C streams have coding rates of r_a, r_b and r_c[14], respectively. Then the following equation must be satisfied for the sake of guaranteeing that the overall coding rate remains 1/2:

$$2 \times (N_a + N_b + N_c) = \frac{N_a}{r_a} + \frac{N_b}{r_b} + \frac{N_c}{r_c}. \quad (22)$$

Again, the A stream is the most important layer, while the B and type C bitstreams are the ELs, where the bitstream B and

[13]Arbitrary overall coding rates such as 2/3, 1/3, 1/4, etc. can be readily applied by changing the channel codec parameters and the puncturers.

[14]In Fig. 22, we used r_0 and r_1 for the scenario of two layers.

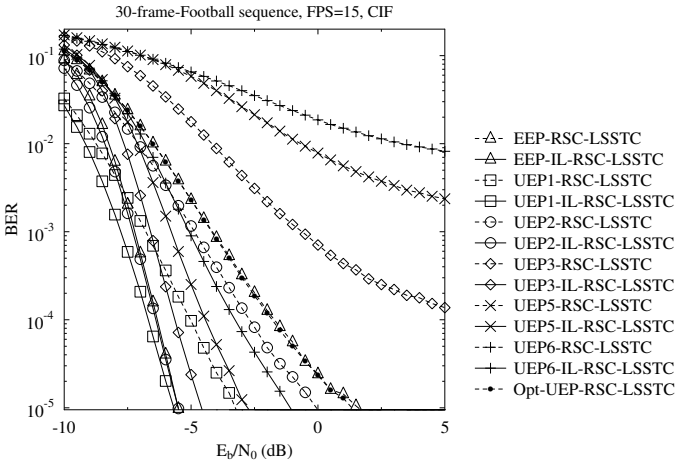


Figure 35: BER versus E_b/N_0 performance for the A partition of the *Football* sequence, including the RSC coding schemes of Table XII and the Opt-UEP-RSC-LSSTC [78]. The schematic of Fig. 22 was used.

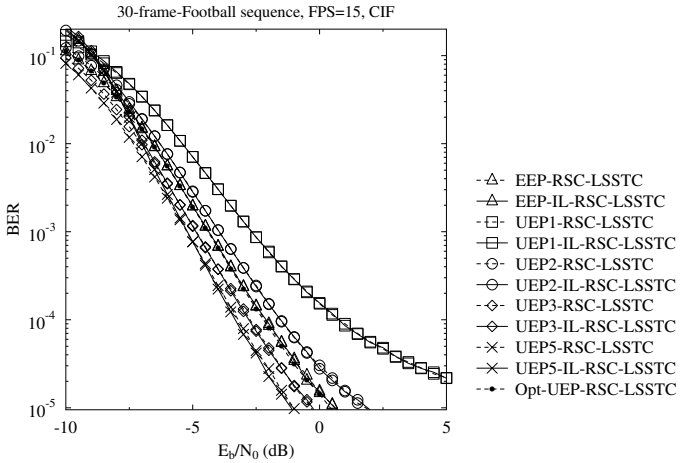


Figure 36: BER versus E_b/N_0 performance for the B partition of the *Football* sequence, including the RSC coding schemes of Table XII and the Opt-UEP-RSC-LSSTC [78]. The schematic of Fig. 22 was used.

C are similarly important. Hence in all the error protection arrangements we have $r_b = r_c$. More specifically, we first select a specific value to r_a, then the value of $r_b = r_c$ was calculated as follows:

$$r_b = \frac{N_b + N_c}{2 \times (N_a + N_b + N_c) - \frac{N_a}{r_a}}. \tag{23}$$

Note that the total number of bits for each partitions of the different video sequences may be different, which results in different protection arrangements. For providing insights into the benefits of IL-FEC schemes, instead of using the "Code Rate Optimization" block of Fig. 22, we empirically selected a number of error protection arrangements, which are listed in Table XII, which may be readily combined with arbitrary EEP or UEP schemes, noting that variable-rate puncturers were designed and employed to achieve a specific coding rate. The "Code Rate Optimization" based system performance will be benchmarked in Section VIII.

B. System Performance using RSC Codec

In this section, we benchmark the system using the RSC codec of Table IX. All the error protection arrangements of Section VII-A will be utilized. Furthermore, in [78] an UEP algorithm was proposed, which the authors of [78] referred to as the optimal UEP. We used this scheme as a benchmarker, which we refer to as the Opt-UEP-RSC-LSSTC arrangement.

The BER curves of the A partition in the Football sequence are displayed in Fig. 35, where the performance of the error protection schemes of Table XII are illustrated. Observe in Fig. 35 that the schemes using the IL-RSC codec achieve a reduced BER compared to their benchmarkers. Specifically, the EEP-IL-RSC-LSSTC scheme outperforms the EEP-RSC-LSSTC benchmarker by about 7.2 dB at a BER of 10^{-5}. Furthermore, among all the error protection arrangements, the UEP1-IL-RSC-LSSTC scheme achieves the best BER performance due to the strong error protection assigned for the A partition. Hence, we may conclude that the UEP aided IL-RSC schemes

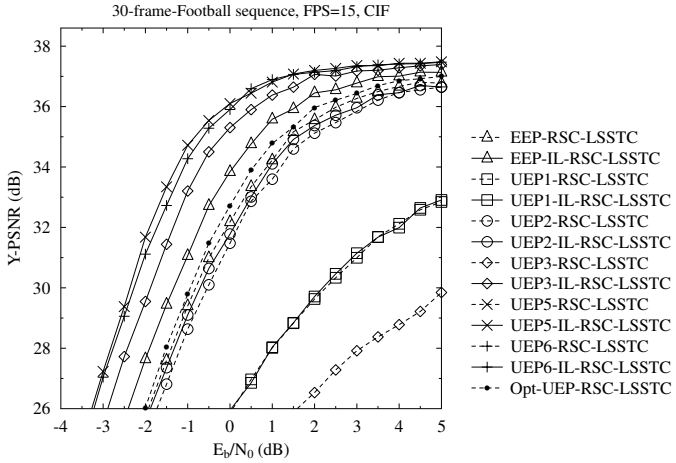


Figure 37: PSNR versus E_b/N_0 performance for the *Football* sequence, including the RSC coding schemes of Table XII and the Opt-UEP-RSC-LSSTC [78]. The schematic of Fig. 22 was used.

are capable of providing an improved system performance compared to the traditional UEP aided RSC codec. On the other hand, the Opt-UEP-RSC-LSSTC system achieves similar BER performance to that of the EEP-RSC-LSSTC scheme.

The BER versus E_b/N_0 performance of the B partition for the Football sequence is presented in Fig. 36. Similar trends were observed for the C partition. Observe in Fig. 36 that the performance of the schemes using IL-RSC is slightly worse than that of their benchmarkers. This is due to the fact that more errors may be introduced into the B partition, when the A partition cannot be correctly decoded. In this scenario the B partition must be dropped in the traditional UEP aided RSC-LSSTC schemes. Hence the error propagation to the B partition does not further degrade the situation.

