



216715 NEWCOM⁺⁺

DR.7.1

First Report on the Activities of WPR7

Contractual Date of Delivery to the CEC: T0+3

Actual Date of Delivery to the CEC: T0+4.5

Editor(s): P. Duhamel, CNRS, with the help of M. Kieffer (CNRS), G. Olmo (CNIT), C. Herzet (CNRS), A. Roumy (CNRS), C. Weidmann (FTW)

Participating institutions: TECHNION, CNIT, CNRS, FTW, AAU

Contributors: G. Olmo (CNIT), A. Roumy (CNRS), M. Varrella (CNIT), E. Paolini (CNIT), M. Chiani (CNIT), C. Weidmann (FTW), C. Marin (CNRS), O. Nemethova (FTW), S. Shamai (Technion), F. Bassi (CNRS), E. Magli (CNIT), C. Herzet (CNRS), A. Diallo (CNRS)

Internal Reviewer(s): P. Duhamel (CNRS)

Workpackage number: DR.7.1

Nature: R

Total Effort Spent: 2.5 pm

Dissemination Level: public

Version: 1

Abstract:

This deliverable is the first deliverable due for WPR7. Its aim is to explain the status of the work on joint source and channel coding/decoding in the various teams collaborating within this workpackage, with respect to the state of the art, and the first directions along which this workpackage intends to contribute. Its structure follows the description of the NEWCOM⁺⁺ proposal, in tasks, the last one being the selection of a few situations which will serve as testbeds for the various proposed algorithms.

Keyword list: Joint source and channel coding/decoding, Unequal error protection, raptor codes, ideal codes, MBMS, packet-level coding, multiple description coding, fading channels, network coding, broadcast, multicast, cross-layer optimization, H264, distributed source coding, multiterminal coding, relay channel, distributed arithmetic codes, error correcting arithmetic codes

TABLE OF CONTENTS

1	Introduction	3
2	Task TR7.1 JSCC/D and the OSI Layers	4
2.1	State of the art	4
2.1.1	Permeable soft protocol stack	4
2.1.2	Optimal JSCD algorithms with reduced complexity	6
2.1.3	Cross-layer optimization	8
2.2	Perspectives	15
2.2.1	Permeable soft protocol stack	15
2.2.2	Optimal JSCD algorithms with reduced complexity	16
2.2.3	Cross-layer optimization	16
3	Task TR7.2 New tools for JSCC/D	17
3.1	State of the art	17
3.1.1	JSCC with fountain codes	17
3.1.2	JSCC using UEP	21
3.1.3	JSCC using MDC	23
3.1.4	JSCC for fading channel without channel state info	23
3.2	Perspectives	24
3.2.1	Comparison between RS and DFC	24
3.2.2	JSCC for fading channel	25
3.2.3	Merging successive refinement and broadcast channel coding	25
3.2.4	Joint source coding and network coding	26
3.2.5	Multiple description with side information at the decoder	26
4	Task TR7.3 Tools for multi-terminal JSCC/D	28
4.1	State of the art	28
4.1.1	Distributed source coding: theoretical background	28
4.1.2	Robust distributed source coding	31
4.2	Source coding with side information of time-varying quality at the decoder	40
4.2.1	Practical design	40
4.3	Perspectives	43
4.3.1	Distributed coding with joint sparsity models	43
4.3.2	Robust DSC	43
4.3.3	Non asymmetric DSC: robust estimate of the sources, when the difference pattern (between the sources) has error	43
4.3.4	Robust DSC employing distributed arithmetic codes	44
5	Task TR7.4 Tools for evaluating the efficiency of JSCC/D methods	46
5.1	State of the art	46
5.1.1	Performance characterization of arithmetic-coding-based JSCC/D using distance spectra	46
5.2	Bounds on the Wyner-Ziv rate distortion function for GE and GBG correlation models	53
5.3	Perspectives	55
5.3.1	Distance spectra of AC-based and general JSC codes	55
5.3.2	Distance properties of distributed arithmetic codes	55
5.3.3	Models and bounds for distributed source coding with sparse correlation structures	55
5.3.4	Distributed joint source channel coding: comparison of source coding (quasi-arithmetic) and channel coding based methods	56

5.3.5	Analysis of joint source channel coding: analysis of the puncturing of quasi-arithmetic coding	56
6	Task TR7.5: Study of practical, long-term solutions plus test-beds	57
6.1	Actual impact of JSCC/D in standards	57
6.1.1	Hidden cases where separation theorem is not really used	57
6.1.2	Recent cases where JSCC/D has contributed to standards	58
6.2	Emergence of the concept in products	58
6.3	Possible impact in the near future depending on the context	58
6.3.1	Point to point communications	59
6.3.2	Broadcasting	59
6.3.3	Multicasting	59
6.4	Longer term situation : Video streaming over peer to peer networks	59
6.5	Testbeds	60
7	Conclusion	61

1 INTRODUCTION

This section sets a framework for the document, and contains a short description of it, how it relates to the overall Project and Work Package(s), how it has been prepared, and ends by defining concisely the content of each section of the document. It shall be followed by:

The document has been written in a cooperative manner : The global outline follows the description of the tasks provided in the joint program of activities of the proposal. A first draft of the table of content was provided during a meeting between Aline Roumy (CNRS), Claudio Weidman (FTW), and Michel Kieffer (CNRS), assigning subsections to the corresponding set of partners. Each section has then been edited by a task leader as follows:

- Task TR7.1 JSCC/D and the OSI Layers (Michel Kieffer, CNRS)
- Task TR7.2 New tools for JSCC/D (Gabriela Olmo, CNIT)
- Task TR7.3 Tools for multi-terminal JSCC/D (Cédric Herzet, Aline Roumy, CNRS)
- Task TR7.4 Tools for evaluating the efficiency of JSCC/D methods (Claudio Weidmann, FTW)
- Task TR7.5: Study of practical, long-term solutions plus test-beds (Pierre Duhamel, CNRS)

2 TASK TR7.1 JSCC/D AND THE OSI LAYERS

Description of the work as in the NEWCOM++ proposal:

- 7.1.1- Compatibility: traditionally, joint source and channel coding (or even decoding) was considered as incompatible with the classical separation of the layers in the classic OSI structure. However, results obtained in NEWCOM demonstrated that in some cases, this incompatibility problem could be circumvented. Compatibility of the JSCC/D strategy will thus be re-considered and studied to address the contradiction vis-a-vis the layer-independent model. In particular, compatible re-synchronization strategies of source codes in the presence of channel noise will be developed. Optimal JSCD algorithms with reduced complexity will be designed.

More generally, techniques allowing or at least facilitating the use of JSCC/D in practical situations, accounting for the layers of the OSI structure should find their place in this task. In order to ensure compatibility, the structure of the encoder should not be modified.

- 7.1.2- Cross-layer optimisation of the redundancy at the various layers. Another related problem is that, in the process of globally optimizing the efficiency of the multimedia transmission, some redundancy is introduced at many layers of the protocol. The best allocation of this redundancy, within the context of JSCC/D, has not been found in the literature, thus it will be addressed here.

This second part goes a step further, allowing the transmitter to be modified.

2.1 State of the art

2.1.1 Permeable soft protocol stack

Due to bandwidth constraints, efficient transmission of multimedia contents requires the use of some source coding scheme [167]. However, compressed data are very sensitive to transmission errors. A single corrupted bit may lead to a loss of a large amount of multimedia data at the receiver. Consequently, the bitstream entering the source decoder has to be almost error-free, according to Shannon's separation principle [176].

This constraint is hardly satisfied when considering transmission over wireless channels. The data stream at receiver side may be heavily corrupted and not directly usable by the source decoder. A first solution to this problem consists in grouping data into packets protected by an error-detection code (CRC or checksum) [26, 106]. Packets which have not been correctly received are identified, and can then be re-transmitted. Nevertheless, retransmissions may become difficult in scenarii with strong delay constraints, *e.g.*, for visiophony, or may even become impossible when broadcasting data, *e.g.*, in satellite television.

In such situations, the standard solutions make use of very strong error-correcting codes (*e.g.*, turbo-codes, LDPC) at *Physical* (PHY) layer combined with packet-erasure codes at intermediate protocol layers [122, 159], see, *e.g.*, MPE-FEC in DVB-H [53]. The redundancy introduced by these codes may however be oversized when the channel is good, reducing the bandwidth allocated for the data. In bad channel conditions, some corrupted packets still cannot be recovered, and are assumed lost. Error-concealment techniques [77, 105] may then be used by the source decoders at *Application* (APL) layer. They exploit the redundancy (temporal and/or spatial) found in the multimedia data for estimating the missing information.

In the recent years, joint source-channel decoding (JSCD) techniques have been proposed to correct damaged packets. These methods involve robust source decoders, which exploit the inherent redundancy in the received packets for correcting errors. Several sources of redundancy have been identified. Constraints in the syntax of variable-length source codes [31, 74, 94, 149, 189] have been used first. Redundancy due to the semantic of the source coders [23, 141, 165, 191] improve significantly the performance of robust decoders. Further redundancy due to the packetization of compressed data has been used in [111]. Altogether, the various redundancies can attain an unexpected amount. Furthermore, redundancy introduced by channel codes at PHY layer can also be used in combination with residual redundancy to build iterative decoders as in [18, 140, 188]. These joint decoding schemes provide improved performance

when compared to classical schemes, and could be of great use in many applications. However, they are not compliant with the standard protocol stacks in several aspects: First, they require exchange of soft information between the channel decoders at PHY layer and the robust source decoders at APL layer.

This part of the workpackage aims at proposing tools to make this feasible (*e.g.*, in a mobile receiver containing all layers of the protocol stack, and thus, able to choose to forward soft values between layers). The main compatibility problem is that standard protocol stacks do not even allow damaged packets to reach the APL layer, the main reason being that the errors may impact some essential information contained in the headers, which is necessary even for the robust APL decoders.

This problem has already motivated the proposal for the so-called MAC-lite variant of the protocol, which already allows the CRC to protect the MAC header only. Nevertheless, there is a variety of headers, as described on Figure 1, and the problem does not seem to be fully solved, as argued below.

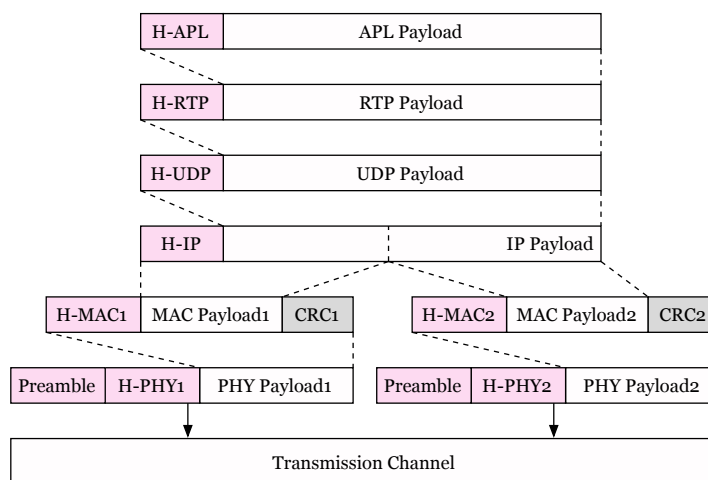


Figure 1: Protocol stack for multimedia transmission over WiFi

The packetized multimedia transmission is usually based on an RTP/UDP/IP protocol stack [106]. Figure 1 illustrates an example of segmentation and encapsulation mechanisms implemented at each protocol layer in case of multimedia packet transmission with the 802.11 standard (WiFi) [7]. Error detection mechanisms implemented at each layer are listed below.

- At PHY layer, a fixed preamble is used to help PHY packet segmentation; a CRC protects the header fields. Received packets with damaged headers are discarded.
- At MAC layer, a CRC protects the corresponding header and payload. When an error occurs, the packet is retransmitted.
- At IPv4 layer, the header fields are protected by a checksum. Received packets with damaged headers are discarded.
- At UDP layer, a checksum protects the header and the payload. When an error occurs, the packet is discarded.

The error detection mechanisms provided by CRCs and checksums, combined with a retransmission mechanism at MAC layer, allow APL layer to receive only error-free packets. The price to be paid is potentially very important transmission delay due to the retransmissions which may become very frequent at MAC level when the channel conditions worsen, or frequent use of error concealment when errors occur at IPV4 or UDP layers (generally due to time out constraint: the limit on the number of retransmissions at MAC level has been exceeded).

