Inter-Layer-Decoding Aided Self-Concatenated Coded Scalable Video Transmission

Yongkai Huo*, Mohammed EI-Hajjar*, M. Fasih Uddin Butt† and Lajos Hanzo*, *Fellow, IEEE*
†Department of EE, COMSATS Institute of Information Technology, Islamabad, Pakistan.
*School of ECS, University of Southampton, UK.
Email: {yh3g09, meh, lh}@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract—In this treatise, we investigate an inter-layer (IL) decoded self-concatenated coded system conceived for scalable video coding (SVC). In SVC, when the base layer (BL) is corrupted due to channel-induced decoding errors, the enhancement layers (ELs) are discarded by the SVC decoder. The proposed IL coding technique implants the systematic information of the BL into the systematic bits of the ELs using exclusive-OR (XOR) operations. The IL coding is combined with a self-concatenated code, which is a near-capacity iterative detection aided channel codec using a single encoder and a single decoder. At the receiver, the IL decoding technique is activated for correcting the errors residing in the error-infested BL. In our experiments, the proposed system is combined with a number of the traditional unequal error protection (UEP) arrangements for benchmarking the system performance. Our simulation results demonstrate that the proposed technique is capable of further improving the system performance of the traditional UEP assisted system by an E_b/N_0 of about 3 dB at a peak signal-to-noise ratio (PSNR) of 36 dB. When viewed from a different perspective, the attainable gain can be quantified as a PSNR improvement of 2.9 dB at $E_b/N_0 = -5$ dB, while imposing a marginal complexity increase of 1.5%.

I. INTRODUCTION

Scalable video coding (SVC) [1] encodes a video sequence into multiple inter-dependent layers, which enables progressive refinement of the reconstructed video. In SVC the most important layer and the less significant layers are referred to as the base layer (BL) and enhancement layers (ELs), respectively. When the BL is corrupted during the transmission, the ELs must be dropped by the receiver. Moreover, specific users can extract a subset of the SVC-compressed bitstreams, depending on their equipments' requirements. For example, a smart phone with a small screen may decode the BL only, while a HDTV requires decoding many more SVC layers. The SVC option is part of the recent H.264/AVC video coding standard [2], which may be used for mobile video broadcasting.

Unequal error protection (UEP) was originally proposed by Masnick and Wolf in [3], which protects the more significant bits by stronger codes and the less significant ones by weaker codes. Since SVC partitions the video sequence into multiple layers of unequal importance, numerous contributions have been focused on the UEP of SVC transmission. A UEP scheme was developed in [4] aiming for SVC video streaming over transmission media imposing packet-loss events. A cross-layer operation aided scheme was proposed in [5] for robust SVC streaming over error-prone networks, where both the available bandwidth and the estimated packet loss ratio (PLR) were used for beneficially configuring the system.

In traditional UEP schemes designed for SVC streaming, different-strength UEP codes are allocated to the different video layers according to their significance. However, the ELs must be dropped by the video decoder, regardless whether they are perfectly recovered or not, when the BL is corrupted. In other words, both the bandwidth and the transmit power allocated to the ELs are wasted, when the BL is lost or corrupted. To combat this problem, a so-called layer-aware forward error correction (LA-FEC) philosophy was designed in [6] for layered video transmission over the binary erasure channel (BEC), where a rateless Raptor codec was employed for recovering the bits erased in the network. In [7], we proposed an inter-layer FEC (IL-FEC) aided UEP scheme employing a recursive systematic convolutional code (RSC) and a turbo codec [8], where the H.264/AVC video encoder configured in its slice-partitioning mode [2] was employed. Specifically, the systematic bits of the BL, namely of the type A partition, were implanted into the systematic bits of the ELs, namely into the type B and type C partitions at the transmitter. A sophisticated IL decoding technique was activated at the receiver for correcting the errors imposed upon the type A partition.

On the other hand, a novel self-concatenated codec [9] was proposed as a powerful near-capacity channel codec, which exhibits a relatively low complexity and relies on a single encoder and a single decoder. More specifically, iteratively-decoded self-concatenated convolutional codes (SECCC-ID) were proposed in [9], which were then further developed in [10]. As another radical performance-improvement, multiple input multiple output (MIMO) systems have been proposed for achieving either a multiplexing or a diversity gain. Layered steered space-time codes (LSSTC) [11] have been developed for combining the benefits of several MIMO techniques, including those of the vertical Bell Labs space-time (VBLAST) structure, of space-time block codes and of beamforming [12].