The PSNR versus E_b/N_0 performance recorded for the Football sequence is shown in Fig. 37, where we observe that

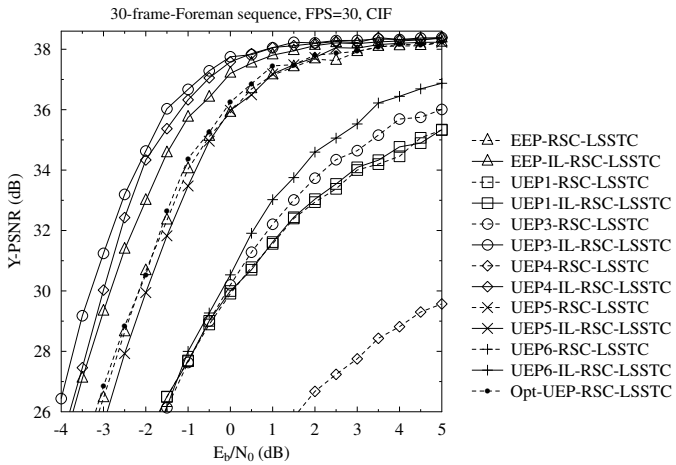


Figure 38: PSNR versus E_b/N_0 performance for the *Foreman* sequence, including the RSC coding schemes of Table XII and the Opt-UEP-RSC-LSSTC [78]. The schematic of Fig. 22 was used.

the EEP-RSC-LSSTC scheme achieves the best performance among all the systems without IL techniques, because the A partition carries only the video header information and fails to assist the H.264/AVC decoder in concealing the residual errors, when the B and C partitions are corrupted. Furthermore, the systems using the IL-RSC-LSSTC model outperform their corresponding benchmarkers. Specifically, the UEP5-IL-RSC-LSSTC constitutes the best protection arrangement among all IL-RSC schemes, which achieves a power reduction of about 3 dB compared to the EEP-RSC-LSSTC scheme at a PSNR of 36 dB. Alternatively, about 3.7 dB of PSNR video quality improvement may be observed at a channel SNR of 0 dB. On the other hand, the Opt-UEP-RSC-LSSTC system dispensing with the IL technique slightly outperforms the EEP-RSC-LSSTC scheme, namely by a power reduction of about 0.5 dB at a PSNR of 36 dB. The UEP5-IL-RSC-LSSTC substantially outperforms the Opt-UEP-RSC-LSSTC arrangement, namely by a power reduction of about 2.5 dB at a PSNR of 36 dB or alternatively, about 3.4 dB of PSNR video quality improvement may be observed at an E_b/N_0 of 0 dB. A subjective comparison of the UEP5-IL-RSC-LSSTC and EEP-RSC-LSSTC arrangements for the Football sequence is presented in Fig. 39.

For providing further insights for video scenes having different motion-activity, the PSNR versus E_b/N_0 performance of the IL-RSC-LSSTC model is presented in Fig. 38 using the Foreman sequence, when employing the protection arrangements of Table XII. Similar to the Football sequence, the traditional UEP technique can hardly improve the reconstructed video quality by allocating more FEC redundancy to the more important layers. By contrast, about 2 dB of power reduction is achieved by the UEP3-IL-RSC-LSSTC arrangement compared to the EEP-RSC-LSSTC scheme at a PSNR of 37 dB. Alternatively, about 3.2 dB of PSNR video quality improvement may be observed at a channel SNR of -1 dB. Similar to the Football sequence, a limited gain can be observed for the Opt-UEP-RSC-LSSTC system compared to the EEP-RSC-LSSTC scheme, while the UEP5-IL-RSC-LSSTC substantially outperforms the Opt-UEP-RSC-LSSTC, namely by about 1.8 dB at a PSNR of 37 dB. A subjective comparison of the UEP3-IL-RSC-LSSTC and EEP-RSC-LSSTC arrangements for the Foreman sequence is presented in Fig. 39.

We may conclude from the above discussion that the A partition should be assigned a code-rate of 0.85 and 0.60 for the Football and Foreman sequence, respectively, for the sake of achieving the best overall system performance, when employing the RSC codec, which contradicts to the traditional UEP strategy. The main reason for this is that the inter-layer aided RSC decoder can still successfully recover the weaker protected A partition relying on the extrinsic information fed back from the B and C partitions with the aid of inter-layer decoding, because B and C are more strongly protected than the A partition.

C. Complexity Analysis

In order to provide insights into the complexity of the scheme, we benchmark the complexity of the IL-FEC-LSSTC scheme using the RSC codec in Fig. 40. We emphasize that if the A partition was corrupted, the corresponding complexity imposed by the B and C partitions was not taken into account, since they cannot be utilized by the video decoder in this case. Therefore, the complexity of both the IL-FEC-LSSTC system and of the benchmarkers is directly proportional to the E_b/N_0 value. Furthermore, in the simulations each NALU was encoded by the FEC as a single packet. The total computational complexity is dominated by that of FEC decoding. Hence, the total number of FEC decoding operations substantially affects the system's complexity, which was hence used for comparing the system's complexity. The y-axis of Fig. 40 represents the average number of FEC decoding operations per NALU, which was averaged over 2221 NALUs in the H.264/AVC encoded Football bitstream for the sake of statistical relevance, where again each NALU was encoded as a single packet in the experiments.

Observe from Fig. 40 that each curve of the IL-RSC-LSSTC schemes may be divided into two regions, where the complexity of the systems increases and decreases upon the increasing E_b/N_0. For example, the curve of the UEP3-IL-RSC-LSSTC scheme can be split at E_b/N_0 of about -6.5 dB. Specifically, in the E_b/N_0 region of $[-10, -6.5]$ dB, the complexity of the UEP3-IL-RSC-LSSTC scheme increases upon increasing the E_b/N_0 value. This is due to the fact that the IL decoding technique was activated frequently for assisting the decoding of A partition. By contrast, for higher E_b/N_0 values the A partition is more likely to be recovered with the aid of the IL technique, which in turn results in decoding the B and C partitions more than once. In the E_b/N_0 region of $[-6.5, 5]$ dB, the complexity of the UEP3-IL-RSC-LSSTC scheme decreases upon increasing E_b/N_0 value. The reason for this phenomenon is that the IL decoding technique is less frequently activated, when the A partition is more likely to be perfectly decoded in its own right at higher E_b/N_0 values. Moreover, the complexity of all the

Figure 39: Video comparison at $E_b/N_0 = -2.5$ dB for the Football and Foreman sequences. The first column indicates the original frames. The second column indicates the EEP-IL-RSC-LSSTC decoded frames. The third column indicates the Opt-UEP-RSC-LSSTC [78] decoded frames. The fourth column represents the UEP5-IL-RSC-LSSTC and UEP3-IL-RSC-LSSTC decoded frames for the Football and Foreman sequences, respectively.

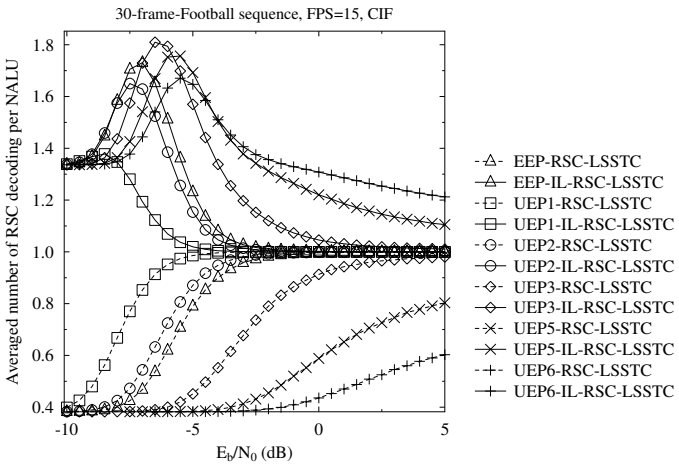


Figure 40: Complexity comparison of the Opt-UEP-RSC-LSSTC system, the IL-RSC-LSSTC schemes and the classic RSC-LSSTC schemes for the error protection arrangements of Table XII for the *Football* sequence. The schematic of Fig. 22 was used.

RSC-LSSTC schemes increases upon increasing E_b/N_0. This may be attributed to the fact that at lower E_b/N_0 the B and C partition were more likely to be dropped by the decoder due to the corruption of the A partition. Since low E_b/N_0 results in unacceptable video quality, here we only focus on higher E_b/N_0 region. More specifically, the UEP5-IL-RSC-LSSTC scheme achieves E_b/N_0 gains of 3 dB and 2.5 dB by imposing about 21% higher complexity than the EEP-RSC-LSSTC and Opt-UEP-RSC-LSSTC schemes at a video quality of 36 dB, respectively. Alternatively, the UEP5-IL-RSC-LSSTC has PSNR gains of 3.7 dB and 3.4 dB at the cost of a 21% complexity increase compared to the EEP-RSC-LSSTC and Opt-UEP-RSC-LSSTC schemes at an E_b/N_0 of 0 dB, respectively.