Joint source-channel decoding techniques allow at APL layer many errors to be corrected based on soft information provided by lower protocol layers. The recently introduced UDP-Lite [110] mechanism, combined with lower permeable protocol layers [89, 90] allow packets with binary errors to be fed to the APL layer. With UDP-lite, a checksum protects a limited number of bytes (generally including the UDP-Lite, RTP, and APL header fields). Thus, packets with erroneous headers are still discarded, even if their payloads contain only few erroneous bits. Considering the order of magnitude of the packets and the various headers in actual wireless communications when tuned for difficult situations [29, 205], this may happen more than expected. The bottleneck of such permeable transmission schemes is the fact that packets are discarded due to erroneous headers.

2.1.2 Optimal JSCD algorithms with reduced complexity

In many source coding standards, entropy coding with variable length codes (VLC) is employed. Huffman-like VLC are used, *e.g.*, in JPEG [82], H263+ [83] or in the baseline profile of H264 [85]. Arithmetic VLC are used in JPEG 2000 [186] or in the main profile of H.264 [85].

It is well known that bitstreams containing VLC encoded data are particularly sensitive to transmission errors, occurring for example when the bitstream is sent through radio-mobile links. This high sensitivity is mainly due to frequent loss of synchronisation of the decoder when errors occur.

Many JSC decoding techniques for VLC rely on specific trellis construction and pruning, which result in various efficiencies. Their complexity may become intractable when realistic Huffman codebooks are considered or when practical arithmetic codes are considered, due to the number of nodes in the trellis. More generally, complexity is a very important issue for JSCD techniques.

Solutions have been proposed, *e.g.*, in [109], where a simplified stack-based [6] decoding algorithm is proposed for Huffman VLCs. When the channel is not too noisy, the optimal MAP estimate of the decoded bitstream may be obtained faster than with a standard trellis-based technique. However, the complexity is reduced mainly for good channel conditions, while it may remain intractable in poor channel conditions. This comes in contrast with trellis-based techniques, for which the complexity remains constant (but high). The M-algorithm [6] and its soft-output variant [218] allow to build reduced but constant complexity estimation algorithms for variable-length codes. The price to be paid in this case is a suboptimality of the decoder when the noise level increases: since only M decoding hypotheses are kept at each iterations, the optimal solution may be dropped due to complexity limitations.

2.1.2.1 Compacting Huffman VLC tables

A nearly optimal reduced-complexity algorithms have been presented in [131], which provides a tool for compacting Huffman VLC tables. The codewords are grouped into a minimum number of classes. Decoding algorithms may then work on a reduced number of classes instead of working on the whole set of codewords. VLC table reduction techniques have been used previously. Some works remove some unlikely codewords [139], at the cost of a reduced coding efficiency.

The VLC table simplification algorithm proposed in [131] gathers VLC codewords of the same length in classes. For example, if one considers the 16 codewords of 8 bits used to encode the texture in H263+,

$$\mathcal{A}_8 = \{00100000, 00100001, \dots, 00101110, 00101111\},$$

one notices that all of them share the same 4-bit prefix 0010, and that all other bits take all possible values. Thus, this set of codewords may be represented by 0010XXXX, where X represents either 0 or 1. Using this simplification, 16 parallel branches in a trellis between nodes $(a - 8)$ and (a) are replaced by a single branch. When the *a priori* probability of each grouped codeword is the same or when ML decoding is performed, this simplification is harmless and decoding techniques inspired from that of punctured convolutional codes [34] may be put at work. When there are discrepancies between *a priori* probabilities, MAP decoding may become suboptimal. While the general case is more intricate than the

simple example above, the simplification algorithm presented in [132] is a generalization of the previous example.

The performance of the texture block localization algorithm presented in [111] is compared using two VLC tables: the initial H263+ VLC table and the compacted VLC table according to [132]. Data have been put in packets of about 1024 bits and sent over an AWGN channel. A soft-input Viterbi algorithm is employed to localize the texture blocks in packets containing INTER encoded texture blocks and in packets containing INTRA encoded texture blocks. Performance is evaluated in terms of block location error rate in Figures 2 and 3, and in terms of average computing time in Table 1.

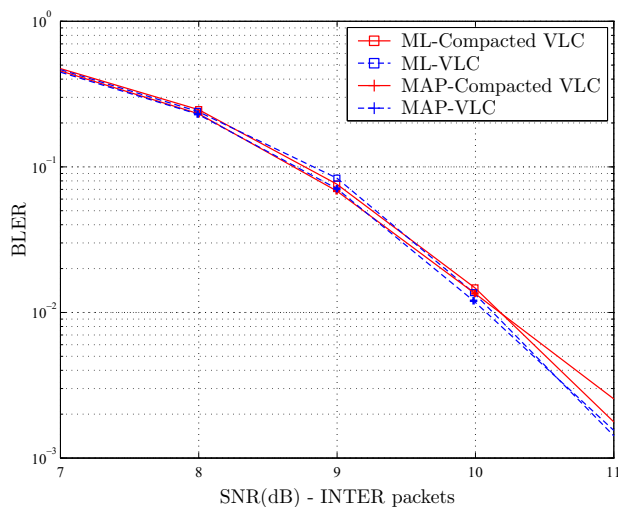


Figure 2: Block location error rate for INTER packets

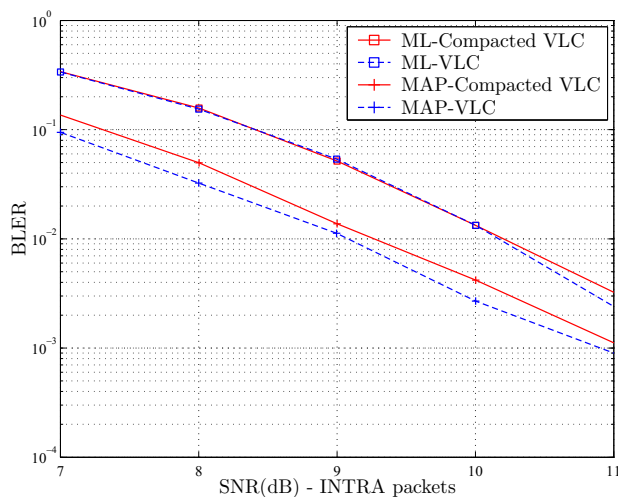


Figure 3: Block location error rate for INTRA packets

When comparing computing times, decoding with a simplified VLC table is three times faster than with the initial VLC table.

	ML		MAP	
	Init. table	Simpl. table	Init. table	Simpl. table
INTRA	165.77	62.53 (-62%)	242.01	92.41 (-62%)
INTER	356.18	108.46 (-70%)	445.76	131.03 (-71%)

Table 1: Mean decoding time (ms/block)

2.1.2.2 State aggregation for Huffman VLC

A novel set of state models and the corresponding trellises for the estimation of the Hidden Markov chain has been proposed in [88]. The state model is defined by both the internal state of the VLC decoder (*i.e.*, the internal node of the VLC codetree) and the rest of the Euclidean division of the symbol clock by a fixed parameter T . Therefore, the approach consists in aggregating states of the bit/symbol trellis which are distant of T instants of the symbol clock. If $T = 1$, the resulting trellis is equivalent to the usual bit-level trellis proposed by Balakirski. If T is greater or equal than the symbol sequence length $L(S)$, the trellis is equivalent to the bit/symbol trellis. The intermediate values of this parameter allow to gracefully trade complexity against the estimation accuracy. The state aggregation leads to close-to-optimum estimations with significantly reduced complexity. The complexity can be further reduced by running separate estimations on trellises of parameters T_1 and T_2 of the sequence of emitted symbols. It has been shown that if T_1 and T_2 are relatively prime, and if the two corresponding sequence estimates S_1 and S_2 obtained with a Viterbi algorithm are such that $S_1 = S_2$, then, this estimate is the same as the one obtained with the trellis of parameter $T_1 \times T_2$.

2.1.3 Cross-layer optimization

Cross-layer optimization offers many opportunities to improve end-to-end distortion without the need to redesign all components of the communication chain, but rather by exploiting information from one layer in existing optimization mechanisms in another layer. An instance of this approach was explored in the NoE NEWCOM, namely the use of feedback mechanisms provided by mobile communications standards in order to optimize the perceived quality of mobile video streaming. Cross-layer scheduling has been chosen as the main tool to accomplish this. The following gives a brief overview of the results reported in [137].

2.1.3.1 Cross-layer scheduling for reduced video distortion

The transmission of real-time video services over mobile networks is particularly challenging, especially due to the fact that limited radio resources lead to high link error probability. The redundancy has to be kept low due to system capacity limitations. Furthermore, the impact of a lost packet on distortion is not uniform across the stream.

In UMTS (Universal Mobile Telecommunication System) DCH (Dedicated Channel) downlink direction, scheduling is performed in the Radio Network Controller (RNC) at the Radio Link Control (RLC) protocol layer. The scheduling mechanism is not standardized, but rather implementation-specific. Scheduling of the packets belonging to one user's stream can considerably improve the quality of service at the receiver, especially in case of streams containing informations with different importance. In [96] and [199] (opportunistic) scheduling algorithms are presented which make use of the characteristics of the streamed video data. In [96], the more important parts of the video stream are transmitted prior to the less important ones, in order to ensure more opportunities for retransmissions in case of error; in [199], the priority-based scheduling exploits the diversity gains from the channel variations when having more than one stream. Another approach is [103], where prediction of link errors is used in connection with call admission control and scheduling algorithms in order to avoid overloading the system, thus improving the quality of the services with higher priority. Rate-distortion optimized (RDO) video streaming is proposed in [95]. Here, a model for delay and loss in the Internet is used to perform scheduling at the

IP layer. In [98], we proposed a simple and efficient IP packet based mechanism that predicts the transmission intervals with low probability of errors and postpones the transmission of packets containing the intra-predicted frames to these intervals, meanwhile transmitting packets containing inter-predicted frames. The following extends that work and proposes scheduling for UMTS at the RLC packet layer that utilizes link error prediction and a novel reference-free distortion estimation. The underlying model considers losses at the RLC PDU (Protocol Data Unit) level, rather than at the frame/slice level; we assume that the video decoder can use the information in a NALU (Network Abstraction Layer Unit, typically encapsulated further in an RTP/UDP/IP packet) up to the first missing RLC PDU as proposed in [136].

2.1.3.2 Distortion Model

Various distortion models have been proposed in literature so far. In [24], a semi-analytical model is proposed for H.264/AVC considering the impairments resulting from the loss or damage of the entire frame. Another approach can be seen in [97], where the visibility of packet loss is modeled depending on a set of motion and spatial video parameters. The method proposed in [208] combines the analysis-by-synthesis approach with model-based estimation. Full decoding of the damaged frame as well as the knowledge of the undamaged original are required to obtain the resulting distortion; the distortion caused by propagation is estimated by an empirical model. In [179], the distortion over the whole sequence is estimated for both compression and packet loss impairments.

Our proposed scheduling mechanism deals with RLC PDUs. We thus need a distortion model with RLC PDU granularity, instead of the more common IP/RTP/NALU granularity. The task of designing such a model is somewhat simplified by the fact that the decoder will not be able to recover information in PDUs after the first lost PDU of a NALU. Thus, if a NALU is made up of k PDUs, only k cases need to be tested. The impairments caused by the loss of a PDU will propagate from the frame of origin to all frames that reference it. Thus, we subdivide the distortion estimation into two steps:

1. determine the *primary distortion*, i.e. the distortion within the frame caused by a lost PDU,
2. determine the *distortion propagation*, i.e. the distortion in the following frames until the end of the GOP.

Note that the distortion is estimated after error concealment, and therefore depends on the performance of the latter. The estimation considers the position and amount of lost data, as well as the size of the corresponding picture area. The number of lost/damaged MBs required by the model can be obtained without fully decoding the video as a proportion of the MBs per NALU packet (if it is fixed) corresponding to the proportion of lost RLC PDUs, by reading of the entropy encoded stream (without decoding the video) or by extra signaling.

The acquisition of the data necessary to build the model is detailed in [137]. In order to build a model based on the obtained data set, the estimated primary distortion $\hat{\epsilon}_0$ has to be represented as a function of the number of missing RLC packets (PDUs) N and the corresponding number of missing macroblocks M :

$$\hat{\epsilon}_0 = f(M, N). \quad (1)$$

The task is to find an appropriate model function $f(\cdot)$ and its best fitting parameters for the given scenario (i.e. encoder settings/error concealment assumptions).