In this treatise, we conceive a novel system for IL coded SECCC aided scalable video streaming using an amalgamated LSSTC transceiver structure (IL-SECCC-LSSTC) employing a RSC as its constituent codes. Our scheme may be combined with the traditional UEP philosophy as a further enhancement, where the rate allocation optimization problem has been solved as in [4], [5]. Furthermore, the SVC mode of the H.264/AVC video codec is employed for encoding a video sequence into multiple layers of unequal significance.

Against this background, our novel contributions are:

1) *We conceive an inter-layer coded SECCC codec for scalable video streaming, which is combined with the traditional UEP philosophy for the sake of achieving an improved system performance. Furthermore, this codec is incorporated into a video communication system em-*
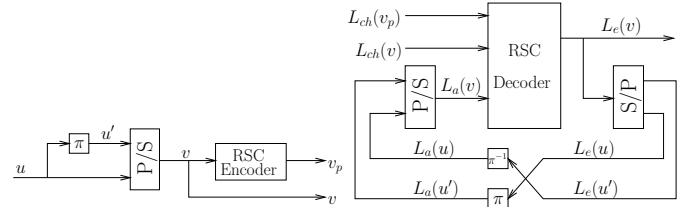
Fig. 2: The encoding (left) and decoding (right) structure of the SECCC codec, where "P/S" and "S/P" represent a parallel to serial and serial to parallel converter, respectively.

ploying a multi-functional MIMO transceiver structure.

2) *A gain of 3 dB is attained at a PSNR of 36 dB, which may also be interpreted as a PSNR gain of 2.9 dB at $E_b/N_0 = -5$ dB against the traditional UEP benchmarkers.*

3) *A modest complexity increase of 1.5% is imposed by our inter-layer decoding scheme.*

In Section II, we briefly introduce the SVC technique of the H.264/AVC followed by Section III, where we describe our proposed IL-SECCC-LSSTC system architecture. In Section IV, the performance of IL-SECCC-LSSTC is benchmarked against the typical SECCC-LSSTC system, when employing SECCC error protection arrangements. Finally, we provide our conclusions in Section V.

II. H.264/AVC SCALABLE VIDEO CODING

The SVC [1] mode was standardized in the Annex G extension of the H.264/AVC video codec [2], which encodes a video sequence into multiple dependent video layers. The most significant layer is referred to as the BL, while the other layers which depend on the BL are referred to as ELs, noting that an EL may be further relied upon by multiple ELs. When either the BL or an EL is corrupted or lost during transmission, the ELs depending on it must be discarded by the SVC decoder. On the other hand, each EL further improves the reconstructed video quality. These characteristics facilitate flexible bitrate and video quality control according to the specific users' preferences. More specifically, some ELs may be removed from the SVC-compressed bitstream for the sake of meeting a specific video bitrate or quality requirement.

A H.264 SVC stream contains a sequence of network abstraction layer units (NALUs) [2], 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 bits represeting a video frame. The parameters dependency_id (DID), temporal_id (TID) and quality_id (QID) contained in the NALU header describe the scalability features of the bitstream. Specifically, DID, TID and QID represent Coarse Grain Scalability (CGS), Temporal Scalability and Medium Grain Scalability (MGS) [1], [13], respectively.

The CGS allows us to adaptively specify the spatial resolution of the video, ranging from quarter common intermediate format (QCIF) to common intermediate format (CIF) for example, where the video can be encoded into a set of sub-streams referred to as dependency-layers. The DID parameter represents the specific dependency-layer the current NALU belongs to. The decoding of a NALU associated with $DID > 0$ depends on all the NALUs belonging to $(DID - 1)$, which are associated with the same TID and QID values. Based on this dependency rule, the video quality may be conveniently improved or reduced by incorporating or removing the specific NALUs having a DID larger than a specific DID parameter. Similar dependency rules are also applied for both temporal scalability and MGS.

III. SYSTEM MODEL

In this section, we detail the architecture of the IL-SECCC-LSSTC conceived for SVC transmission over our LSSTC multi-functional MIMO system. A block diagram of the proposed system is illustrated in Fig. 1a. Firstly, the SECCC codec

will be detailed in Section III-A. Then, the structure of the transmitter employing the IL-SECCC encoder and the LSSTC MIMO system will be detailed in Section III-B. Afterwards, the structure of receiver employing the IL SECCC decoder and the LSSTC MIMO receiver will be illustrated in Section III-C. Moreover, for simplifying our related discussions, we assume that two layers are generated by the SVC encoder, which we refer to as layer L_0 and layer L_1. For the sake of simplifying our elaborations, these two layers are assumed to contain an equal number of n bits, but our algorithm may be readily extended to the scenario of having an unequal length as in [7]. The system's architecture is displayed in Fig. 1a, while the structures of the variable node decoder (VND) and of the check node decoder (CND) [14] are detailed in Fig. 1b. Assuming that u_1, u_2 and $u_3 = u_1 \oplus u_2$ represent random binary variables, the action of the VND of Fig. 1b may be formulated as $L_{o_3}(u_1) = L_{i_1}(u_1) + L_{i_2}(u_1)$, while the boxplus operation $L(u_1 \oplus u_2) = L(u_1) \boxplus L(u_2)$ [15] is expressed as follows

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

The CND of Fig. 1b may be formulated as $L_o(u_3) = L_i(u_1) \boxplus L_i(u_2)$.