D. Summary

We may arrive at the following conclusions from Section VII:

1) In the RSC based systems, the most important layer should be assigned less redundancy than partitions B and C for the sake of achieving the best overall system performance, which is in contrast to the traditional UEP strategy. For example, the system arrangement having channel coding rates of 0.85, 0.44 and 0.44 for the A, B and C partitions, respectively, achieves the best system performance when employing the RSC code for the transmission of the Football sequence.
2) As jointly observed from Fig. 37 of Section VII-B and Fig. 40 of VII-C, the proposed IL coding technique is capable of achieving 2.5 dB of E_b/N_0 again or alternatively, 3.4 dB of PSNR gain over the traditional UEP technique at the cost of a 21% complexity increase.

VIII. Results of IL-FEC Using Scalable Video Coding: Optimized Coding Rates

In comparison to the empirical coding rates derived for the DP-H.264 scheme in Section VII, we characterize the system's performance based on the rates r_0 and r_1 determined by the "Code Rate Optimization" block of Fig. 22. We refer to this scheme as the Optimized-IL-RSC for simplicity. We benchmark the 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

Table XIII: The parameters of the system of Fig. 22.

System Parameters	Value
FEC	RSC$[1011, 1101, 1111]$
Overall Coding rate	1/2
Modulation	BPSK
Channel	Uncorrelated Rayleigh Fading Channel

Table XIV: The parameters of the video sequences employed in the evaluation of the system of Fig. 22.

	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

Rayleigh fading channels. These parameters are listed in Table XIII. The 30-frame Foreman and Football video clips represented in (176×144)-pixel 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 the IL-FEC 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 I-frame was inserted every 15 frames. Correspondingly, both of the two video sequences were encoded into GOPs, consisting of an I-frame, followed by 14 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 the JSVM configuration. Additionally, only the MGS [20], [154] feature is 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 QP of 40, 32 and 24, respectively. 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. These parameters of the video sequences are listed in Table XIV.

Moreover, each SVC H.264/AVC-compressed bitstream was channel encoded and transmitted on a NALU by NALU [9] 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 the experiments, the compressed bitstreams were transmitted 300 times in order to generate statistically sound performance curves.

Table XV: Example of the LUTs used by the system of Fig. 22.

(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	$p(l)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

A. Off-line LUT Generation

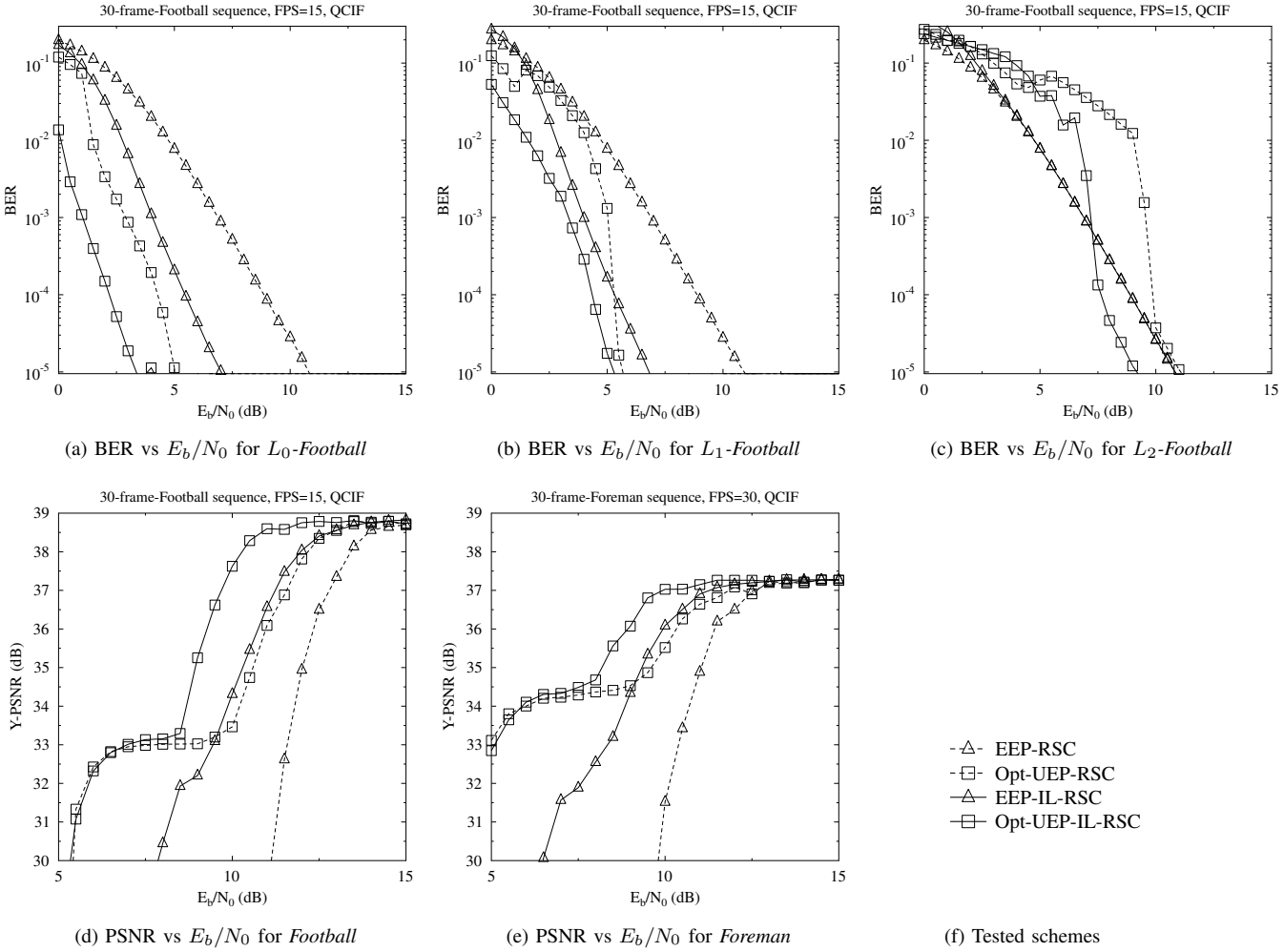
In experiments, the vectors of $[0 : 0.5 : 15]$, $[0 : 0.01 : 1]$, $[0.33 : 0.02 : 1]$[15] are utilized for the variables SNR, I_s, r, respectively, for generating the LUTs, which result in $n_{snr} = 31$, $n_I = 101$, $n_r = 33$. Firstly, we recorded the MI values of the LLRs output by the BPSK demodulator for the SNR values of $[0 : 0.5 : 15]$ for the sake of generating the LUT $T_b\,(SNR)$. Then for each SNR, the MIMI 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 LUTLUTs 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 $\left(C \times n_r^2 = C \cdot 1089\right)$, thereby finding the optimized coding rates of r_0, r_1 and r_2. A number of entries of the LUTs generated for system are displayed in Table XV.

B. System Performance

In this section, we benchmark 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 [78] 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. 41a, 41b and 41c, respectively, where the Opt-UEP-IL-RSC system always outperforms the Opt-UEP-RSC system. Moreover, as seen in Fig. 41c, the EEP schemes outperform the Opt-UEP-IL-RSC and Opt-UEP-RSC

[15]These values can be stored as floats in 8 bytes each.



(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

Figure 41: BER and PSNR versus E_b/N_0 performance comparison of the proposed system and the benchmarkers, namely the EEP-RSC scheme, the Opt-UEP-RSC scheme [78], the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme for the *Football* and *Foreman* sequences. The schematic of Fig. 22 configured with the parameters in Tables XIII and XIV was used.

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.

The PSNR versus E_b/N_0 performance recorded for the Football sequence is displayed in Fig. 41d, 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 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. 41c. 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 benchmarkers recorded for the Football sequence is presented in Fig. 42.

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. 41e 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.