The distortion grows monotonically with increasing number of missing RLC PDUs. We observed that this growth is better described by a rational rather than exponential function. With increasing number of missing MBs, the distortion decreases. Thus, we propose the distortion model

$$\hat{\epsilon}_0 = e^{a \cdot M + \frac{b}{N} + c} \quad (2)$$

with parameters a , b , and c determined by least square fitting. For the model data set, the following values were obtained: $a = -0.041$, $b = -6.371$, and $c = 4.482$.

In order to model the error propagation in a GOP, all obtained distortion sequences $\{\varepsilon_k\}_{k=1,2,\dots}$, where k is the frame index, were normalized by the corresponding primary distortion ε_0 . On average, an exponential decrease was observed:

$$\widehat{\varepsilon}_k = \widehat{\varepsilon}_0 \cdot e^{-s \cdot k}, \quad (3)$$

where s determines the speed of decay, which depends strongly on the nature of the motion in the sequence. Since we assume that only M and N are available, the error propagation has to be estimated by an average over all sequences in the model data set. The decay thus obtained is $s = 0.08$.

Video sequence	$\rho(\varepsilon_0)$	$\rho(\varepsilon_k)$
1	0.86	0.84
2	0.88	0.81
3	0.85	0.78
4	0.68	0.80
5	0.84	0.77
6	0.87	0.85
7	0.79	0.67
8	0.83	0.73

Table 2: Performance of the distortion estimation for video sequences in the model data set (1–5) and new sequences (6–8).

The estimation performance of the proposed method was tested for all video sequences in the model data set, as well as for three additional ones. Table 2 summarizes the results in terms of Pearson correlation coefficient ρ for both primary distortion and error propagation model, obtained by decoding of all possible RLC PDU loss error patterns.

The performance of the distortion estimation varies for different video sequences. Especially scene cuts and frames with significant amount of I-MBs cause lower correlation. However, in the scheduling application absolute distortion values are not as important as the relation between the distortions caused by different error patterns.

2.1.3.3 Link Error Model

Based on the analysis of the link error characteristics from UMTS DCH (Dedicated Channel) measurements in live networks [100], we know that there is correlation and thus memory in the link. Furthermore, in [99] the authors have shown that the impact of the UMTS DCH quality based power control — which provides error prediction properties — on the link error characteristics becomes significant especially in static scenarios.

Defining the number of subsequently received erroneous TBs (Transport Blocks) as error burst, the successive error-free TBs as error gap, and erroneous TBs which are separated by less than or equal to 12 (for the 384 kbit/s bearer) error-free TBs as error cluster (Fig. 4), we can describe the UMTS DCH link error process in the following way.

Due to the effects of jointly coded and interleaved TBs within one TTI (Transmission Timing Interval), the UMTS radio link produces error bursts which are separated by short gaps (gap length ≤ 12). These bursts and short gaps, lying within up to three successive TTIs, then form an error cluster. The power control algorithm of the UMTS DCH separates successive clusters by long gaps in order to reach the link quality target (block error rate of 1% in the considered networks) as illustrated in Fig. 4.

For a prediction of future link errors, the knowledge of past transmission errors is necessary at the transmitter. That is accomplished in UMTS via the RLC AM (Radio Link Control Acknowledged Mode). Of course, the transmitter receives the RLC AM error feedback information from the receiver after a certain feedback delay d_{FB} which is in the order of three TTIs (= 30 ms for the measured 384 kbit/s

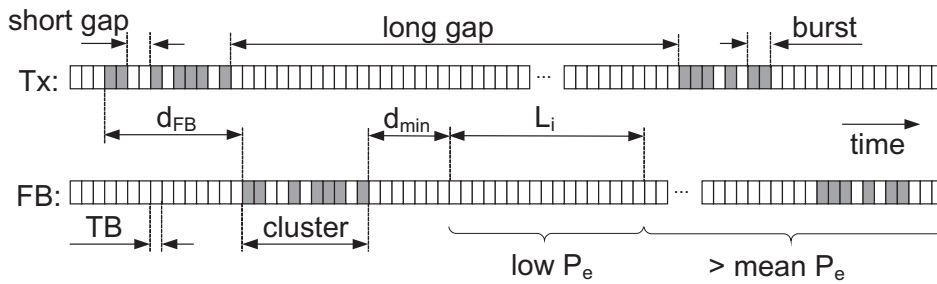


Figure 4: Schematic illustration of the link error characteristics for UMTS DCH in downlink direction.

bearer). Thus, error prediction within one error cluster is not possible but the location of the error clusters or in other words the length of the long gaps can be estimated.

We have shown in [99] that the lengths of the long gaps are not geometrically distributed, thus providing memory. Furthermore, it was argued that the long gaps follow a Weibull distribution with shape parameter $a > 1$, meaning the distribution is convex from zero up to the inflection point (mode). This leads to a region with a very small probability of having gaps between the maximum of the short gap lengths ($\cong 12$) and about 300 TBs [99]. Therefore, if within a time interval d_{min} (larger than the maximum of the short gaps) after the last error there are no error reports, we can assume to be out of an error cluster, within a long error gap and thus at the beginning of an interval with a very low probability for the occurrence of another error cluster, due to the convex distribution of the long gaps.

This capability of predicting the start of intervals with very low error probability has already been used by the authors for cross-layer optimized video streaming (presented in [98]), where a significant gain in quality of the streamed video was obtained. For the scheduling mechanism presented in this work it is not enough to detect the beginning of the intervals with very low error probability at d_{min} after the last error report, but it is also necessary to estimate the length L_i of these intervals.

To determine the length L_i of the intervals with lower error probability, we make use of a theoretic analysis with long gaps separating single error events only. In Fig. 5 the simulated error probability P_e at a point x TBs after an error event (in practice corresponding to the occurrence of an error cluster) is presented with Weibull-distributed gap lengths. We can see from this curve that the error probability is very small right after an error event, staying below the total mean error probability until approximately 2500 TBs. In practice, we may estimate L_i via the intersection of this conditional mean curve with the unconditional mean \bar{P}_e .

To obtain an analytic expression for L_i we first compute the steady state error probability $P_{e,ss}$. Let L be the random variable measuring the time (distance) between two error events. Assuming that the error process is stationary, Kac's lemma [93] implies that

$$P_{e,ss} = \frac{1}{E\{L\}}, \quad (4)$$

where $E\{L\} = \sum_{i=1}^{\infty} i \cdot P(L = i)$ is the average recurrence time corresponding to the average gap length.

It turns out that the stationarity assumption holds over longer ranges, but not in immediate succession to an error event, where the (measured) error probability is actually lower than the steady state error probability $P_{e,ss}$. Therefore, we approximate the probability $P_e(l)$ of having an error at the l th TB after an error event assuming that the steady state is reached before the third error after the initial error event:

$$P_e(l) \cong P(L_1 = l) + P(L_1 + L_2 = l) + P_{e,ss} \cdot P(L_1 + L_2 < l). \quad (5)$$

Here, L_1 and L_2 denote the gap lengths before the first and the second error, respectively. The result (blue dashed curve in Fig. 5) shows that taking the intersection between this curve and the total mean as the end of L_i gives a good approximation, whereas the assumption of reaching the steady state after the first error leads to an underestimation of L_i (grey dashed curve).

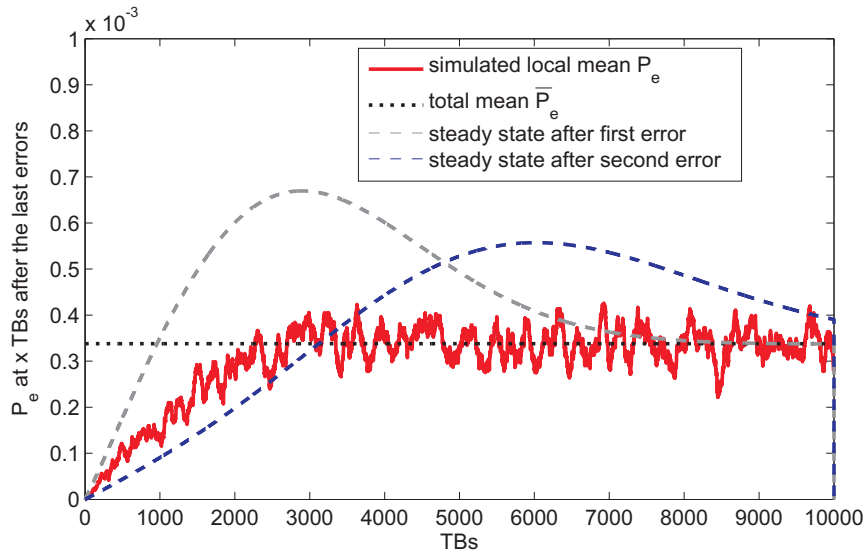


Figure 5: P_e for Weibull distributed long gaps (Weibull shape parameter $a = 2.176$, scale parameter $b = 3350$).

2.1.3.4 Scheduling Mechanism

The proposed scheduling mechanism using link error prediction and distortion estimation is illustrated in Fig. 6.

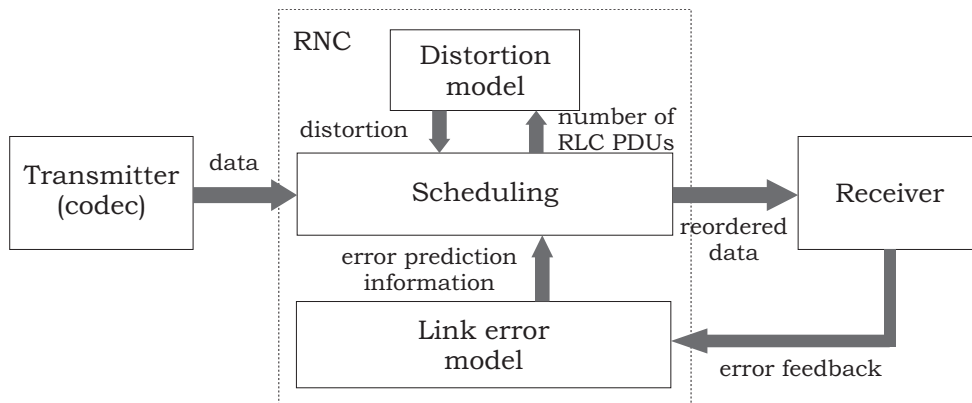


Figure 6: Functional scheme of the proposed scheduling mechanism.

The performance of the scheduler can be controlled by the scheduling buffer length $L_b \geq L_i$ and the maximum acceptable delay d_{\max} . The delay d_{\max} depends on the requirements of the service; it should be chosen smaller than or equal to the application's play-out buffer length. Moreover, the maximum gap length provides a guiding value for determining the size of d_{\max} .

Based on the RLC feedback from the receiver, the link error model identifies the transmission intervals with low error probability and delivers this information to the scheduler. The distortion model provides an estimation of the *cumulative distortion* $\hat{\epsilon}_\Sigma$ that would be caused by the loss of a particular RLC PDU. This is calculated for every RLC PDU in the scheduling buffer as

$$\hat{\epsilon}_\Sigma = \sum_{k=n}^m \hat{\epsilon}_{k-n}, \quad (6)$$

where n denotes the number of the frame where the RLC PDU under consideration is located and m is the number of frames in the GOP.

At each step, if an RLC PDU remained in the scheduling buffer longer than d_{\max} , then it is scheduled immediately. Otherwise, at each step within the time interval L_i , the RLC PDU associated with the highest estimated cumulative distortion $\hat{\epsilon}_{\Sigma}$ is chosen from the scheduling buffer. Outside the interval L_i , the RLC PDU with the lowest $\hat{\epsilon}_{\Sigma}$ is scheduled. If several RLC PDUs have the same (lowest or highest) value of $\hat{\epsilon}_{\Sigma}$, the oldest among them is scheduled.

2.1.3.5 Experimental Results and Conclusions

We implemented all presented methods using the Joint Model reference software [73]. We chose the music video clip sequence “videoclip” as a test sequence, since it contains a variety of different scenes separated by cuts and gradual transitions. The sequence was encoded with the same settings as those used for the distortion model. Its transmission was simulated 1500 times (representing ~ 3 h of video) for each of the considered methods to obtain sufficient statistics. The CRC of the RLC PDUs ($q = 320$ bits) was used to detect errors. Decoding of a video slice was performed up to the first erroneous RLC PDU within the corresponding packet, as proposed in [136]. The following methods were compared:

- **Without scheduling:** decoding using unchanged link error traces obtained from measurements in a live UMTS network,
- **Scheduling limit:** achievable limit under assumptions that the position of errors is known and both d_{\max} and L_b are larger than one entire video sequence,
- **I scheduling:** prioritized scheduling of the I frames as proposed in [98], but assuming the usage of the correctly received data from NALU up to the first missing RLC PDU,
- **RLC scheduling:** configurations of the proposed method for different d_{\max} and L_b .