A. SECCC Codec

The encoder architecture of the SECCC codec [10] is displayed in Fig. 2, where a RSC codec is employed. The original bits u are firstly interleaved by the interleaver π of Fig. 2 for generating the bit-sequence u'. Then, the original bits u and the interleaved bits u' are converted to the bit sequence v by the "P/S" converter, which implies that the resultant bit sequence v contains twice the bits contained in the bit sequence u. Afterwards, the bit sequence v is encoded by a classic RSC codec for generating the parity bits v_p, as seen in Fig. 2. The parity bits v_p and the systematic bits v are the final output of the SECCC encoder of Fig. 2.

The decoder architecture of the SECCC codec [10] is also shown in Fig. 2, where the LLR information $L_{ch}(v)$ and $L_{ch}(v_p)$ related to the systematic bits v and the parity bits v_p, respectively, are input to a classic RSC decoder employed as the core of the SECCC codec [10]. After each round of RSC decoding, the extrinsic information $L_e(v)$ is generated, which will be split into two parts by the "S/P" converter of Fig. 2, namely the extrinsic information $L_e(u)$ related to the original bits u and the extrinsic information $L_e(u')$ of the interleaved bits u'. Then these two types of information will be appropriately interleaved/deinterleaved in order to generate the
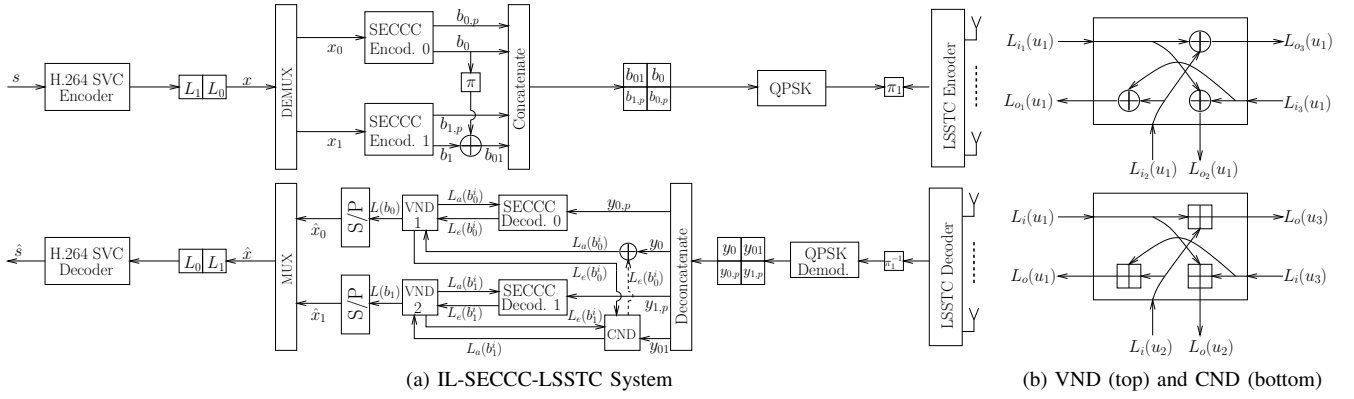
(a) IL-SECCC-LSSTC System

(b) VND (top) and CND (bottom)

Fig. 1: IL-SECCC-LSSTC encoding/decoding architecture of SVC H.264/AVC coded video, where $\oplus$ and $\boxplus$ indicate the addition and boxplus operation, respectively, while $L_i(\cdot)$ and $L_o(\cdot)$ indicate the input and output LLRs, respectively.

a-priori information $L_a(u')$ for the interleaved bit sequence u' and the *a-priori* information $L_a(u)$ for the original bit sequence u, which is converted by the "P/S" converter into the *a-priori* information $L_a(v)$ for the bit sequence v. Afterwards, the next decoding iteration will be performed based on the new *a-priori* information. We emphasize that the extrinsic information of the original bits u and the interleaved bits u' is exchanged in each decoding iteration, which enables the SECCC to carry out a turbo-like decoding process and hence to attain a turbo-like error correction cability, while using the single RSC decoder shown in Fig. 2.