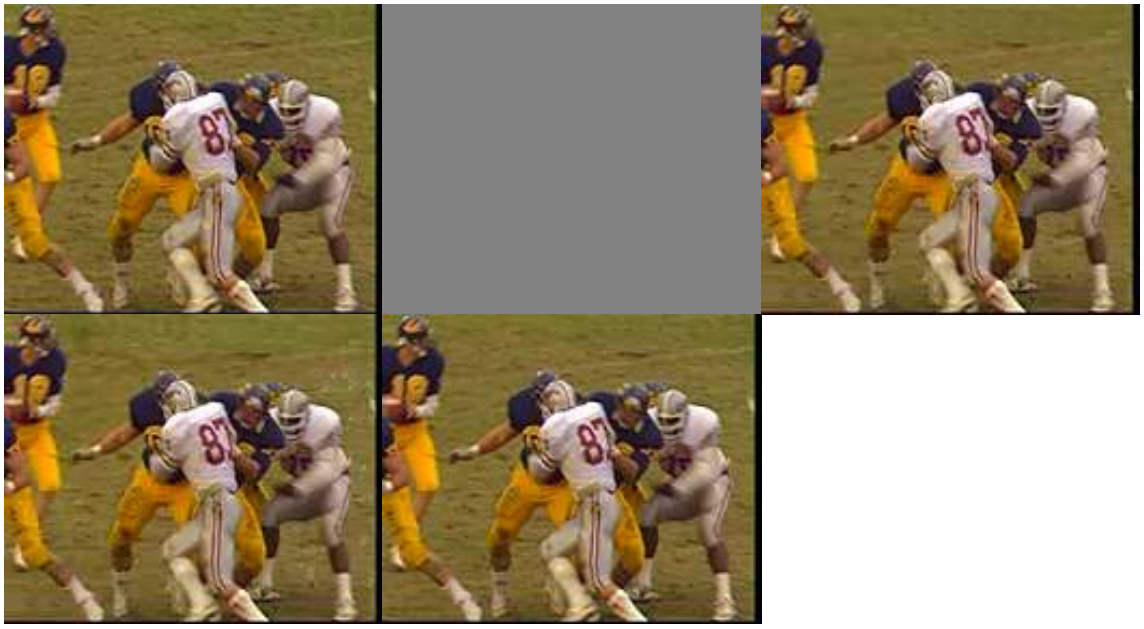


Figure 42: Comparison of frames at E_b/N_0 of 10 dB for the *Football* sequences. The five columns (from left to right, from top to bottom) indicate frames of the original video, the EEP-RSC scheme, the Opt-UEP-RSC [78] scheme, the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme, respectively. The schematic of Fig. 22 configured with the parameters in Tables XIII and XIV was used.

C. Optimized Coding Rates

The optimized coding rates found by the Opt-UEP-IL regime for the Football sequence are shown in Fig. 43 in comparison to the benchmarkers. Specifically, the y axis of Fig. 43 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. 43 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. Moreover, similar trends are observed for the Foreman sequence.

D. Summary

We may arrive at the following conclusions from Section VIII:

1) Compared to the Opt-UEP system, the Opt-UEP-IL assigns a higher coding rate to the most important layer for the sake of achieving the best possible system performance, as displayed in Fig. 43. This is due to the fact that the protection assigned to the ELs also protects the BL. This conclusion is consistent with the conclusions of Section VII-D.
2) As observed from Fig. 37, the proposed Opt-UEP-IL system outperform the traditional UEP aided system by about 1.9 dB of E_b/N_0 or 3.3 dB of PSNR.

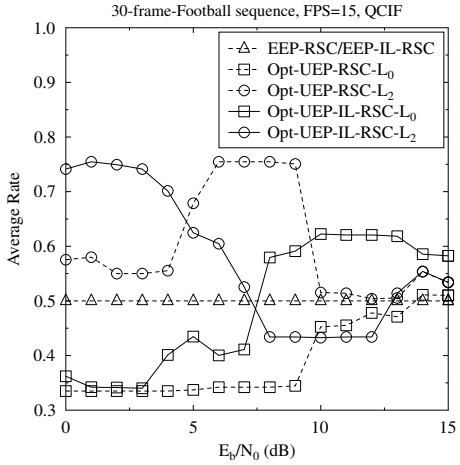


Figure 43: Coding rates comparison of the benchmarkers, namely the EEP-RSC scheme, the Opt-UEP-RSC [78] scheme, the EEP-IL-RSC scheme and the Opt-UEP-IL-RSC scheme, for the *Football* sequence. The schematic of Fig. 22 configured with the parameters in Tables XIII and XIV was used.

IX. Conclusions

In this concluding chapter, a summary of the tutorial and the main findings of our investigations will be presented.

A. Summary

In this tutorial, we commenced with a brief historical perspective on video compression, leading up to the evolution of layered video compression. Then, we reviewed the state-of-the-art in layered video techniques, followed by the rudiments

of LA-FEC and IL-FEC coded layered video streaming, with a special emphasis on soft-decoded bit-level IL-FEC techniques. This research is still widely open.

We detailed the bit-level IL-FEC coded video scheme, where the information of the BL was incorporated into the systematic bits of the ELs with the aid of an XOR operation. When the BL can be successfully decoded in its own right, the systematic bits of the ELs can be extracted by flipping the sign of the appropriately generated check information received. However, when the BL cannot be correctly decoded without the assistance of the ELs, the IL-FEC decoding philosophy exchanging information between the BL and the ELs will be activated to assist in decoding the BL. Furthermore, efficient techniques of finding the optimized coding rates of the individual coded bitstreams "on-the-fly" at the transmitter, were discussed for the sake of optimizing the IL-FEC coded system's performance. Specifically, the coding rates were found for achieving the minimum video quality distortion with the aid of the MI between the LLRs and the corresponding video bits.

EXIT charts were invoked for analyzing the benefits of the IL-FEC scheme. Finally, the partitioned H.264/AVC video codec and the widely employed SVC-H.264 codec were invoked for the sake of benchmarking the IL-FEC system.

B. Design Guidelines

Based on this tutorial, we summarize our general design guidelines conceived for IL-FEC aided layered video communication schemes:

1) **Delay**
   Some of the layered video applications impose strict constraints on the tolerable delay, such as lip-synchronized tele-conferencing, while some others may be insensitive to delay, such as video storage. *Hence, we have to carefully control the delay imposed, when designing real-time, interactive video systems.*
2) **Complexity**
   Since the system imposes extra processing in comparison to the conventional EEP systems, *the attainable performance improvements should be carefully considered in the light of the complexity imposed.*
3) **UEP of layers in a GOP**
   In layered video coding, we typically assume that each GOP contains F number of frames, each of which is encoded into L layers. The F frames of each GOP tend to have unequal importance. Furthermore, the L layers of each frame also tend to have a different importance. Hence, each GOP contains $(F \times L)$ different-significance layers. It is intuitive to aim for allocating more transmit/coding power to the more important layers. *Hence, we have to carefully optimize the power allocation among the $(F \times L)$ layers for the sake of improving the attainable video quality in the face of transmission errors.*
4) **Inter-layer dependency**
   In layered video streams, each layer may depend on other layers for decoding. Hence this characteristic may be interpreted as a type of redundancy within the bit-streams. It is beneficial to exploit this characteristic for the sake of achieving an improved video quality. *Therefore the corresponding design lesson for layered video transmission is to jointly encode multiple layers for the sake of exploiting their dependency at the receiver in order to improve the robustness against transmission errors.*
5) **LUT Aided Design**
   It is difficult to theoretically model a complex system component, such as RSC codes in a wireless system. *We may treat this component as a black box and model its BER vs SNR curve using LUTs for the sake of holistically optimizing the entire system.*
6) **EXIT Chart**
   EXIT charts are widely employed for analyzing wireless transmission systems. *Hence we may also use EXIT chart for analyzing and optimizing video systems.*