Here we present simulations for combinations $d_{\max} = 8400$ TBs, $L_b = 3600$ TBs and $d_{\max} = 3600$ TBs, $L_b = 1200$ TBs (denoted further as 8400/3600 and 3600/1200), which corresponds to the maximum delay of 7 and 3 seconds, respectively.

Fig. 7 shows the empirical CDF of frame PSNR for the presented methods. It illustrates well the difference between the I scheduling and the RLC scheduling method. All CDFs approach one at the PSNR = ∞ (error-free frames). The I scheduling method globally reduces the number of erroneous frames

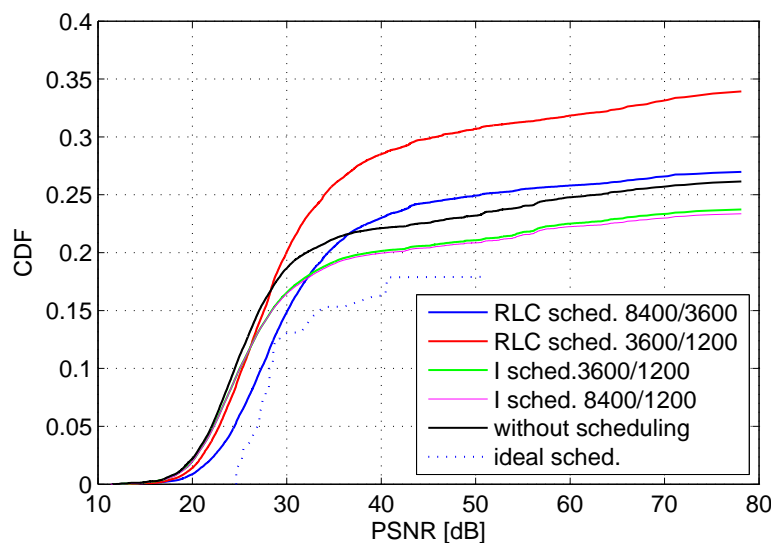


Figure 7: CDF of frame PSNR achieved by presented scheduling methods.

since it moves the errors from the first frame in GOP to the later ones, making the error propagation chain shorter. The proposed RLC scheduling reduces the number of frames with higher distortion and

consequently increases the number of frames with lower distortion. It can even happen that the total number of frames containing an error (even small) increases after applying RLC scheduling. This is caused by burst errors possibly affecting multiple frames, since we do not consider the combinations of errors in time; i.e. if an error occurs followed by another error in the next frame at the same position, then the cumulative distortion is smaller or equal to the sum of the two distortions. However, as we do not have an information about the motion between the two frames, we cannot reliably determine the additional distortion caused by another error. Moreover, considering all possible combinations of two and more errors would increase the complexity of the scheduler considerably.

Fig. 8 depicts the PSNR over the frames of the tested “videoclip” sequence. Again, the difference between the I scheduling and RLC scheduling can be seen. The PSNR at the end of the sequence decreases for the RLC scheduling. This effect is caused by the high predicted distortions for a loss of the first frame. In our simulation we encoded the video sequence 1500 times, thus the first frame follows after the last one again. Since the first frame is concealed only by spatial error concealment, the corresponding distortion caused by a loss of its parts is higher.

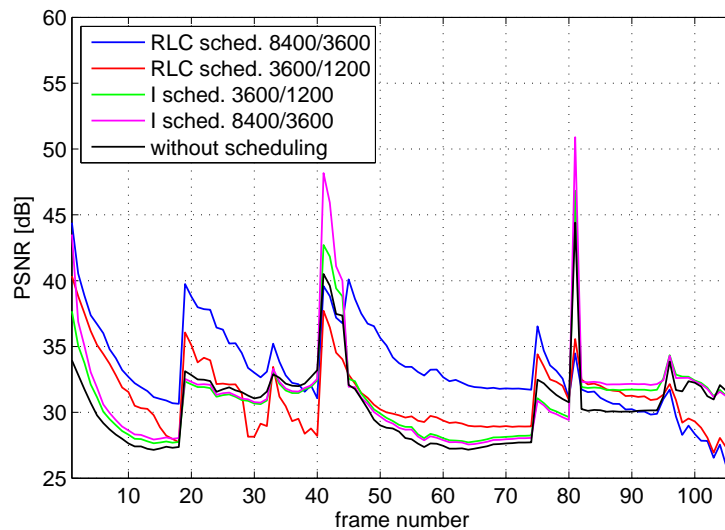


Figure 8: PSNR in time: comparison of presented scheduling methods.

In Table 3, the proportion of error-free frames is summarized together with the mean distortion for all presented methods. The I scheduling is not as sensible to smaller d_{min} as the RLC scheduling.

Conclusions: We investigated an RLC layer based scheduling for UMTS that makes use of both distortion estimation and link error prediction. We proposed a distortion model that allows for estimation of distortion caused by loss of RLC PDUs without decoding the video stream. Based on the prediction of intervals with low error probability, the scheduler prioritizes the RLC PDUs whose loss would lead to highest distortion. We used error traces obtained from live UMTS network measurements for our experiments. The proposed scheduling method gains in average 2 dB compared to common (in-order)

Method	Error-free frames [%]	Mean PSNR [dB]
Without scheduling	73.86	29.29
I sched. ($d_{max} = 3600$ TBs)	76.27	30.26
I sched. ($d_{max} = 8400$ TBs)	76.63	30.32
RLC sched. (3600/1200)	66.07	30.44
RLC sched. (8400/3600)	73.01	32.03
Scheduling limit	82.09	35.53

Table 3: Number of erroneous frames and mean PSNR for the presented scheduling methods.

scheduling. Since the loss of packets containing I frames usually causes the highest distortions, we compared the proposed scheduler (based on the distortion model) with a scheduler that prioritizes the packets belonging to I frames. The gain achieved by using the distortion model is then still about 1 dB. These gains are significant from the point of view of user perception.

2.2 Perspectives

2.2.1 Permeable soft protocol stack

one aim of the work in this topic consists in using the various source of redundancy in the protocol stack to recover headers, increasing the amount of packets that can be used for robust decoding at APL layer. As a result, the efficiency of joint source-channel decoding techniques at APL layer may be improved in any case: (i) when retransmissions are allowed (point to point communication), they are performed the least often, (ii) when higher layer redundancy has been introduced to circumvent the problem (e.g. the so-called MPE-FEC of the MAC layer in DVB-H), our strategy at least allows to reduce it, and finally (iii) when no retransmission is allowed, it improves the quality of the video, because error concealment is used less frequently.

Two main ideas may be put at work. First, intra-layer and inter-layer redundancy is present in the protocol stack of Figure 1, and in many other protocol stacks. This redundancy has been exploited in the *Robust Header Compression (ROHC)* mechanism [28] which replaces the headers introduced by the RTP, UDP, and IP layers by a compressed version. Here, this redundancy is used to build some *a priori* information on erroneous headers, which will facilitate their estimation. Second, CRC and checksums are used as error-correcting codes. This technique has been previously proposed in [142] and in [125].

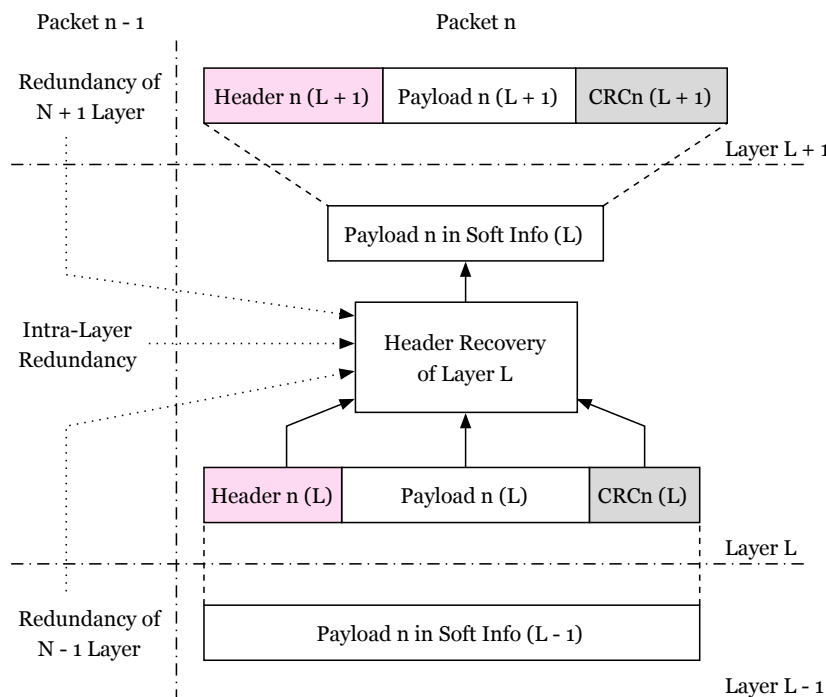


Figure 9: Proposed permeable layer mechanism

Header recovery combines soft information provided by the lower protocol layers, properties of the CRC or checksum, as well as *a priori* information obtained by a careful examination of the protocol. Figure 9 illustrates the way soft information is transmitted from layer $L-1$ to layer $L+1$ through the permeable layer L . This figure also summarizes the various sources of redundancy used to facilitate

the header recovery at layer L . Albeit the header is removed after decoding, its information fields are necessary to help in the delivery of the payload to layer $L + 1$, for further processing.

Emphasis will put first on the PHY and MAC layers of WiFi, as examples of a very generic situation, but the permeable layer mechanisms may be applied to many protocol layers.

In this context, ARQ mechanisms have to be modified to account for the fact that erroneous packets may reach upper layers to be protected. *Soft* ARQ mechanisms have to be defined.

The data encryption at a given protocol layer may also be exploited in a useful way to allow error detection/correction of headers of the upper layers. The interaction between JSCD and data decryption has also to be examined.

2.2.2 *Optimal JSCD algorithms with reduced complexity*

Optimal or near-optimal decoder may be obtained by combining VLC grouping tools as proposed by [131] and trellis state aggregation tools presented in [88].

Optimal JSCD algorithms with reduced complexity may be especially useful when considering JSCD techniques to allow packets to reach upper protocol layers. Information contained in headers at layer L have to be correctly estimated at that layer using soft information provided by layer $L - 1$. Payloads at layer L will be processed at layer $L + 1$ and are not significant at for layer L . When a channel code (CRC or checksum) protects both header and payload at layer L , it may be useful to resort to marginalisation to focus on the header. Efficient techniques for decoding these header and using the redundancy of the CRC may be developed using tools inspired from [217].

2.2.3 *Cross-layer optimization*

In the process of globally optimizing the efficiency of the multimedia transmission, some redundancy is introduced at many layers of the protocol. The best allocation of this redundancy, within the context of JSCC/D, has not been found in the literature, thus it will be addressed here.

Traditionnally, protocol stacks are designed in such a way that the channel codes at PHY layer provide approximatively error-free packets to upper layers. The intermediate layers then check whether packets still contain errors and discards those which are not error-free. The lower layers of the protocol stack thus transform a channel introducing noise at a bit or symbol level into a packet-erasure channel. Packet-erasure correcting codes may then be used at upper layers, see *e.g.*, MPE-FEC in DVB-H [53].

Clearly, this design is fairly compatible with the needs of JSCD techniques, which were first designed to work at APL using soft information on bits. The use of JSCD techniques at lower protocol layers may help to feed the APL layer with soft information available at PHY layer. It is thus a preliminary unavoidable to consider a cross-layer optimization accounting for the presence of JSCD techniques at APL layer.

Once this first step is completed, several other problems have to be investigated. The efficiency of the redundancy introduced at several protocol layers has to be investigated. This problems is partly investigated in WP R7.4. The study of distance spectra of channel *and* source codes (especially when they show some error-correcting capabilities) is an interesting tool for that purpose. EXIT charts [187] is now a widely used tool to evaluate the performance of concatenated codes. It has been used to evaluate the performance of concatenation of source and channel codes [87]. These tools allow a joint optimization of source and channel coding. Nevertheless, usually source and channel codes are situated in very distant layers of the protocol stack, which does not facilitate analysis. Good models of the efficiency of JSCD techniques at lower layers may thus be useful. Among the other problems which have to be focused on, one may need to consider adaptive link layer techniques, intelligent packet scheduling, video distortion modeling...