B. Transmitter Model

At the transmitter, the video signal s is encoded by the SVC of the H.264/AVC codec as shown in Fig. 1a, generating the BL L_0 and the EL L_1. Then, the layers L_0 and L_1 are de-multiplexed into two bitstreams by the *DEMUX* component of Fig. 1a, namely into the stream x_0 representing the layer L_0 and the stream x_1 carrying the data of layer L_1.

Then, the bits x_0 representing layer L_0 will be encoded by the SECCC encoder 0 of Fig. 1a, which generates the coded bits, including the systematic bits b_0 and the parity bits $b_{0,p}$. The bits representing the EL x_1 will firstly be encoded into the systematic bits b_1 and the parity bits $b_{1,p}$ by the SECCC encoder 1. Then, the exclusive-OR (XOR) operation will implant the systematic information bits b_0 into the systematic bit sequence b_1 without changing the parity bits of layer L_1, namely $b_{1,p}$. More specifically, the XOR-based implantation process generates the check bits $b_{01}^i = b_0^i \oplus b_1^i$. Overall, the check bits b_{01} and the parity bits $b_{1,p}$ are the final coded bits.

Finally, the bit sequences b_0, $b_{0,p}$, b_{01} and $b_{0,p}$ will be concatenated into a joint bitstream that is mapped to the quadrature phase shift keying (QPSK) modulator of Fig. 1a. Note that a bit-level interleaver is employed for interleaving the BL bit sequence b_0 before the XOR-based implantation of the BL bits into the EL bits b_1. Additionally, two puncturers are employed for generating specific code rates of both the BL L_0 and of the EL L_1, respectively. However, the puncturers are omitted in Fig. 1a for simplifying the system's architecture.

Following the IL-SECCC encoding process, the coded bits will be QPSK modulated and then transmitted by the LSSTC based MIMO transmitter [11]. The transmission structure displayed in Fig. 1a is equipped with $N_t = 4$ transmit antennas, which are spaced sufficiently for apart for encountering independent fading. Correspondingly, the receiver also has $N_r = 4$ receive antennas. The LSSTC system employed is characterized by a diversity order of 2 and multiplexing order
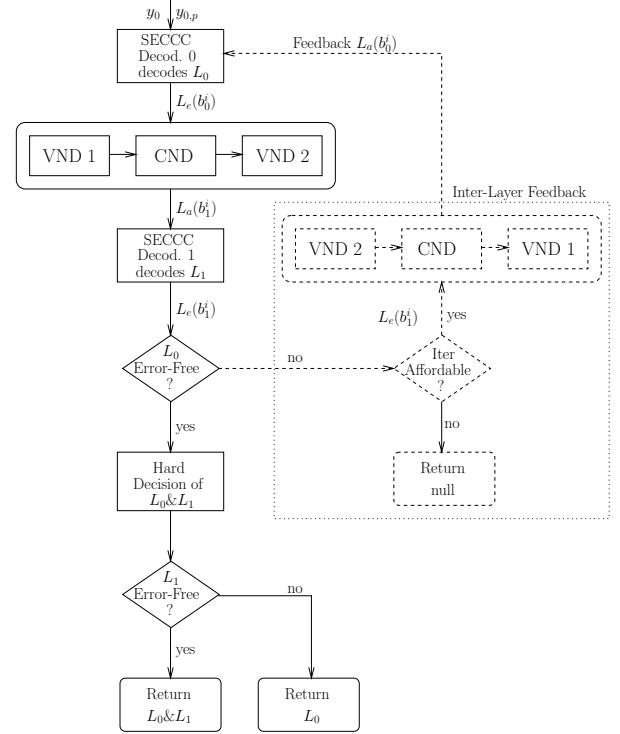

Fig. 3: The flow chart of inter-layer aided SECCC decoding.

of 2, thereby providing twice the data rate of a single antenna system, while also attaining a second-order diversity gain.

C. Receiver Model

At the receiver[1], the LSSTC decoding process is performed as detailed in [11], where the LSSTC output is passed to the QPSK demapper that outputs the log-likelihood ratios (LLR) of both the systematic information y_0, y_{01} and of the parity information $y_{0,p}$, $y_{1,p}$, for the BL L_0 and the EL L_1, respectively. Following the demodulation process, IL SECCC decoding will be performed for both layers.

The IL aided SECCC decoding process is illustrated by the flow-chart of Fig. 3. Firstly, the SECCC decoder 0 will decode the received information y_0 and $y_{0,p}$ for estimating the LLRs of the bits b_0 of the layer L_0. Then, the resultant extrinsic LLR information of layer L_0 will be input to the "VND1-CND-VND2" block of Fig. 3 for extracting the *a-priori* LLRs $L_a(b_1^i)$ of layer L_1, which is carried out by following the processing of the LLRs in the VND 1, CND and VND 2 components of Fig. 1a. Specifically, the "VND1-CND-VND2" block of Fig. 3 performs the following operations step-by-step:

---

[1]The deinterleaver π^{-1} is ignored at the receiver for the sake of simplifying the system architecture.