GLOSSARY

AA	Antennas Array.
ACK	Acknowledgement.
ACS	Adaptive Channel Selection.
APSO	Additive Particle Swarm Optimization.
ASP	Advanced Simple Profile.
AVC	Advanced Video Coding.
AVS	Advanced Video Standard.
BEC	Binary Erasure Channel.
BER	Bit Error Ratio.
BL	Base Layer.
BOP	Block of Packets.
BP	Belief Propagation.
BPSK	Binary Phase-Shift Keying.
CBP	Coded Block Pattern.
CC	Convolutional Codes.
CGS	Coarse Grain Scalability.
CIF	Common Intermediate Format.
CND	Check Node Decoder.
CodedCM	Coded Cooperative Multicast.
CQI	Channel Quality Information.
CR	Code Rate.
CRC	Cyclic Redundancy Check.
DCT	Discrete Cosine Transform.
DID	the parameter dependency_id in scalable video stream.
DP	Data Partitioning.
DPCM	Differential Pulse Code Modulation.
DV	Digital Video.
DVC	Distributed Video Coding.
DVD	Digital Versatile Disc.
EEP	Equal Error Protection.
EL	Enhancement Layer.
EWF	Expanded Window Fountain.
EXIT	EXtrinsic Information Transfer.
FEC	Forward Error Correction.
FPS	Frames Per Second.
GE	Gilbert-Elliot.
GOP	Group Of Pictures.
HD	High Definition.
HEVC	High-Efficiency Video Coding.
HNC	Hierarchical Network Coding.
HQAM	Hierarchical Quadrature Amplitude Modulation.
I-frame	Intra-coded frame.
IEC	International Electrotechnical Commission.
IL	Inter-Layer.
IL-FEC	Inter-Layer Forward Error Correction.
IP	Internet Protocol.
IPTV	Internet Protocol TeleVision.
ISDN	Integrated Services Digital Networks.
ISO	International Organization for Standardization.
ITU	International Telecommunication Union.

ITU-T	International Telecommunication Union Telecommunication standardization sector.
JPEG	Joint Photographic Experts Group.
JSVM	Joint Scalable Video Model.
JVT	Joint Video Team.
kbps	KBits Per Second.
LA	Layer-Aware.
LA-FEC	Layer-Aware Forward Error Correction.
LAN	Local Area Network.
LDPC	Low-Density Parity-Check.
LLR	Log-Likelihood Ratio.
LR-PET	Limited-Retransmission Priority Encoding Transmission.
LSSTC	Layered Steered Space-time Codes.
LT	Luby Transform.
LUT	LookUp Table.
LW-EZEP	Layer-Weighted Expected Zone of Error Propagation.
M-UEP	Multistream Unequal Error Protection.
MAC	Media Access Control.
MB	Macro Block.
MCER	Motion Compensated Error Residual.
MDS	Maximum Distance Separable.
MGM	Multi-Generation Mixing.
MGS	Medium Grain Scalability.
MI	Mutual Information.
MIMO	Multiple Input Multiple Output.
MPEG	Moving Picture Experts Group.
MVC	Multiview Video Coding.
MVP	MultiView Profile.
NALU	Network Abstraction Layer Unit.
NTSC	National Television System Committee.
OFDM	Orthogonal Frequency-Division Multiplexing.
OFDMA	Orthogonal Frequency-Division Multiplexing Access.
OppCM	Opportunistic Cooperative Multicast.
OSI	Open Systems Interconnection.
P-frame	Predicted frame.
PAL	Phase Alternating Line.
PC	Personal Computer.
PEC	Packet Erasure Channel.
PER	Packet Error Ratio.
PET	Priority Encoding Transmission.
PHY	Physical.
PLR	Packet Loss Ratio.
PSNR	Peak Signal-to-Noise Ratio.
PSO	Particle Swarm Optimization.
QCIF	Quarter Common Intermediate Format.
QID	the parameter quality_id in scalable video stream.
QoE	Quality of Experience.
QoS	Quality of Service.
QP	Quantization Parameter.
QPSK	Quadrature Phase-Shift Keying.
RF	Radio-Frequency.
RFC	Request for Comments.
RS	Reed-Solomon.
RSC	Recursive Systematic Convolutional.
SBC	Short Block Code.
SHVC	Scalable High Efficiency Video Coding.
SI	Side Information.
SMPTE	Society of Motion Picture and Television Engineers.
SNR	Signal-to-Noise Ratio.
SP	Simple Profile.
SPIHT	Set Partitioning relying on Hierarchical Trees.
SVC	Scalable Video Coding.
TCP	Transmission Control Protocol.
TID	the parameter temporal_id in scalable video stream.
TS	Temporal Scalability.
TV	TeleVision.
UDE	Unequal error protection Density Evolution.
UEP	Unequal Error Protection.
UHD	Ultra High Definition.
VCD	Video Compact Disc.
VCEG	Video Coding Experts Group.
VLC	Variable Length Coding.
VND	Variable Node Decoder.
WAN	Wide Area Network.
WMV9	Windows Media Video version 9.
WZ	Wyner-Ziv.
XOR	eXclusive OR operation.

REFERENCES

[1] C. Hellge, D. Gomez-Barquero, T. Schierl, and T. Wiegand, "Layer-aware forward error correction for mobile broadcast of layered media," *IEEE Transactions on Multimedia*, vol. 13, pp. 551–562, June 2011.

[2] Y. Huo, M. El-Hajjar, and L. Hanzo, "Inter-layer FEC aided unequal error protection for multilayer video transmission in mobile TV," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, pp. 1622–1634, September 2013.

[3] Y. Huo, C. Zhu, and L. Hanzo, "Spatio-temporal iterative source-channel decoding aided video transmission," *IEEE Transactions on Vehicular Technology*, vol. 62, pp. 1597–1609, May 2013.

[4] Y. Huo, T. Wang, R. G. Maunder, and L. Hanzo, "Iterative source and channel decoding relying on correlation modelling for wireless video transmission," *IET Communications*, vol. 7, pp. 1465–1475, September 2013.

[5] 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: John Wiley, 2007.

[6] COST 211 and CCITT, *ITU-T Rec. H.120: Codecs for videoconferencing using primary digital group transmission*, November 1988.

[7] ITU-T, *H.261: Video codec for audiovisual services at p x 384 kbit/s*, November 1988.

[8] ITU-T, *H.263 : Video coding for low bit rate communication*, January 2005.

[9] 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*, March 2010.

[10] ITU-T, *H.262 : Information technology - Generic coding of moving pictures and associated audio information: Video*, Febrary 2012.

[11] G. K. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, pp. 30–44, 1991.

[12] ISO/IEC 11172, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s*, 1993.

[13] ITU-R, *BT.470 : Conventional analogue television systems*, Febrary 2005.

[14] IEC 61834-2, *Recording - Helical-scan digital video cassette recording system using 6,35 mm magnetic tape for consumer use*, August 1998.

[15] P. Cherriman and L. Hanzo, "Programmable H.263-based wireless video transceivers for interference-limited environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, pp. 275–286, June 1998.

[16] P. Cherriman, T. Keller, and L. Hanzo, "Orthogonal frequency division multiplex transmission of H.263 encoded video over highly frequency-selective wireless networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, pp. 701–712, August 1999.

[17] G. Conklin, G. Greenbaum, K. Lillevold, A. Lippman, and Y. Reznik, "Video coding for streaming media delivery on the Internet," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 269–281, March 2001.

[18] L. Bouchard, "Multimedia software for mobile phones," *IEEE Software*, vol. 27, pp. 8–10, May 2010.

[19] "SMPTE VC-1 Receiving Industry wide Support." Available: http://www.microsoft.com/en-us/news/press/2006/apr06/04-24VC1PR.aspx.

[20] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1103–1120, September 2007.

[21] Y. Chen, Y.-K. Wang, K. Ugur, M. Hannuksela, J. Lainema, and M. Gabbouj, "The emerging MVC standard for 3D video services," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1–13, January 2009.

[22] AVS Workgroup of China, *Audio and Video Coding Standard*, January 2005.

[23] ITU-T, *H.265 : High efficiency video coding*, April 2013.