3 TASK TR7.2 NEW TOOLS FOR JSCC/D

This section summarizes the state of the art and the main technical challenges related to novel tools for JSCC/D.

Digital fountain codes represent an innovation in the field of coding for multimedia applications, due to their adaptability to time varying network condition, and limited computational complexity at both the encoder and the decoder side. The principles of digital fountain codes are recalled in sect. 3.1.1. Unequal error protection is a tool that fits the scalable nature of modern multimedia encoders, such as H.264/SVC. In fact, the hierarchy among layers makes it necessary to allocate different protection levels to each of them. UEP schemes that have been proposed in the scientific literature, for binary erasure as well as other channel models, are described in sect. 3.1.2. Multiple description coding methods are briefly described in sect. 3.1.3. The state of the art related to JSCC for fading channel without channel state information is described in sect. 3.1.4. Finally, in sect. 3.2 the main open issues related to innovative JSCC/D tools, and which are likely to be addressed in the technical community in the near future, are discussed and commented.

3.1 State of the art

3.1.1 JSCC with fountain codes

Raptor codes [177], a class of fountain codes [33] for packet erasure recovery, have been recently selected for Multimedia Broadcast Multicast Service (MBMS) download delivery [2]. Such codes show appealing properties, among which linear time encoding, near-optimum erasure recovery capability under maximum likelihood (ML) decoding, and efficiency with a large range of file sizes. Similar properties have been recently obtained also using judiciously designed low-density parity-check codes [146, 147]. Thanks to these attractive features, Raptor codes represent excellent candidate for reliable file delivery over mobile broadcast networks.

Fixed-rate Raptor codes

Raptor codes were introduced by Shokrollahi in [177]. They are an instance of the concept of *fountain code*, which is commonly used to refer to a code which can produce on-the-fly any desired number of encoded symbols from k information symbols. [33] and, thanks to the large degrees of freedom in parameter choice, they can be applied to several systems, increasing their reliability. Recently, a fully specified version of Raptor codes has been approved as a means to efficiently disseminate data over a broadcast network [2, Annex B]. A (n, k) fixed-rate Raptor code can be obtained by limiting to n the amount of symbols produced by the Raptor encoder. Fixed-rate Raptor codes derived from the MBMS standard [2, Annex B] are currently under investigation for the multi-protocol encapsulation (MPE) level protection within the DVB standards family [48]. In the following, we will provide first a description of the Raptor codes specified in [2, Annex B], including some insights on their encoding and decoding algorithms.

The Raptor code can be viewed as the concatenation of several codes. For example the Raptor encoder specified in [2] is depicted in Fig. 10. The most-inner code is a non systematic Luby-transform (LT) code [120] with L input symbols \mathbf{F} , producing the encoded symbols \mathbf{E} . The symbols \mathbf{F} are known as *intermediate symbols*, and are generated through a pre-coding, made up of some outer high-rate block coding, effected on the k symbols \mathbf{D} . The s intermediate symbols \mathbf{D}_s are known as *LDPC symbols*, while the h intermediate symbols \mathbf{D}_h are known as *half symbols*. The combination of pre-code and LT code produces a non systematic Raptor code. The parameters s and h are functions of k , according to [2]. Some pre-processing is to be put before the non-systematic Raptor encoding to obtain a systematic one. Such a pre-processing consists in a rate-1 linear code generating the k symbols \mathbf{D} from the k information symbols \mathbf{C} . LT codes are the first practical implementation of fountain codes. An unique encoded symbol ID (ESI) is assigned to each encoded symbol. Starting from an ESI i , the encoded symbol E_i is computed by xor-ing a subset Θ_i of d_i intermediate symbols. The number d_i , known as the *degree* associated with the encoded symbol E_i , is a random integer between 0 and $L - 1$: the d_i intermediate symbols are chosen

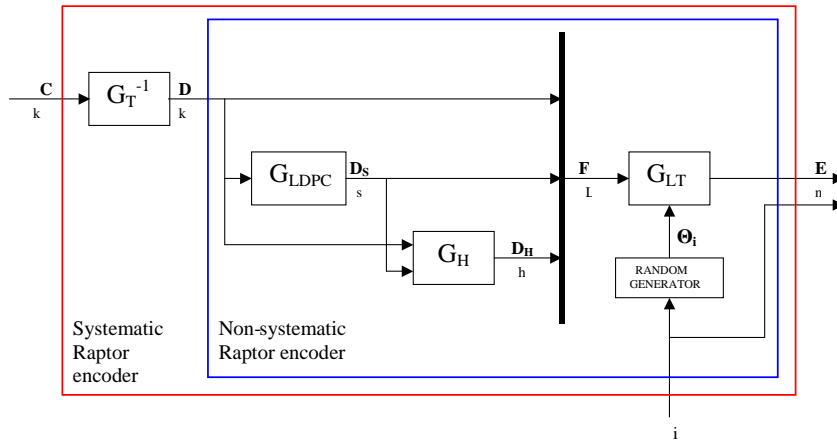


Figure 10: Block diagram of the systematic Raptor encoder specified in [2].

at random according to a specific probability distribution. As a consequence, in order to recover the information symbols the decoder needs both the set of encoded symbols E_i and of the corresponding Θ_i . This last information can either be explicitly transmitted or obtained by the decoder through the same pseudo-random generator used for the encoding, starting from ESIs, which have therefore to be sent together with the corresponding encoded symbols.

Some of the main properties of LT codes are that the encoder can generate as many encoded symbols as desired and that the decoder is able to recover the block of source symbols from any set of received encoded symbols, whose number is only slightly greater than that of the source symbols (in fact the code claims a low level of overhead). A Raptor code, whose core consists of an LT code, inherit such properties. The addition of a pre-coding phase is used to obtain an encoding/decoding complexity linear with k ; a feature which is missing in the mere LT code.

Fixed-rate Raptor generator matrix

Considering a systematic Raptor code as a finite length (n, k) linear block code (fixed-rate Raptor code), we can ask what is the structure of its generator matrix. This problem is addressed next for the Raptor code specified in [2]

The generator matrix of the first pre-coding stage is given by $[\mathbf{I}_k | \mathbf{G}_{LDPC}^T]^T$. According to the specifications in [2], \mathbf{G}_{LDPC} consists in columns all of weight equal to 3, regardless the value of k . On the other hand, the generator matrix of the second pre-coding stage is given by $[\mathbf{I}_{s+k} | \mathbf{G}_H^T]^T$, where \mathbf{G}_H is a $(h \times (s+k))$ matrix consisting in columns all of constant weight: each column is an element of the Grey sequence of weight h' , where $h' = \lceil h/2 \rceil$. Finally, let us denote by \mathbf{G}_{LT} the $(n \times L)$ LT code generator matrix (regarded as a finite length n linear block code). It is built in such a way that the row of index i has d_i ones in Θ_i positions, where d_i and Θ_i are derived from the ESI i , through pseudo-random algorithms described in in [2]. Next, we use the notation $\mathbf{G}_{LT}(i_1, i_2, \dots, i_r)$ to denote the $(r \times L)$ submatrix of \mathbf{G}_{LT} composed of the rows with indexes (i_1, i_2, \dots, i_r) . The notation \mathbf{G}_{LT} is equivalent to $\mathbf{G}_{LT}(1, \dots, n)$.

The $L = k + s + h$ intermediate symbols \mathbf{F} are obtained from \mathbf{D} as

$$\mathbf{F} = \begin{bmatrix} \mathbf{D} \\ \mathbf{D}_s \\ \mathbf{D}_h \end{bmatrix}$$

through the relations

$$\mathbf{D}_s = \mathbf{G}_{LDPC} \cdot \mathbf{D} \quad (7)$$

$$\mathbf{D}_h = \mathbf{G}_H \cdot \begin{bmatrix} \mathbf{D} \\ \mathbf{D}_s \end{bmatrix}. \quad (8)$$

The intermediate symbols \mathbf{F} are the inputs to the LT encoder for deriving the n encoded symbols \mathbf{E} as

$$\mathbf{E} = \mathbf{G}_{\text{LT}} \cdot \mathbf{F}. \quad (9)$$

Let us subdivide \mathbf{G}_{LT} as

$$\mathbf{G}_{\text{LT}} = [\mathbf{G}_{\text{LT}}^{\text{I}} \mid \mathbf{G}_{\text{LT}}^{\text{II}} \mid \mathbf{G}_{\text{LT}}^{\text{III}}],$$

where the sizes of the three submatrices are $(n \times k)$, $(n \times s)$ and $(n \times h)$, respectively. If also \mathbf{G}_{H} is subdivided as

$$\mathbf{G}_{\text{H}} = [\mathbf{G}_{\text{H}}^{\text{I}} \mid \mathbf{G}_{\text{H}}^{\text{II}}],$$

that is into two submatrices whose sizes are $(h \times k)$ and $(h \times s)$, respectively, then the non-systematic Raptor code generator matrix can be expressed as

$$\begin{aligned} \mathbf{G}_{\text{R},n\text{-sys}} &= \mathbf{G}_{\text{LT}}^{\text{I}} + \mathbf{G}_{\text{LT}}^{\text{II}} \cdot \mathbf{G}_{\text{LDPC}} \\ &\quad + \mathbf{G}_{\text{LT}}^{\text{III}} (\mathbf{G}_{\text{H}}^{\text{I}} + \mathbf{G}_{\text{H}}^{\text{II}} \cdot \mathbf{G}_{\text{LDPC}}) \end{aligned}$$

which satisfies the relation:

$$\mathbf{E} = \mathbf{G}_{\text{R},n\text{-sys}} \cdot \mathbf{D}.$$

Let's now subdivide $\mathbf{G}_{\text{R},n\text{-sys}}$ into the two submatrices $\mathbf{G}_{\text{R},n\text{-sys}}^{\text{I}}$ and $\mathbf{G}_{\text{R},n\text{-sys}}^{\text{II}}$, whose sizes are $(k \times k)$ and $((n - k) \times k)$, respectively:

$$\mathbf{G}_{\text{R},n\text{-sys}} = \begin{bmatrix} \mathbf{G}_{\text{R},n\text{-sys}}^{\text{I}} \\ \mathbf{G}_{\text{R},n\text{-sys}}^{\text{II}} \end{bmatrix}.$$

For a systematic code it must be valid the following

$$E_i \equiv C_i \quad \forall i = 1, \dots, k,$$

and therefore

$$\begin{bmatrix} \mathbf{G}_{\text{R},n\text{-sys}}^{\text{I}} \\ \mathbf{G}_{\text{R},n\text{-sys}}^{\text{II}} \end{bmatrix} \cdot \mathbf{D} = \begin{bmatrix} \mathbf{E}_{[1,\dots,k]} \\ \mathbf{E}_{[k+1,\dots,n]} \end{bmatrix} \quad (10)$$

$$= \begin{bmatrix} \mathbf{C} \\ \mathbf{E}_{[k+1,\dots,n]} \end{bmatrix}. \quad (11)$$

We have introduced in (10) the notations $\mathbf{E}_{[1,\dots,k]}$ and $\mathbf{E}_{[k+1,\dots,n]}$ to denote the first k and the last $n - k$ encoded symbols, respectively.

We can state that the pre-processing matrix generating \mathbf{D} from \mathbf{C} can be obtained by

$$\mathbf{G}_{\text{T}}^{-1} = (\mathbf{G}_{\text{R},n\text{-sys}}^{\text{I}})^{-1}$$

and, as a consequence, the systematic Raptor code generator matrix is

$$\mathbf{G}_{\text{R},\text{sys}} = \begin{bmatrix} \mathbf{I}_k \\ \mathbf{G}_{\text{R},n\text{-sys}}^{\text{II}} \end{bmatrix} \quad (12)$$

In (12) \mathbf{I}_k denotes the $(k \times k)$ identity matrix. Obviously, $\mathbf{G}_{\text{R},n\text{-sys}}^{\text{I}}$ can be inverted if and only if it has full rank k . By initializing the random generator of inner LT code through the so-called *systematic index* (defined in [2]), this property is fulfilled for all $k = 4, \dots, 8192$.

k	s	h	
\mathbf{G}_{LDPC}	\mathbf{I}_s	\mathbf{Z}	s
\mathbf{G}_H		\mathbf{I}_h	h
$\mathbf{G}_{\text{LT}}(1, \dots, n)$			n

Figure 11: Structure of the encoding matrix \mathbf{A} for an (n, k) Raptor code specified in [2] ($L = k + s + h$).