1) The extrinsic LLR $L_e\left(b_0^i\right)$ generated by the SECCC decoder 0 is input to the VND 1 block of Fig. 1a, which extracts the extrinsic LLR information $L_e\left(b_0^i\right)$ and forwards it to the CND block of Fig. 1a.

2) The LLR information $L_e\left(b_0^i\right)$ and the received check information y_{01} is input to the CND block of Fig. 1a for extracting the LLR information of the systematic bit b_1^i, namely the soft input $L_a\left(b_1^i\right)$ of VND 2.

3) The LLR information $L_a\left(b_1^i\right)$ extracted by the CND is input to the VND 2 block of Fig. 1a, which extracts the LLR information input $L_a\left(b_1^i\right)$ of the SECCC decoder 1 of Fig. 1a.

Then, the SECCC decoder 1 of Fig. 3 will decode the layer L_1 information with the aid of the resultant *a-priori* LLR $L_a(b_1^i)$ and the soft parity information received from the channel, namely $y_{1,p}$ of Fig. 1a. Afterwards, the classic cyclic redundancy check (CRC) is invoked for detecting, whether the recovered BL L_0 is error-free or not, as shown in Fig. 3. This check results in two possible decoding processes, as shown in Fig. 3 and described as follows:

1) With Inter-Layer Feedback: When the bits b_0 of the BL are not successfully decoded, the iterative IL technique will be activated for exploiting the extrinsic information of layer L_0 fed back from the SECCC decoder 1. In this case, both the solid lines and the dashed lines shown in the decoder of Figs. 1a and 3 will be activated. More explicitly, the "VND2-CND-VND1" block of Fig. 3 will be utilized for extracting the extra LLR information $L_e(b_0^i)$ for BL L_0 based on both the extrinsic LLR $L_e(b_1^i)$ and the soft check information y_{01}. Generally, the "VND2-CND-VND1" block of Fig. 3 represents a similar process to the "VND1-CND-VND2" block of Fig. 3. After this stage, improved *a-priori* information is generated for the BL, which concludes the current IL decoding iteration. Finally, the receiver will return to the beginning of the flow chart shown in Fig. 3. 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. 3.

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. 3. 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 Figs. 1a and 3 will be activated.

Moreover, after decoding BL L_0, the recovered error-free hard bits b_0 may be represented using infinite LLR values, indicating the hard bits $0/1$, respectively. Then, the CND process invoked for generating the LLR $L\left(b_1^i\right)$ shown in Fig. 1a may be derived as follows using the boxplus operation

$$L\left(b_1^i\right) = L(b_0^i) \boxplus L(b_{01}^i) = \operatorname{sign}\left(\tilde{b}_0^i\right) \cdot L\left(b_{01}^i\right), \qquad (2)$$

where $\tilde{b}_0^i$ is the modulated version of the bit b_0^i and the LLR input $L\left(b_{01}^i\right)$ 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. (2) is essentially an LLR sign-flipping operation, it does not affect the absolute value of the LLR information of b_1. This means that in this scenario our proposed IL technique is equivalent to the traditional UEP techniques, where L_0 and L_1 are encoded and decoded independently. Moreover, since L_0 is decoded independently

without feedback from the layer 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.

IV. PERFORMANCE RESULTS

In this section, we benchmark our proposed IL-SECCC-LSSTC system against the SECCC-LSSTC system employing the traditional UEP technique. 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 our proposed system. These sequences were scanned at 15 and 30 FPS, respectively. The "frame-copy" based error concealment tool built into the JSVM H.264/AVC reference codec was activated for combating the effects of channel impairments. The group of pictures (GOP) duration was set to 15, hence an Intra-frame-coded (I) picture was inserted every 15 frames. Correspondingly, both of the two video sequences were encoded into GOPs, consisting of an I frame, followed by 14 predicted (P) frames. Since the bi-directionally predicted (B) frames may impose error propagation on their forward- and backward-predicted dependent-frames, the B frames are disabled in our JSVM configuration. Additionally, only the MGS feature is enabled, when encoding the video sequences into three different quality-ELs, namely layers L_0, L_1 and L_2 using the standardized quantization parameters (QP) of 40, 32 and 24, respectively.