[24] T. Zhang and Y. Xu, "Unequal packet loss protection for layered video transmission," *IEEE Transactions on Broadcasting*, vol. 45, pp. 243–252, June 1999.

[25] Y. Huo, M. El-Hajjar, and L. Hanzo, "Wireless video: An interlayer error-protection-aided multilayer approach," *IEEE Vehicular Technology Magazine*, vol. 9, pp. 104–112, September 2014.

[26] H. Imaizumi and A. Luthra, *Three-Dimensional Television, Video and Display Technologies*, ch. MPEG-2 Multiview Profile, pp. 169–181. Berlin, Heidelberg, and New York: Springer Verlag, 2002.

[27] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proceedings of the IEEE*, vol. 99, pp. 626–642, April 2011.

[28] O. Lehtoranta, J. Suhonen, M. Hännikäinen, V. Lappalainen, and T. D. Hämäläinen, "Comparison of video protection methods for wireless networks," *Signal Processing: Image Communication*, vol. 18, no. 10, pp. 861–877, 2003.

[29] Nasruminallah and L. Hanzo, "EXIT-chart optimized short block codes for iterative joint source and channel decoding in H.264 video telephony," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 4306–4315, October 2009.

[30] Nasruminallah and L. Hanzo, "Near-capacity H.264 multimedia communications using iterative joint source-channel decoding," *IEEE Communications Surveys and Tutorials*, vol. 14, pp. 538–564, Second Quarter 2012.

[31] S. Wenger, "H.264/AVC over IP," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 645–656, July 2003.

[32] M. Wien, H. Schwarz, and T. Oelbaum, "Performance analysis of SVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1194–1203, September 2007.

[33] A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243–250, June 1996.

[34] B.-J. Kim, Z. Xiong, and W. Pearlman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 1374–1387, December 2000.

[35] E. N. Leith and J. Upatnieks, "Scalable video coding and transport over broadband wireless networks," *Proceedings of the IEEE*, vol. 89, pp. 6–20, January 2001.

[36] T. Schierl, T. Stockhammer, and T. Wiegand, "Mobile video transmission using scalable video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1204–1217, September 2007.

[37] M. Wien, R. Cazoulat, A. Graffunder, A. Hutter, and P. Amon, "Real-time system for adaptive video streaming based on SVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1227–1237, September 2007.

[38] M. Hassan, A. Nayandoro, and M. Atiquzzaman, "Internet telephony: services, technical challenges, and products," *IEEE Communications Magazine*, vol. 38, pp. 96–103, April 2000.

[39] D. Wu, Y. T. Hou, and Y.-Q. Zhang, "Transporting real-time video over the Internet: challenges and approaches," *Proceedings of the IEEE*, vol. 88, no. 12, pp. 1855–1877, 2000.

[40] B. Sklar, "Rayleigh fading channels in mobile digital communication systems .I. characterization," *IEEE Communications Magazine*, vol. 35, pp. 90–100, July 1997.

[41] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Channel-adaptive resource allocation for scalable video transmission over 3G wireless network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 1049–1063, August 2004.

[42] D. Song and C. W. Chen, "Scalable H.264/AVC video transmission over MIMO wireless systems with adaptive channel selection based on partial channel information," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1218–1226, September 2007.

[43] D. T. Nguyen and J. Ostermann, "Congestion control for scalable video streaming using the scalability extension of H.264/AVC," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, pp. 246–253, August 2007.

[44] J. Monteiro, C. Calafate, and M. S. Nunes, "Evaluation of the H.264 scalable video coding in error prone IP networks," *IEEE Transactions on Broadcasting*, vol. 54, pp. 652–659, September 2008.

[45] Y. Fallah, H. Mansour, S. Khan, P. Nasiopoulos, and H. Alnuweiri, "A link adaptation scheme for efficient transmission of H.264 scalable video over multirate WLANs," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, pp. 875–887, July 2008.

[46] Y. Wang, S. Wenger, J. Wen, and A. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol. 17, pp. 61–82, July 2000.

[47] A. Sehgal, A. Jagmohan, and N. Ahuja, "Wyner-Ziv coding of video: an error-resilient compression framework," *IEEE Transactions on Multimedia*, vol. 6, pp. 249–258, April 2004.

[48] Y. Zhang, C. Zhu, and K.-H. Yap, "A joint source-channel video coding scheme based on distributed source coding," *IEEE Transactions on Multimedia*, vol. 10, pp. 1648–1656, December 2008.

[49] Y. Guo, Y. Chen, Y.-K. Wang, H. Li, M. Hannuksela, and M. Gabbouj, "Error resilient coding and error concealment in scalable video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, pp. 781–795, June 2009.

[50] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Transactions on Information Theory*, vol. 13, pp. 600–607, October 1967.

[51] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Transactions on Information Theory*, vol. 42, pp. 1737–1744, November 1996.

[52] T. Brüggen and P. Vary, "Unequal error protection by modulation with unequal power allocation," *IEEE Communications Letters*, vol. 9, pp. 484–486, June 2005.

[53] V. Pavlushkov, R. Johannesson, and V. Zyablov, "Unequal error protection for convolutional codes," *IEEE Transactions on Information Theory*, vol. 52, pp. 700–708, February 2006.

[54] N. Rahnavard and F. Fekri, "New results on unequal error protection using LDPC codes," *IEEE Communications Letters*, vol. 10, pp. 43–45, January 2006.

[55] V. Kumar and O. Milenkovic, "On unequal error protection LDPC codes based on Plotkin-type constructions," *IEEE Transactions on Communications*, vol. 54, pp. 994–1005, June 2006.

[56] C. Gong, G. Yue, and X. Wang, "Message-wise unequal error protection based on low-density parity-check codes," *IEEE Transactions on Communications*, vol. 59, pp. 1019–1030, April 2011.

[57] N. Rahnavard, H. Pishro-Nik, and F. Fekri, "Unequal error protection using partially regular LDPC codes," *IEEE Transactions on Communications*, vol. 55, pp. 387–391, March 2007.

[58] 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 Journal on Selected Areas in Communications*, vol. 22, pp. 777–790, May 2004.

[59] M. Gallant and F. Kossentini, "Rate-distortion optimized layered coding with unequal error protection for robust internet video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 357–372, March 2001.

[60] L. Thornton, J. Soraghan, R. Kutil, and M. Chakraborty, "Unequally protected SPIHT video codec for low bit rate transmission over highly error-prone mobile channels," *Signal Processing: Image Communication*, vol. 17, no. 4, pp. 327–335, 2002.

[61] M. van der Schaar, S. Krishnamachari, S. Choi, and X. Xu, "Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 1752–1763, December 2003.

[62] Y.-C. Su, C.-S. Yang, and C.-W. Lee, "Optimal FEC assignment for scalable video transmission over burst error channel with loss rate feedback," *Signal Processing: Image Communication*, vol. 18, no. 7, pp. 537 – 547, 2003.

[63] J. Kim, R. Mersereau, and Y. Altunbasak, "Error-resilient image and video transmission over the internet using unequal error protection," *IEEE Transactions on Image Processing*, vol. 12, pp. 121–131, February 2003.

[64] J. Kim, R. Mersereau, and Y. Altunbasak, "A multiple-substream unequal error-protection and error-concealment algorithm for SPIHT-coded video bitstreams," *IEEE Transactions on Image Processing*, vol. 13, pp. 1547–1553, December 2004.

[65] F. Marx and J. Farah, "A novel approach to achieve unequal error protection for video transmission over 3G wireless networks," *Signal Processing: Image Communication*, vol. 19, no. 4, pp. 313–323, 2004.

[66] 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 Proceedings Communications*, vol. 152, pp. 1116–1124, December 2005.

[67] H. Wang, F. Zhai, Y. Eisenberg, and A. Katsaggelos, "Cost-distortion optimized unequal error protection for object-based video communications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, pp. 1505–1516, December 2005.

[68] Y. C. Chang, S.-W. Lee, and R. Komiya, "A low-complexity unequal error protection of H.264/AVC video using adaptive hierarchical QAM," *IEEE Transactions on Consumer Electronics*, vol. 52, pp. 1153–1158, November 2006.