Raptor Encoding

The relations (7), (8) and (9) can conveniently be represented as:

$$\mathbf{A} \cdot \mathbf{F} = \begin{bmatrix} \mathbf{0} \\ \mathbf{E}_{[1, \dots, n]} \end{bmatrix}$$

whereby \mathbf{A} is a $((s + h + n) \times (s + h + k))$ binary matrix called *encoding matrix*, whose structure is shown in Fig. 11. In this figure, \mathbf{I}_s is the $(s \times s)$ identity matrix, \mathbf{I}_h is the $(h \times h)$ identity matrix and \mathbf{Z} is the $(s \times h)$ all-zero matrix. The matrix \mathbf{A} doesn't properly represent the Raptor code generator matrix (which is defined in (12) instead), but includes the set of constraints imposed by the pre-coding and LT coding together. We use next the notation $\mathbf{A}(i_1, i_2, \dots, i_r)$ to indicate the $((s + h + r) \times L)$ submatrix of \mathbf{A} obtained by selecting only the rows of \mathbf{G}_{LT} with indexes (i_1, i_2, \dots, i_r) . Again, \mathbf{A} is equivalent to $\mathbf{A}(1, \dots, n)$.

A possible Raptor encoding algorithm exploits a submatrix of \mathbf{A} . Such a matrix, consisting of the first L rows of \mathbf{A} , is used to obtain \mathbf{F} solving the system of linear equations:

$$\mathbf{A}(1, \dots, k) \cdot \mathbf{F} = \begin{bmatrix} \mathbf{0} \\ \mathbf{C} \end{bmatrix}.$$

At this point it is sufficient to multiply \mathbf{F} by the LT generator matrix to produce the encoded symbols \mathbf{E} , according to (9).

Raptor Decoding

The most direct way to decode the received sequence lies in inverting each encoding step of Fig. 10; in this case you work on individual sub-codes. When using ML decoding at each sub-code, such a method requires the inversion of a matrix for each code, so it doesn't appear to be the best solution from the computational viewpoint [121]. Moreover, if the number of received encoded symbols is not larger enough than (which in many cases may mean much higher than) the number of source symbol k , it shows an high failure probability.

For example, let's assume that only a subset of encoded symbols of ESIs (i_1, i_2, \dots, i_r) are available at the decoder. The first step the decoder should perform is to solve the system of linear equations:

$$\mathbf{G}_{\text{LT}}(i_1, i_2, \dots, i_r) \cdot \mathbf{F} = \mathbf{E}_{[i_1, i_2, \dots, i_r]}.$$

The matrix $\mathbf{G}_{\text{LT}}(i_1, i_2, \dots, i_r)$ has $(r \times L)$ size and, obviously, the necessary condition to solve the system is that $r \geq L$. If such a condition is not fulfilled, the decoding fails. It means that to recover the source symbols the decoder requires at least L encoded symbols (let's recall that $L = k + s + h$).

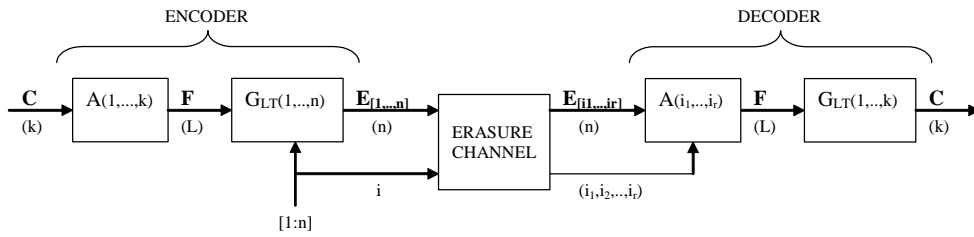


Figure 12: Overview of the encoding and decoding process for the systematic Raptor code specified in [2].

Such a method doesn't exploit the fact that the L intermediate symbols are not independent from each other, but subject to the pre-coding constraints, instead. Therefore, to obtain the intermediate symbols \mathbf{F} by using a submatrix of \mathbf{A} (which consider such constraints) turns out to be a far more efficient solution.

According to the above-mentioned assumption, the first decoding step will turn into:

$$\mathbf{A}(i_1, i_2, \dots, i_r) \cdot \mathbf{F} = \begin{bmatrix} \mathbf{0} \\ \mathbf{E}_{[i_1, i_2, \dots, i_r]} \end{bmatrix}$$

where $\mathbf{A}(i_1, i_2, \dots, i_r)$ is a $((s+h+r) \times L)$ matrix, as defined above. The system can be solved by Gaussian elimination (ML decoding) if and only if $s+h+r \geq L$, that is $r \geq k$. In this way the number of encoded symbols required at the decoder is definitely lower compared to that in the previous case and, notably, is close to the number of source symbols k . Once \mathbf{F} is known, the source symbols \mathbf{C} are easily recovered by

$$\mathbf{C} = \mathbf{G}_{\text{LT}}(1, \dots, k) \cdot \mathbf{F}.$$

To sum up, when the described encoding and decoding algorithms are employed, both the encoding and the decoding are performed by making use of operations which are analogous in the two case (Fig. 12).

If we take into consideration the first decoding step, an algorithm to perform Gaussian elimination in a more efficient way on $\mathbf{A}(i_1, i_2, \dots, i_r)$ has been proposed in [annex E]. This algorithm share some similarities with that proposed in [30] for LDPC codes. In both cases, the erased symbols are solved by mean of a structured gaussian elimination (GE), exploiting the sparse nature of the equations to reduce the size of the matrix on which brute-force GE is performed. The targets of the structured GE are $\mathbf{H}_{\bar{K}}$ for low-density parity-check (LDPC) codes and \mathbf{A} for Raptor codes. Consider now a (n, k) LDPC code and its fixed-rate Raptor counterpart. Suppose also an erasure pattern (introduced by the communication channel) leading to a small overhead δ , i.e., that the amount of correctly received symbols is $k + \delta$. On the LDPC code side, the structured GE will be performed on $\mathbf{H}_{\bar{K}}$ with size $(n-k) \times (n-k-\delta)$. For the Raptor code, the structured GE will work on \mathbf{A} with size $(k+\delta+s+h) \times (k+s+h)$. Hence, while for the LDPC code the complexity of the ML decoder is driven by $(n-k)$ (i.e., the amount of redundancy, thus from the code rate R), for the Raptor code the complexity depends just on k (i.e., it's code rate independent). The result is that for code rates greater than $1/2$ LDPC codes have an inherent advantage in complexity, while for lower code rates Raptor codes shall be preferable.

3.1.2 JSCC using UEP

Modern multimedia communications encompass novel scenarios in which contents can be consumed using different terminals and access networks. For example, users may want to stream their own video contents to a virtual community, using terminals equipped with different capabilities in terms of power consumption, memory, computational resources and visual resolution, including hand-held terminals or cell phones. The contents may be accessed using broadband networks such as DSLs, optical cable or WiMax, but also full mobility GPRS/UMTS or beyond 3G networks. In home distribution using WiFi connection is a today's reality.

In this situation, scalability is a key feature. It allows the video quality to be matched with the capabilities of the terminal at hand, and enables graceful degradation capabilities in case of unreliable transmission. Of course, the hierarchical data partitioning of a scalable encoder demands for unequal error protection strategies.

Let us focus on a scenario where data are transmitted on unreliable, bandwidth limited, non prioritized networks, with all packets encompassing the same loss probability. We assume that packets may be lost due to congestion or errors detected at the physical layer; we do not consider the situation where packets affected by bit errors are delivered to the upper layers. In other words, we employ a binary erasure channel (BEC) transmission model where the basic transmission units are packets. We assume that retransmission strategies cannot be addressed to ensure reliable transmission, due to delay constraints of real time video streaming, and possibly network flooding problems typical of broadcast or multicast applications. In this context, it is important that all the received packets be exploited at the application layer; on the other hand, if packets built on top of a layered or progressive source, the loss of a base layer packet prevents one from exploiting the subsequent ones, even though correctly received. At present, unequal error protection (UEP) is implemented using forward error correction (FEC) codes at the application layer. The most popular UEP methods for BEC transmission model (also known as unequal loss protection - ULP - schemes) make use of Reed-Solomon (RS) codes [129, 133, 134]. The basic concept is to arrange fragments of progressively encoded data in the rows of a matrix, allocating different RS codes to each fragment, and transmitting the columns of such data matrix. The rates of the codes to be employed are obtained as the result of an optimization procedure, aiming at maximizing the expected peak signal-to-noise ratio (PSNR) at the receiver side, given the packet loss probability and the source RD characteristics. As the code allocation procedure is very demanding, alternative, suboptimal yet efficient algorithms have been proposed (see for example [154]). Recently, UEP systems have been designed using LDPC and turbo codes, for the situation of transmission over binary symmetric (BSC) and fading channels. For a thorough review of such methods, the interested reader can refer to [75].

A recent topic is the use of rateless codes, as those described in Sect. 3.1.1, to enable unequal loss protection capabilities. In fact, such codes do not assume any knowledge of the channel; therefore, they are suitable for lossy multicast channels, non uniform or time varying channels, BEC channels with unknown or varying erasure probability. A design of rateless codes with UEP capabilities is proposed in [155]. These codes enjoy the property that some portions of the data turn out to be natively more protected than others. This feature is achieved imposing that not all the input nodes have the same probability of being selected in forming each output node. The variable nodes are partitioned into subsets, and the probability of selecting a node is uniform within each subset, but different from set to set. To increase the recovery probability of nodes in a given set, one has to increase the selection probability of the nodes in that set. The performance of such generalized rateless codes have been analyzed asymptotically, showing that they can provide very low error rates for the most important symbols at the expenses of a limited performance loss as for the least important ones. Finite length codes have also been studied, and bounds to the error rates of such codes under ML decoding have been worked out. Similarly, in [49], a short block length fountain code with built-in UEP characteristics is designed for video multicast; it is based on defining different classes of importance of the input symbols, and different probabilities of being randomly selected to packets belonging to each class.

Other proposals [107, 182, 209] apply different digital fountains to each data layer. The overhead of each digital fountain is tuned to as to match the required probability of success in decoding that layer, as a function of the maximum tolerable delay imposed by the application. Unequal error protection can also be achieved by means of sliding window mechanisms, extending the principle proposed in [27]. For example, in [168], the input data block is ordered so as to be progressive (the first symbols are the most important ones). Then, the block is divided into embedded windows, so as the number of windows a symbol belongs to is proportional to its relevance. Digital fountains are successively enabled on such windows, so leading to a large probability of decoding the most important part of the stream.

3.1.3 JSCC using MDC

Multiple description coding represents a possible alternative to UEP for the protection of multimedia data. In MDC [64], non hierarchical representations of the source are generated, yielding mutually refinable information. The quality at the receiver side only depends on the number of successfully received packets, and not on the particular subset or arrival order. Such nice features are paid in terms of some impairment of the compression efficiency, generally measured in terms of the extra rate required by the MDC scheme, compared to a single description reference system achieving the same performance. However, as redundancy is unavoidable in every resilient scheme, including those based on FEC, a more sensible evaluation is in terms of the resiliency/redundancy trade off. The performance of MDC, as well as UEP, depends on the effectiveness of the policies to trade redundancy for robustness. It is generally recognized that MDC may be competitive with FEC based techniques, especially when the probability of packet loss is relevant [112].

Many MDC methods proposed in the scientific literature address the situation of two balanced descriptions, so enabling two quality levels at the receiver side (*central* and *side* qualities). Such a class encompasses MD scalar quantization (MDSQ) [201] and its numerous variants [202, 203]. MDSQ has been used for image coding in [86, 170]. However, it requires the modification of the quantization stage of standards; moreover, it can hardly be generalized to $N > 2$ descriptions [9, 22]. Pairwise correlating transforms have been proposed, operating on the coefficients in order to introduce redundancy among the descriptions [50, 63, 210]. Again, for more than two descriptions, grouping coefficients and allocating redundancy is not trivial [161]. In the perspective of generating an arbitrary number of balanced descriptions and controlling the trade off between central and side qualities, it is worth mentioning methods based on the polyphase decomposition followed by selective quantization [91].