These configurations jointly result in a bitrate of 2297 kbps and a peak-signal to noise ratio (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. Moreover, each SVC H.264/AVC-compressed bitstream was channel encoded and transmitted on a NALU by NALU [2] basis, which is the smallest unit to be decoded by the SVC decoder. Each NALU was protected by CRC codes. At the receiver, each decoded NALU failing to pass the CRC check process was removed before the SVC video decoding process. In all of our experiments, the compressed bitstreams were transmitted 300 times in order to generate statistically sound performance curves.

A. Error Protection Arrangements

Error Protection Arrangements	Code Rates			
	Layer 0	Layer 1	Layer 2	Average
EEP	0.5/0.5	0.5/0.5	0.5/0.5	0.5/0.5
UEP1	0.64/0.64	0.54/0.55	0.46/0.47	0.5/0.5
UEP2	0.41/0.47	0.47/0.49	0.54/0.51	0.5/0.5
UEP3	0.37/0.41	0.46/0.47	0.57/0.54	0.5/0.5
UEP4	0.33/0.36	0.45/0.45	0.60/0.56	0.5/0.5

TABLE I: Coding rates of different error protection arrangements for the *Football/Foreman* sequence.

In our simulations, an overall coding rate of 0.5 was employed for both the equal error protection (EEP) and the UEP schemes for the sake of having a fair comparison. All NALUs within each encoded bitstream were taken into account, when calculating the total number of bits of each layer.

We consider the error protection arrangements shown in Table I for the Football and Foreman sequences. All the arrangements can be applied in both the EEP or UEP schemes with the aid of variable-rate puncturers.

B. System Performance

In this section, we characterize the performance of our proposed IL-SECCC-LSSTC system, where a RSC codec using the generator polynomials $G = [111, 101]$ was employed for the SECCC. Moreover, the system was investigated for all the arrangements presented in Section IV-A.

In Figs. 4a and 4b, we present the BER versus E_b/N_0 curves for different layers of the encoded Football bitstream, noting that as expected for random bits, similar trends were observed for the Foreman sequence. Specifically, the BER versus E_b/N_0 curves recorded for L_0 are displayed in Fig. 4a, which allow us to compare the attainable performance for all the different error protection arrangements of Table I. Observe in Fig. 4a that the IL-SECCC-LSSTC system achieved a reduced BER compared to its SECCC-LSSTC benchmarker. Specifically, the EEP-IL-SECCC-LSSTC scheme outperforms the EEP-SECCC-LSSTC benchmarker by about 1.1 dB at a BER of 10^{-5}. Moreover, among all the error protection arrangements of Table I, the UEP4-IL-SECCC-LSSTC system achieves the best BER performance due to the powerful error protection assigned for layer L_0. Therefore, we may conclude that the UEP aided IL-SECCC schemes are capable of achieving a better performance than that of the traditional UEP aided SECCC-LSSTC systems.

The BER versus E_b/N_0 curves of L_2 of the Football sequence are displayed in Fig. 4b, where the performance of the schemes using the IL technique is seen to be worse than that of the SECCC-LSSTC systems. This is due to the fact that errors may propagate to L_2 from the error-infested layer L_1. Note that this cross-layer error propagation only occurs, when the layer L_1 is corrupted, while L_2 must be discarded in the traditional UEP aided SECCC-LSSTC schemes.

The PSNR versus E_b/N_0 performance recorded for the Football sequence is displayed in Fig. 4c. Observe in Fig. 4c that among all the traditional UEP error arrangements, the UEP3-SECCC-LSSTC scheme of Table I achieves the best performance, attaining an E_b/N_0 gain of 3.1 dB compared to the EEP-SECCC-LSSTC system of Table I at a PSNR of 36 dB. Moreover, the IL-SECCC-LSSTC systems of Table I outperform their corresponding SECCC-LSSTC benchmarkers. More explicitly, the UEP2-IL-SECCC-LSSTC scheme achieves an E_b/N_0 reduction of 2.4 dB compared to the UEP3-SECCC-LSSTC scheme at a PSNR of 37 dB. Alternatively, about 2.8 dB of PSNR improvement may be achieved at an E_b/N_0 of -5 dB. Note that in the E_b/N_0 region of $[-8, -6]$ dB only a modest PSNR improvement can be observed with the increase of the E_b/N_0 value for the UEP4-IL-SECCC-LSSTC scheme in Fig. 4c. This may be attributed to the fact that a weaker error protection was assigned to the ELs, hence this may result in corrupting the ELs at these low E_b/N_0 values. A subjective comparison of the EEP-SECCC-LSSTC, UEP3-SECCC-LSSTC and UEP2-IL-SECCC-LSSTC schemes of Table I employing the Football sequence is illustrated in Fig. 5.