[69] T. Fang and L.-P. Chau, "GOP-based channel rate allocation using genetic algorithm for scalable video streaming over error-prone networks," *IEEE Transactions on Image Processing*, vol. 15, pp. 1323–1330, June 2006.

[70] B. Barmada, M. Ghandi, E. Jones, and M. Ghanbari, "Combined turbo coding and hierarchical QAM for unequal error protection of h.264

coded video," *Signal Processing: Image Communication*, vol. 21, no. 5, pp. 390–395, 2006.

[71] T. Gan, L. Gan, and K.-K. Ma, "Reducing video-quality fluctuations for streaming scalable video using unequal error protection, retransmission, and interleaving," *IEEE Transactions on Image Processing*, vol. 15, pp. 819–832, April 2006.

[72] M. Ghandi and M. Ghanbari, "Layered H.264 video transmission with hierarchical QAM," *Journal of Visual Communication and Image Representation*, vol. 17, no. 2, pp. 451–466, 2006.

[73] Q. Qu, Y. Pei, and J. Modestino, "An adaptive motion-based unequal error protection approach for real-time video transport over wireless IP networks," *IEEE Transactions on Multimedia*, vol. 8, pp. 1033–1044, October 2006.

[74] S. Yingbo Shi, C. Wu, and J. Du, "A novel unequal loss protection approach for scalable video streaming over wireless networks," *IEEE Transactions on Consumer Electronics*, vol. 53, pp. 363–368, May 2007.

[75] J. Huusko, J. Vehkaperä, P. Amon, C. Lamy-Bergot, G. Panza, J. Peltola, and M. Martini, "Cross-layer architecture for scalable video transmission in wireless network," *Signal Processing: Image Communication*, vol. 22, no. 3, pp. 317–330, 2007.

[76] H. Ha, C. Yim, and Y. Y. Kim, "Packet loss resilience using unequal forward error correction assignment for video transmission over communication networks," *Computer Communications*, vol. 30, no. 18, pp. 3676–3689, 2007.

[77] M. Aydinlik and M. Salehi, "Turbo coded modulation for unequal error protection," *IEEE Transactions on Communications*, vol. 56, pp. 555–564, April 2008.

[78] H. Ha and C. Yim, "Layer-weighted unequal error protection for scalable video coding extension of H.264/AVC," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 736–744, May 2008.

[79] 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 Transactions on Consumer Electronics*, vol. 54, pp. 1066–1073, August 2008.

[80] 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 Transactions on Consumer Electronics*, vol. 55, pp. 1089–1097, August 2009.

[81] Y. Shan, S. Yi, S. Kalyanaraman, and J. W. Woods, "Adaptive two-stage FEC scheme for scalable video transmission over wireless networks," *Signal Processing: Image Communication*, vol. 24, no. 9, pp. 718–729, 2009.

[82] D. Sejdinović, D. Vukobratović, A. Doufexi, V. Šenk, and R. Piechocki, "Expanding window fountain codes for unequal error protection," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2510–2516, 2009.

[83] M. Jubran, M. Bansal, L. Kondi, and R. Grover, "Accurate distortion estimation and optimal bandwidth allocation for scalable H.264 video transmission over MIMO systems," *IEEE Transactions on Image Processing*, vol. 18, pp. 106–116, January 2009.

[84] Y. Shan, I. Bajic, J. Woods, and S. Kalyanaraman, "Scalable video streaming with fine-grain adaptive forward error correction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, pp. 1302–1314, September 2009.

[85] D. Vukobratović, V. Stanković, D. Sejdinović, L. Stanković, and Z. Xiong, "Scalable video multicast using expanding window fountain codes," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1094–1104, 2009.

[86] P. Li, Y. Chang, N. Feng, and F. Yang, "A novel hierarchical QAM-based unequal error protection scheme for H.264/AVC video over frequency-selective fading channels," *IEEE Transactions on Consumer Electronics*, vol. 56, pp. 2741–2746, November 2010.

[87] E. Maani and A. Katsaggelos, "Unequal error protection for robust streaming of scalable video over packet lossy networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, pp. 407–416, March 2010.

[88] S. Ahmad, R. Hamzaoui, and M. Al-Akaidi, "Unequal error protection using fountain codes with applications to video communication," *IEEE Transactions on Multimedia*, vol. 13, pp. 92–101, February 2011.

[89] K. Nguyen, T. Nguyen, and S.-C. Cheung, "Video streaming with network coding," *Journal of Signal Processing Systems*, vol. 59, pp. 319–333, June 2010.

[90] H. Xiao, Q. Dai, X. Ji, and W. Zhu, "A novel JSCC framework with diversity-multiplexing-coding gain tradeoff for scalable video transmission over cooperative MIMO," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, pp. 994–1006, July 2010.

[91] H. Zhang, Y. Zheng, M. Khojastepour, and S. Rangarajan, "Cross-layer optimization for streaming scalable video over fading wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, pp. 344–353, April 2010.

[92] S. Dumitrescu, G. Rivers, and S. Shirani, "Unequal erasure protection technique for scalable multistreams," *IEEE Transactions on Image Processing*, vol. 19, pp. 422–434, February 2010.

[93] M. Halloush and H. Radha, "Network coding with multi-generation mixing: A generalized framework for practical network coding," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 466–473, February 2011.

[94] A. A. Khalek, C. Caramanis, and R. W. Heath, "A cross-layer design for perceptual optimization of H.264/SVC with unequal error protection," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 1157–1171, August 2012.

[95] H. Kim, P. Cosman, and L. Milstein, "Motion-compensated scalable video transmission over MIMO wireless channels," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, pp. 116–127, January 2013.

[96] M. Li, Z. Chen, and Y.-P. Tan, "Scalable resource allocation for svc video streaming over multiuser MIMO-OFDM networks," *IEEE Transactions on Multimedia*, vol. 15, pp. 1519–1531, November 2013.

[97] S.-C. Wang and W. Liao, "Cooperative multicasting for wireless scalable video transmissions," *IEEE Transactions on Communications*, vol. 61, pp. 3980–3989, September 2013.

[98] R. Xiong, D. Taubman, and V. Sivaraman, "PET protection optimization for streaming scalable videos with multiple transmissions," *IEEE Transactions on Image Processing*, vol. 22, pp. 4364–4379, November 2013.

[99] Y. Zhang, S. Qin, B. Li, and Z. He, "Rate-distortion optimized unequal loss protection for video transmission over packet erasure channels," *Signal Processing: Image Communication*, vol. 28, pp. 1390–1404, November 2013.

[100] J. Micallef, R. Farrugia, and C. Debono, "Correlation noise based unequal error protected rate-adaptive codes for distributed video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2013.

[101] A. Midya, R. Ranjan, and S. Sengupta, "Scene content driven FEC allocation for video streaming," *Signal Processing: Image Communication*, vol. 29, no. 1, pp. 37–48, 2014.

[102] Y. Huo, M. El-Hajjar, R. G. Maunder, and L. Hanzo, "Layered wireless video relying on minimum-distortion inter-layer FEC coding," *IEEE Transactions on Multimedia*, vol. 23, pp. 319–331, January 2014.

[103] S. Cicalo and V. Tralli, "Distortion-fair cross-layer resource allocation for scalable video transmission in OFDMA wireless networks," *IEEE Transactions on Multimedia*, vol. PP, no. 99, pp. 1–1, 2014.

[104] T. Stockhammer, M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 657–673, 2003.

[105] K. Alajel, W. Xiang, and Y. Wang, "Unequal error protection scheme based hierarchical 16-QAM for 3-D video transmission," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 731–738, 2012.

[106] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *Conference Record of the 33rd Asilomar Conference on Signals, Systems, and Computers*, vol. 1, (Pacific Grove, CA), pp. 342–346, October 1999.

[107] W. Xiang, C. Zhu, C.-K. Siew, Y. Xu, and M. Liu, "Forward error correction-based 2-D layered multiple description coding for error-resilient H.264 SVC video transmission," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, pp. 1730–1738, December 2009.