An interesting topic is the generation of MDC streams that are indeed compatible with standard image and video codecs. Some recently proposed methods make use of H.264/AVC options such as redundant slices, data partitioning, flexible macroblock ordering [192, 206] to create two H.264 streams, which can be either separately decoded, or pre-processed prior to be input to a standard H.264 decoder. These methods yield good performance when 2 descriptions are generated, but may suffer from suboptimal redundancy allocation for a larger number of descriptions. In [180], the authors employ a uniformly distributed random mask for each subband of a wavelet transformed image, so as to generate different groups. These are then quantized at different rates and combined together to generate the descriptions, which are compressed with EBCOT [185]. Although EBCOT is the main engine of JPEG 2000, the implementation of this algorithm for JPEG 2000 requires the modification of the quantization stage, resulting in non-compliant streams. Moreover, the efficiency of EBCOT is impaired due to the perturbation of the coefficient statistics. In [193], a method is proposed, based on the generation of two JPEG2000 streams encoded at different rates, and on the splitting of codeblocks (CBs) encoded at either rate into two descriptions. The method is generalized to H.264/AVC video in [164], where intermediate, unbalanced bit streams are split into two descriptions.

In some situations, generating more than two descriptions may be useful. Generally speaking, the most suitable number of descriptions to be generated should be input to an optimization procedure, which takes into account the data rate/distortion (RD) characteristics, the available bandwidth, the packetization strategy, and the probability of packet loss. The problem of generating more than 2 descriptions is addressed in [10, 11] for JPEG2000 data. However, both algorithms suffer of suboptimal redundancy allocation. A first attempt to overcome this limitation is provided in [190], where the concept of using multiple rates is introduced, and some simulation results are reported for a Gaussian source. The comparison of MDC schemes with many descriptions and UEP methods represents an interesting open issue.

3.1.4 JSCC for fading channel without channel state info

Fading channel occurs naturally as a model in wireless communications. For slow fading, the receiver can usually recover the channel state information (CSI) accurately, however the transmitter only knows

the probability distribution of CSI, but not the realization. Such uncertainty can cause significant system performance degradation, and the broadcast strategy was used in [172], [173] as an approach to combat this detrimental effect. In this strategy, some information can only be decoded when the fading is less severe, which is superimposed on the information that can be decoded under more severe fading. Thus the receiver can decode the information adaptively according to the realization of the channel state. The similarity to the degraded broadcast channel [39] (see also [43]) is clear in this context, particularly for channels with a finite number of fading states. Generalizing this view, when the fading gain can take continuous values, the receiver can be taken as a continuum of users in a broadcast channel. The broadcast strategy naturally matches the successive refinement (SR) source coding framework [102]- [160], as the information decodable under the most severe fading is protected the most, and should be used to convey the base layer information in the SR source coding. As more information can be decoded when the channel is subject to less fading, more SR encoded layers can be decoded, and the reconstruction quality improves. In this work, we consider this scheme for a quadratic Gaussian source on a single input single output (SISO) channel. In order to minimize the expected distortion at the receiver, it is essential to find the optimal power allocation in the broadcast strategy, and this is indeed our focus. It is worth noting here that though in [173] the objective function to be maximized is the expected rate, the cross layer design approach of combining SR source coding with broadcast strategy was in fact suggested (though not treated) in that work.

Initial effort on this problem was made by Sesia et al. in [171], where the broadcast strategy coupled with SR source coding was compared with several other schemes. The optimization problem was formulated by discretizing the continuous fading states, and an algorithm was devised when the source coding layers are assumed to have the same rate. This algorithm, however, does not directly yield the optimal power allocation when the fading states are discrete and pre-specified, nor does it give a closed-form solution for the continuous case. Etemadi and Jafarkhani also considered this problem in [52], and provided an iterative algorithm by separating the optimization problem into two sub-problems. In two more recent works [138], [1], Ng et al. provided a recursive algorithm to compute the optimal power allocation for the case with M possible fading states, with worst case complexity of $O(2^M)$; moreover, by directly taking the limit of the optimal solution for the discrete case, a solution was given for the continuous case optimal power allocation, under the assumption that the optimal power allocation is concentrated in a single interval. Similar problems were considered in [36]- [25] in the high SNR regime from the perspective of distortion exponent.

3.2 Perspectives

3.2.1 Comparison between RS and DFC

The main pros and cons of RS and DF codes are well known. RS codes optimized for erasure channels are optimal from a theoretical point of view, i.e. they require the minimum possible amount of encoded symbols to recover the whole data block. RS have been used for many years in several applications. They are non proprietary, and co-decoding algorithms are freely available. On the other hand, when using RS codes, one has to fix the coding rate in advance; therefore, RS are not adaptive with respect to varying network conditions. RS codes operate over Galois fields, and this limits the flexibility in defining symbol and block sizes. Consequently, their performance is affected on the symbol and block dimensions, as padding may be necessary to match the actual code constraints. Last but not least, RS impose rather demanding processing at both the decoder and the encoder side, often requiring hardware implementation.

DF codes are rateless, so they can be adapted to varying network conditions. However, it should be noted that the delay in the decoding process is indeed dependent on the erasure probability, and it should be kept into control if such codes are applied to delay sensitive data such as real time video. When using DF codes, a code overhead must be accounted for, depending on the desired probability of decoding success. Some efficient DF codes (such as raptor codes) are commercial. On the other hand, DF codes do

not impose any limitation as for the symbol and block sizes. The co-decoding complexity of raptor codes is linear with the block size, hence significantly less than RS.

From this brief discussion, it turns clear that the relative advantages of each class of codes depends on many application details:

- Actual need of flexibility (to what extent are the network parameters time varying?).
- Available computational resources.
- Delay requirements.
- Block and symbol sizes.
- Required probability of correct decoding (related to the necessary coding overhead).

For all these reasons, we can conclude that a thorough comparison of these two code families is not available at present, for sensible case studies and applications.

3.2.2 JSCC for fading channel

Our perspective is to focus on two central aspects of the problem: firstly, devising a new algorithm that can compute in linear time, i.e., of $O(M)$ complexity, the optimal power allocation for the case with M fading states; secondly, we aim to derive for the continuous case optimal power allocation solution by the classical variational method [61]. Such a result will be more general than that in [1] as it removes the restriction that the optimal power allocation is concentrated in a single interval. Both the algorithm and the derivation rely on an alternative representation of the Gaussian broadcast channel capacity, which appeared in [194]. The dual problem of minimizing power consumption subject to a given expected distortion constraint is also of interest. The broadcast strategy coupled with SR source coding can be considered as a source-channel separation approach with optimized cross-layer resource allocation. More explicit joint source-channel coding approaches, for example those in [35], [130]- [158], may provide better performance in the scenario being considered, however this aspect is beyond the current scope of the work. We nevertheless note that the algorithm which will be developed in the current effort makes it possible to use the broadcast strategy coupled with SR coding as a benchmark for future investigation into this joint source-channel coding problem.

3.2.3 Merging successive refinement and broadcast channel coding

The work [212] looked at a problem similar to the one of Section 3.2.2, with the main distinction that the broadcast nature is retained (i.e. one considers a set of receivers with different fading levels), rather than averaged over fading realizations. The focus in [212] is on code designs for a continuously degraded Gaussian broadcast channel, that is a continuous range of channel SNRs (CSNRs), with the assumption of perfect channel state information at the receiver (CSIR). However, instead of averaging over a fading distribution, it formulated broadcast strategies that have a constant rate loss for every CSNR in a range, compared to a code that maximizes the receiver rate at that CSNR. Thus the goal was to build codes for the memoryless AWGN channel that are universal over a range of CSNRs, where the price for this universality is a rate loss compared to the capacity at a given CSNR.

The approach in [212] was to match a successive refinement source code to a superposition broadcast channel code, thus maintaining source-channel separation in a multi-terminal setting where it does not hold in general. Future work in this direction could consider generalizing the broadcast strategy via a rate loss profile, which will be a function of the operational rate-distortion curve and the desired degradation behavior. Such broadcast strategies can be seen as a first step towards a class of universal codes for a family of Gaussian channels. One interesting quantity to be studied in this regard is the minimax rate loss

$$\rho^* = \lim_{n \rightarrow \infty} \min_{\mathcal{C}_n} \max_{\gamma \in [\gamma_0, \gamma_1]} C(\gamma) - R_{\mathcal{C}_n}(\gamma), \quad (13)$$

where $\{\mathcal{C}_n\}$ is a sequence of random universal codebooks and $C(\gamma)$ is the capacity of a channel with CSNR γ . The superposition strategy of [212] provides an upper bound on ρ^* , respectively an inner bound to the region \mathcal{R} of achievable triples $(\gamma_0, \gamma_1, \rho)$.

If the separation between source and channel coding is not needed, the relevant quantity is the mini-max distortion gap

$$\Delta^* = \lim_{n \rightarrow \infty} \min_{\mathcal{C}_n} \max_{\gamma \in [\gamma_0, \gamma_1]} \frac{D_{\mathcal{C}_n}(\gamma)}{D(C(\gamma))}, \quad (14)$$

where we assumed exponential rate-distortion behavior, so that the distortion ratio is meaningful.

The quantity ρ^* could be useful to benchmark separation-based digital JSC systems, while Δ^* could play the same role for digital JSC without separation.

3.2.4 *Joint source coding and network coding*

In the approaches described above, the data flows are coded at the source only. However, this strategy may be suboptimal for multicast traffic, where many sources send their data to many destinations over a common network. More precisely, a new paradigm has been proposed, where the coding is performed by the intermediate nodes of the network [5]. This approach is referred to as Network coding and allows significant gain with respect to the traditional routing approach, where intermediate nodes can only forward the received packets. In other words, traditional approaches consist in separating the source coding from the transport (routing). No general separation results exist. On the other hand, some counterexamples show that the separation may be suboptimal.

The goal is to study the possible interaction of source and network coding and to propose new coding strategy at the intermediate nodes that take into account the possible correlation between the data flows. Special effort will be devoted to the design of low complexity decoder in a realistic scenario. This realistic scenario will consist into considering deterministic network coding rather than the more theoretical random coding. The effect of additional informations [101] about the data flows (correlation), the knowledge of the local topology at each intermediate node [215] will be studied and algorithms that explicitly exploit this information will be proposed.

3.2.5 *Multiple description with side information at the decoder*

The problem addressed is robust compression of a source when correlated side information is available at the decoder. This problem can be seen as a generalization of the multiple description problem with the inclusion of side information. The rate-distortion region for this problem has been established in [47]. To our knowledge, no practical scheme based on this theoretical problem has been designed yet. Multiple description coding has been introduced as a generalization of source coding subject to a fidelity criterion for communication systems that use diversity to overcome channel impairments. Several correlated coded representations of the signal are created and transmitted on different channels. The design goals are therefore to achieve the best average rate-distortion performance when all the channels work, subject to constraints on the average distortion when only a subset of the channels is received correctly. MDC is an interesting tool for robust communication over lossy networks such as the Internet, peer-to-peer, diversity wireless networks and sensor networks. A resilient peer-to-peer streaming approach is proposed in [145] based on multiple distribution trees introducing diversity in network paths used together with multiple description coding introducing redundancy in the transmitted data. Jointly optimized multi-path routing and multiple description coding is also shown in [12] to improve the end-to-end quality of service in dense mesh networks. The above approaches can be regarded as joint source-network coding techniques. The objective is to go one step further and consider the case where correlated side information about the transmitted source is available at the receiver. Multiple description coding with side information at the receiver is particularly relevant for robust transmission in sensor networks where correlated data are being transmitted to a common receiver, as well as for robust video compression. The rate-distortion region for multiple description source coding when side information about a correlated random process

is only known at the decoder has been established in [47]. Analytical expressions of the rate-distortion bounds are derived for Gaussian sources and a Gaussian correlation model, assuming the side information to be common to the two descriptions. We will focus on the design of a practical multiple description coding scheme with side information at the receiver. It will build upon both multiple description coding principles and Slepian-Wolf coding principles.

4 TASK TR7.3 TOOLS FOR MULTI-TERMINAL JSCC/D

4.1 State of the art

4.1.1 Distributed source coding: theoretical background

DSC finds its foundation in the seminal Slepian-Wolf (SW) [178] and Wyner-Ziv (WZ) [220] theorems. In this section, the principles of Slepian-Wolf and Wyner-Ziv coding as well as the corresponding rate bounds are first reviewed. For more details, the readers are referred to [62, 152, 221].