For providing further insights into the system's behavior for video sequences exhibiting lower motion-activity, the PSNR versus E_b/N_0 performance of the IL-SECCC-LSSTC model employing the Foreman sequence is illustrated in Fig. 4d, where the error protection arrangements of Table I were tested. Similar to the Football sequence, about 3 dB of power reduc-

Fig. 5: Comparison of frames at E_b/N_0 of -6 dB for the *Football* and *Foreman* sequences. The three columns (from left to right) indicate frames of the original video, of the traditional UEP scheme, and of the IL-UEP scheme, respectively.

tion is achieved by the UEP3-IL-SECCC-LSSTC compared to the best traditional UEP4-SECCC-LSSTC scheme at a PSNR of 36 dB. Alternatively, about 2.9 dB of PSNR improvement may be observed at an E_b/N_0 of -5 dB. A subjective comparison of the EEP-SECCC-LSSTC, UEP4-SECCC-LSSTC and UEP3-IL-SECCC-LSSTC schemes employing the Foreman sequence is illustrated in Fig. 5.

C. Complexity Analysis

Below we benchmark the complexity of our proposed IL-SECCC-LSSTC system against that of the traditional SECCC-LSSTC systems for the sake of providing insights into the complexity of our scheme. The related performance comparisons are plotted in Fig. 4e. Note that when layer i is corrupted, the complexity imposed by the decoding of the higher-order layers is not taken into account, since they will be discarded by the SVC decoder. In the simulations, each NALU was encoded by the SECCC as a stand-alone packet. Therefore, the number of SECCC decoding operations is representative of the system's complexity. Specifically, the y-axis of Fig. 4e indicates the average number of SECCC decoder activations per NALU. There are a total of 127 NALUs within the SVC H.264/AVC encoded Foreman bitstream, where each NALU was streamed as a stand-alone packet in the simulations.

The complexity of all the traditional SECCC-LSSTC schemes increases upon increasing the E_b/N_0 value due to the fact that at lower E_b/N_0 values layer L_1 and layer L_2 were more likely to be discarded by the decoder, upon receiving a corrupted layer L_0. Moreover, at E_b/N_0 values around -8 dB complexity peaks appear in Fig. 4e for the IL-SECCC-LSSTC schemes due to the fact that the BL cannot be successfully decoded in its own right, which in turn activates the IL feedback more frequently. Here we are only interested in the relatively higher E_b/N_0 values, where achieving a sufficiently high video quality is possible. Specifically, the UEP3-IL-SECCC-LSSTC scheme achieves an E_b/N_0 gain of 3 dB by imposing about 1.5%[2] higher complexity than the UEP4-SECCC-LSSTC scheme at a PSNR of 36 dB. Alternatively, the UEP3-IL-SECCC-LSSTC achieves a PSNR gain of 2.9 dB at the cost of a 1.5% complexity increase compared to the UEP4-SECCC-LSSTC of Table I at a E_b/N_0 of -5 dB.

---

[2]The complexity imposed is read from $E_b/N_0 = -5$ dB and -2 dB for the UEP3-IL-SECCC-LSSTC and UEP4-SECCC-LSSTC, respectively.