[108] M. Luby, "LT codes," (Vancouver, BC, Canada), pp. 271–280, November 2002.

[109] C.-L. Huang and S. Liang, "Unequal error protection for MPEG-2 video transmission over wireless channels," *Signal Processing: Image Communication*, vol. 19, no. 1, pp. 67–79, 2004.

[110] W.-J. Hwang, C.-M. Ou, R.-C. Lin, and W.-W. Hu, "Layered video transmission based on genetic programming for lossy channels," *Neurocomputing*, vol. 57, no. 0, pp. 361–372, 2004.

[111] E. Namjoo, A. Aghagolzadeh, and J. Museviniya, "Robust transmission of scalable video stream using modified LT codes," *Computers & Electrical Engineering*, vol. 37, no. 5, pp. 768–781, 2011.

[112] C. Hellge, T. Schierl, and T. Wiegand, "Multidimensional layered forward error correction using rateless codes," in *IEEE International Conference on Communications*, (Beijing, China), pp. 480–484, May 2008.

[113] J. Xiao, T. Tillo, and Y. Zhao, "Real-time video streaming using randomized expanding Reed Solomon code," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, pp. 1825–1836, Nov 2013.

[114] C. Diaz, C. Hellge, J. Cabrera, F. Jaureguizar, and T. Schierl, "Enhancement of Pro-MPEG COP3 codes and application to layer-aware FEC protection of two-layered video transmission," in *IEEE International Conference on Image Processing (ICIP)*, pp. 1598–1602, Sept 2013.

[115] Y. Huo, X. Zuo, R. G. Maunder, and L. Hanzo, "Inter-layer FEC decoded multi-layer video streaming," in *IEEE Global Telecommunications Conference (GLOBECOM)*, (Anaheim, CA, USA), pp. 2113–2118, December 2012.

[116] Y. Huo, M. El-Hajjar, M. F. U. Butt, and L. Hanzo, "Inter-layer-decoding aided self-concatenated coded scalable video transmission," in *IEEE Wireless Communications and Networking Conference (WCNC)*, (Shanghai, China), pp. 4647–4652, April 2013.

[117] 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.), pp. 239–280, Boca Raton, FL: CRC Press, 2008.

[118] L. Hanzo, T. Liew, B. Yeap, and R. Tee, *Turbo Coding, Turbo Equalisation and Space-Time Coding*. New York: John Wiley, 2011.

[119] S. X. Ng and J. Y. Chung and P. Cherriman and L. Hanzo, "Burst-by-Burst Adaptive Decision Feedback Equalised TCM, TTCM and BICM for H.263-Assisted Wireless Video Telephony," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, pp. 363–374, March 2006.

[120] Y. Huo, T. Wang, R. G. Maunder, and L. Hanzo, "Motion-aware mesh-structured trellis for correlation modelling aided distributed multi-view video coding," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 319–331, 2014.

[121] Y. Huo, T. Wang, R. Maunder, and L. Hanzo, "Two-dimensional iterative source-channel decoding for distributed video coding," *IEEE Communications Letters*, vol. 18, pp. 90–93, January 2014.

[122] I. J. 11, "Dis 23008-10: Information technology - high efficiency coding and media delivery in heterogeneous environments - part 10: Mpeg media transport forward error correction (fec) codes," January 2014.

[123] A. Shokrollahi, "Raptor codes," *IEEE Transactions on Information Theory*, vol. 52, pp. 2551–2567, June 2006.

[124] J. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1528–1540, October 2002.

[125] N. Rahnavard, B. Vellambi, and F. Fekri, "Rateless codes with unequal error protection property," *IEEE Transactions on Information Theory*, vol. 53, pp. 1521–1532, April 2007.

[126] R. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA, USA: MIT Press, 1963.

[127] Shokrollahi and M. Luby., "Raptor codes," *Foundations and Trends in Communications and Information Theory*, vol. 6, no. 3-4, pp. 213–322, 2011.

[128] M. Luby, A. Shokrollahi, M. Watson, and T. Stockhammer, "Raptor forward error correction scheme for object delivery," *RFC 5053*, October 2007.

[129] M. Luby, A. Shokrollahi, M. Watson, T. Stockhammer, and L. Minder, "RaptorQ forward error correction scheme for object delivery," *RFC 6330*, August 2011.

[130] J. Xiao, T. Tillo, C. Lin, and Y. Zhao, "Dynamic sub-GOP forward error correction code for real-time video applications," *IEEE Transactions on Multimedia*, vol. 14, pp. 1298–1308, August 2012.

[131] Pro-MPEG Forum Code of Practice 3 release 2, *Transmissions of professional MPEG-2 transport streams over IP networks*, July 2004.

[132] R. Tanner, "A recursive approach to low complexity codes," *IEEE Transactions on Information Theory*, vol. 27, pp. 533–547, September 1981.

[133] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, pp. 498–519, Febrary 2001.

[134] C. Hellge, *Prioritized Transmission of Layered Media with Layer-Aware FEC*. Mensch & Buch, 2013.

[135] DVB, "Blue book a115 dvb application layer fec evaluations," 2006.

[136] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Transactions on Communications*, vol. 52, pp. 670–678, April 2004.

[137] R. Maunder and L. Hanzo, "Extrinsic information transfer analysis and design of block-based intermediate codes," *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 762–770, March 2011.

[138] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, pp. 429–445, March 1996.

[139] J. Chen, A. Dholakia, E. Eleftheriou, M. Fossorier, and X.-Y. Hu, "Reduced-complexity decoding of LDPC codes," *IEEE Transactions on Communications*, vol. 53, pp. 1288–1299, August 2005.

[140] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation*. John Wiley & Sons, IEEE press, 2009.

[141] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *IEEE International Conference on Communications*, (Geneva, Switzerland), pp. 1064–1070, May 1993.

[142] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, October 2001.

[143] R. Khalili and K. Salamatian, "A new analytic approach to evaluation of packet error rate in wireless networks," in *Communication Networks and Services Research Conference*, (Halifax, Nova Scotia, Canada), pp. 333–338, May 2005.

[144] 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 Processing: Image Communication*, vol. 18, no. 3, pp. 157–167, 2003.

[145] Y. Liu and S. Yu, "Adaptive unequal loss protection for scalable video streaming over IP networks," *IEEE Transactions on Consumer Electronics*, vol. 51, pp. 1277–1282, November 2005.

[146] C. Zhang, H. Yang, S. Yu, and X. Yang, "Gop-level transmission distortion modeling for mobile streaming video," *Signal Processing: Image Communication*, vol. 23, no. 2, pp. 116–126, 2008.

[147] M. Clerc and J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.

[148] D. Schonfeld and N. Bouaynaya, "A new method for multidimensional optimization and its application in image and video processing," *IEEE Signal Processing Letters*, vol. 13, no. 8, pp. 485–488, 2006.

[149] Z.-H. Zhan, J. Zhang, Y. Li, and H.-H. Chung, "Adaptive particle swarm optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, pp. 1362–1381, December 2009.

[150] P. Cherriman, C. Wong, and L. Hanzo, "Turbo- and BCH-coded wideband burst-by-burst adaptive H.263-assisted wireless video telephony," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 1355–1363, December 2000.

[151] Nasruminallah and L. Hanzo, "Short block codes for guaranteed convergence in soft-bit assisted iterative joint source and channel decoding," *Electronics Letters*, vol. 44, pp. 1315–1316, October 2008.

[152] R. Otnes and M. Tüchler, "EXIT chart analysis applied to adaptive turbo equalization," in *Nordic Signal Processing Symposium*, (Hurtigruten, Norway), pp. 1–6, October 2002.

[153] M. El-Hajjar and L. Hanzo, "Layered steered space-time codes and their capacity," *Electronics Letters*, vol. 43, pp. 680–682, June 2007.

[154] 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 *IEEE Symposium on Computers and Communications, ISCC*, (Sousse, Tunisia), pp. 36–41, July 2009.