(a) BER vs E_b/N_0 for L_0-*Football*

(b) BER vs E_b/N_0 for L_2-*Football*

(c) PSNR vs E_b/N_0 for *Football*

(d) PSNR vs E_b/N_0 for *Foreman*

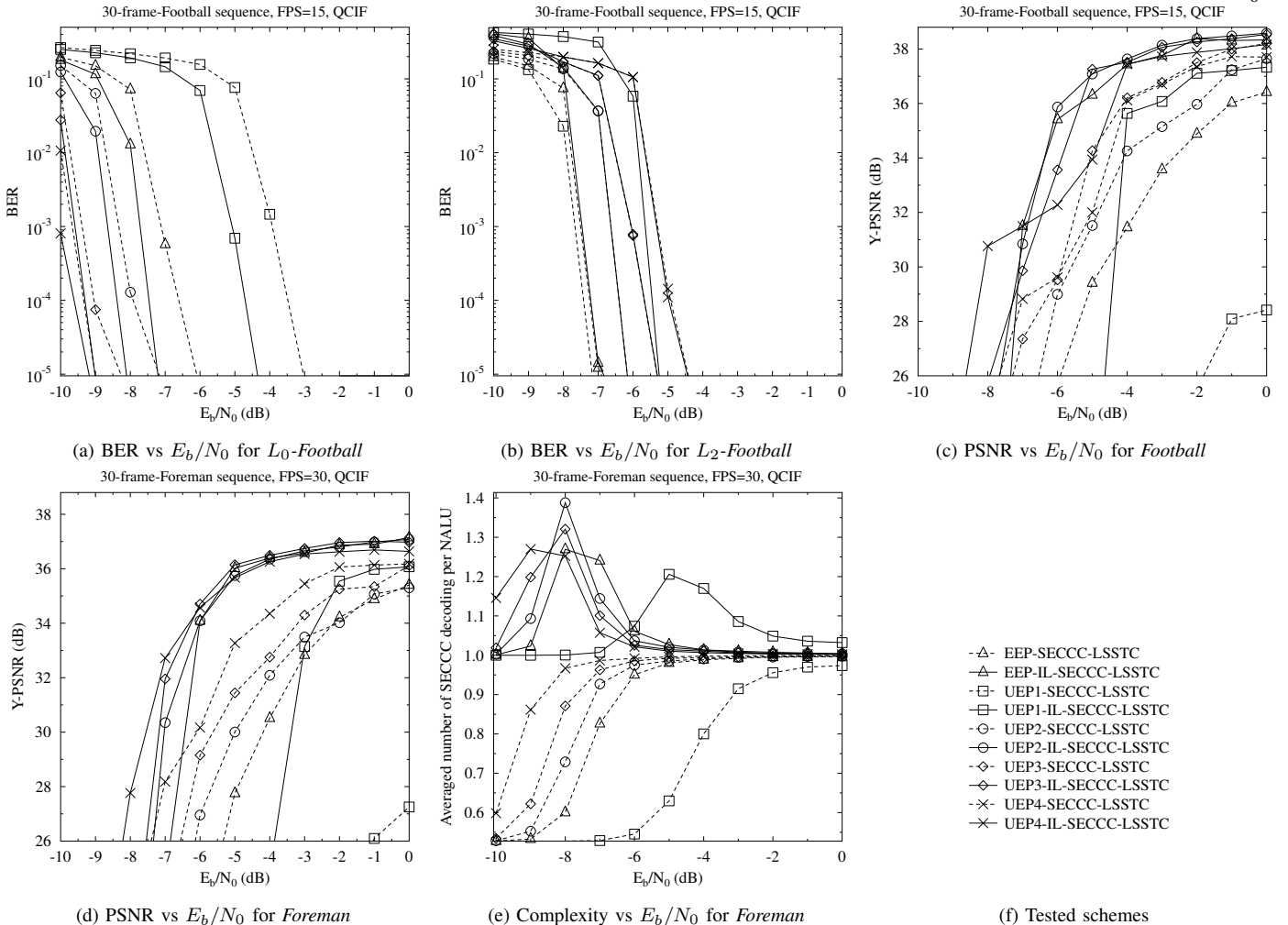(e) Complexity vs E_b/N_0 for *Foreman*

(f) Tested schemes

Fig. 4: BER, PSNR and complexity versus E_b/N_0 performance for *Football* and *Foreman* sequences, where the systems used all the parameters of Table I for transmission over uncorrelated Rayleigh channels.

V. CONCLUSIONS

An IL-SECCC coded video streaming scheme relying on a novel multi-functional MIMO technique was proposed, where SVC based H.264/AVC video coding was employed and the systematic bits of layer L_0 were implanted into the systematic bits of layer L_1 and layer L_2 by employing an XOR operation. At the receiver, the IL-SECCC technique was activated for the sake of achieving an improved system performance. Our simulation results showed that the proposed scheme substantially outperforms the traditional UEP SECCC systems. Explicitly, our proposed technique is capable of achieving a 3 dB E_b/N_0 gain at a PSNR of 36 dB or an approximately 2.9 dB of improved PSNR at a 1.5% complexity increase.

REFERENCES

[1] 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.

[2] 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.

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

[4] 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.

[5] 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.

[6] 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.

[7] Y. Huo, X. Zuo, R. G. Maunder, and L. Hanzo, "Inter-layer FEC decoded multi-layer video streaming," *IEEE Global Telecommunications Conference, 2012*, to appear. Available at http://eprints.soton.ac.uk/343750/1/1569586913.pdf.

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

[9] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Self-concatenated trellis coded modulation with self-iterative decoding," in *IEEE Global Telecommunications Conference, GLOBECOM*, vol. 1, pp. 585–591, 1998.

[10] M. Butt, S. X. Ng, and L. Hanzo, "Self-concatenated code design and its application in power-efficient cooperative communications," *IEEE Communications Surveys Tutorials*, vol. 14, pp. 858–883, Third Quarter 2012.

[11] 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.

[12] L. Hanzo, J. Blogh, and S. Ni, *3G, HSPA and FDD versus TDD networking : smart antennas and adaptive modulation*. Chichester: John Wiley & Sons [s.l.], 2008.

[13] 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 2009*, pp. 36–41, July 2009.

[14] 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.

[15] 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.