4.1.1.1 Lossless coding of correlated sources: Slepian-Wolf coding

Let X and Y be two binary correlated memoryless sources to be losslessly encoded. If the two coders communicate, it is well known from Shannon's theory that the minimum lossless rate for encoding X and Y is given by the joint entropy $H(X, Y)$. Slepian and Wolf have established in 1973 [178] that this lossless compression rate bound can be approached with a vanishing error probability for long sequences, even if the two sources are coded separately, provided that they are decoded jointly and that their correlation is known to both the encoder and the decoder. The achievable rate region is thus defined by $R_X \geq H(X|Y)$, $R_Y \geq H(Y|X)$ and $R_X + R_Y \geq H(X, Y)$, where $H(X|Y)$ and $H(Y|X)$ denote the conditional entropies between the two sources. This region is shown in figure 13. Let us consider the particular case where Y is available at the decoder, and has been coded without information on the source X at its entropy rate $H(Y)$. According to the SW theorem, the source X can be coded losslessly at a rate arbitrarily close to the conditional entropy $H(X|Y)$ (which is function of the innovation of X given Y), if the sequence length tends to infinity. The minimum total rate for the two sources is thus $H(Y) + H(X|Y) = H(X, Y)$. This set-up where one source is transmitted at full rate (e.g., $R_Y = H(Y)$) and used as side information (SI) to decode the other one (implying $R_X = H(X|Y)$ or reciprocally) corresponds to one of the corner points of the SW rate region (see [152]). This operating point is called the asymmetric case and corresponds to the point A (and reversely point B) in figure 13 (b).

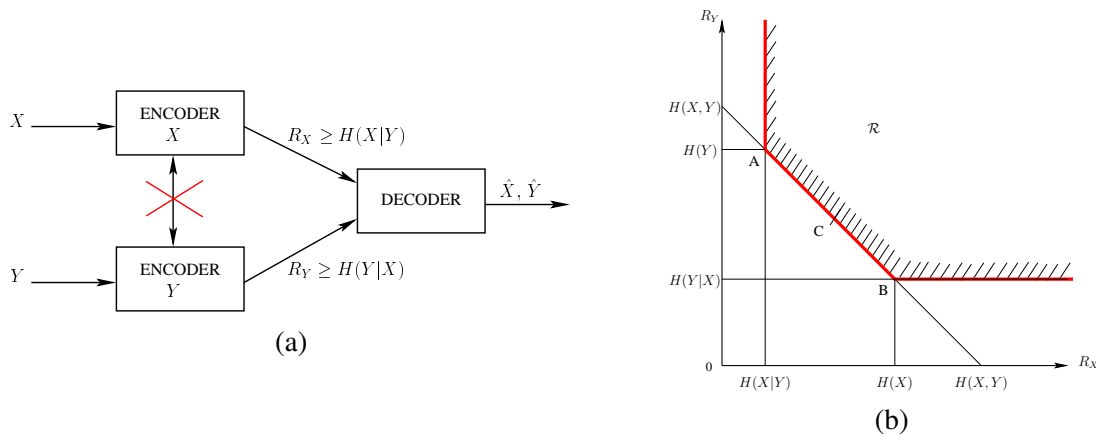


Figure 13: Distributed source coding of statistically dependent i.i.d. discrete random sequences X and Y . Set-up (left); achievable rate region (right).

SW code construction: the syndrome approach. Due to the random code generation in the SW theorem, the SW proof is non-constructive. However, some details in the proof do give insights on how to construct SW bound achieving codes. More precisely, the proof relies on binning, i.e. a partition of the space of the source sequences into subsets (or bins) such that the vectors in the bins are as far apart (or as much “jointly non-typical”) as possible. In 1974, Wyner suggested the use of parity check codes to approach the corner points of the SW rate region [219]. These corner points (see points A and B in figure 13

(b) correspond to the asymmetric case where one source is transmitted at full rate (*e.g.*, $R_Y = H(Y)$) and used as SI to decode the other one (implying $R_X = H(X|Y)$ or reciprocally).

The bins partitioning the space of all possible source realizations are constructed from the cosets of the parity check code. The correlation between X and the side information Y is modeled as a “virtual” channel, where Y is regarded as a noisy version of X . The compression of X is achieved by transmitting only a bin index, *i.e.*, a syndrome. The decoder corrects the “virtual” channel noise, and thus estimates X given the received syndrome and the SI Y regarded as a noisy version of the original sequence X . More precisely, the decoder computes the codeword closest to Y in the coset indexed by the syndrome of X . This approach called “syndrome approach” is shown to be optimal [219], in the sense that if one can build an optimal code for the virtual channel, then this code will also achieved the corner points of the Slepian-Wolf region (points A and B in figure 13 (b)).

Recent channel capacity-achieving codes, block codes, convolutional codes [153], LDPC codes [118], or turbo-codes [195, 196] have been shown to approach the corner points of the SW region. In all these methods, the encoder computes a syndrome by applying a parity check matrix H . The decoder has to estimate X s.t.

$$\hat{X} = \arg \min_{X:HX=s} d(X,Y)$$

Note that the decoder differ from conventional decoders (used in channel decoding), since the optimization has to be performed in a coset that may differ from the set of codewords (with syndrome $S = 0$).

For block codes [153], syndrome decoding is performed. For convolutional codes, [153] proposes to apply the Viterbi decoding on a modified trellis. If the syndrome is zero the trellis is left unaltered. However, if the syndrome is non zero, then one arbitrary codeword of the coset is added to all the codewords that label the edges of the trellis. Finding such a representative is easy if the code is systematic. Therefore the method work for systematic convolutional only. In [196], the authors propose to first add an arbitrary codeword (belonging to the same coset as X), then decode (in the coset of the syndrome 0), and then retrieve the arbitrary codeword. The first operation is called Inverse Syndrome Former. The advantage of this method is to use a conventional Viterbi decoder. This method can also be applied to turbo codes [195]. For LDPC codes, the belief propagation decoder can be adapted to take into account the syndrome [118]. More precisely, the syndrome bits are added to the graph such that. each syndrome bit is connected to the parity check equation it is related to. The update rule at a check node is modified in order to take into account the value of the syndrome bit (known perfectly at the decoder). The authors in [163] propose a SW scheme based on convolutional codes and turbo codes that can be used for any code (not only systematic convolutional code). Rather than considering the usual trellis based on the generator matrix of the code, the decoder is based on a syndrome trellis.

SW code construction: the parity approach. The above detailed syndrome approach is optimal [219], however it may be difficult to construct rate-adaptive codes by puncturing the syndrome. Therefore another approach called “parity approach” has been proposed. The compression of the source X is achieved by transmitting only the parity bits of the source X . Here again, the correlation between the source X and the SI Y is modeled as a “virtual” channel, where the concatenation of the SI and the parity bits is regarded as a noisy version of the concatenation of the source X and its parity bits. The channel is therefore a parallel combination of a BSC and a perfect channel. The decoder corrects the “virtual” channel noise, and thus estimates X given the parity bits and the SI regarded as a noisy version of X . Therefore, the usual ML decoder must be adapted to take into account that some bits of the received sequence (the parity bits) are perfectly known at the decoder.

4.1.1.2 Lossy coding of a source with side information at the decoder: Wyner-Ziv coding

In 1976, Wyner and Ziv considered the problem of coding, with respect to a fidelity criterion, of two correlated sources X and Y [220]. They have established the rate-distortion function $R_{X|Y}^*(D)$ for the case where the SI Y is perfectly known to the decoder only. For a given target distortion D , $R_{X|Y}^*(D)$ in general

verifies $R_{X|Y}(D) \leq R_{X|Y}^*(D) \leq R_X(D)$, where $R_{X|Y}(D)$ is the rate required to encode X if Y is available to both the encoder and the decoder, and R_X is the minimal rate for encoding X without SI. Wyner and Ziv have shown that, for correlated Gaussian sources and a mean square error distortion measure, there is no rate loss with respect to joint coding and joint decoding of the two sources, i.e., $R_{X|Y}^*(D) = R_{X|Y}(D)$. This no rate loss result has been extended in [151] to the case where only the innovation between X and Y needs to be Gaussian, that is where X and Y can follow any arbitrary distribution.

Coding under a fidelity criterion finds its foundations in quantization theory. Practical code constructions based on the Wyner-Ziv theorem thus naturally rely on a quantizer (source code) followed by a SW coder (channel code). The source is first quantized using a fine source code. The quantizer partitions the continuous-valued source space into 2^{R_s} regions (or quantization cells), where R_s is defined as the source rate in bits/sample. A codeword is associated to each region, thus constructing the source codebook. The SW coder then partitions the source codebook into 2^R cosets, each containing 2^{R_c} (with $R = R_s - R_c$) codewords, and computes the index of the coset containing the source codeword. Only the index I of the coset is transmitted with a transmission rate $R \leq R_s$. By exploiting the remaining correlation between X_q (the quantized version of X) and the SI Y , the SW decoder recovers the source codeword in a given coset by finding the codeword which is the closest to the observed SI Y . The SW decoder is followed by an MMSE estimation which searches for \hat{X} , the reconstructed value of X , minimizing the expectation $E[(X - \hat{X})^2 | \hat{X}_q, Y]$, where X_q is the quantization index of X . A graphical illustration of the WZ coding steps can be found in [153] with the example of scalar quantization. Under ideal Gaussianity

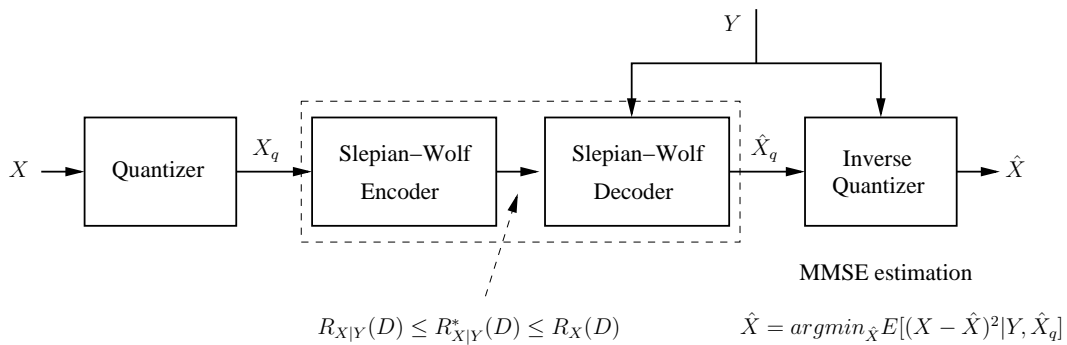


Figure 14: Wyner-Ziv coding principles.

assumptions, the WZ limit can be asymptotically achieved with nested lattice quantizers. Nested lattice constructions are proposed for WZ coding in [169], [211], [115]. The source is first quantized using a fine source code, the coset indexes being then encoded with a SW code which exploits the remaining correlation between the quantized version of X and the SI Y . However, to minimize the quantization loss, the lattice quantizer may require high dimensionality, hence high complexity. Solutions based on trellis coded quantizers [223] are shown to approach the performance of Lattice codes with reduced complexity.

4.1.1.3 Lossy coding of correlated sources: Multiterminal coding.

Gaussian sources Multiterminal (MT) Source Coding is an extension of Slepian-Wolf coding to the case when one allows distortion in the reconstruction of the two sources [21], [198], [143]. As shown in Fig. 15, there are two classes of MT source coding problems: the *direct* and *indirect* MT source coding problems. The problem of *direct* MT source coding consists in finding the achievable rate regions at which two correlated continuous-valued sources Y_1 and Y_2 can be separately compressed so that they can be jointly recovered under the constraints of given distortions D_1 and D_2 . In *indirect* MT source coding, the two sources Y_1 and Y_2 are noisy observations of another source X . The problem is then of finding the rate regions for the noisy observations Y_1 and Y_2 so that one can then estimate the source X under the constraint of a given distortion D . Inner and outer bounds of the rate regions have been derived for

different MT coding problems, and in particular for the Gaussian quadratic case, [21, 143, 144, 150, 198]. Note however that the general multiterminal lossy source coding problem remains unsolved.

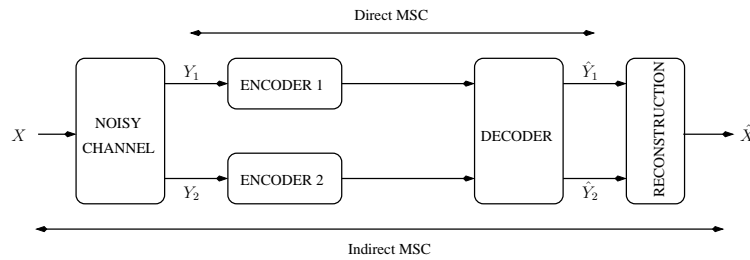


Figure 15: Direct and Indirect Multiterminal Source Coding (MSC).

Let us consider the problem of direct MT coding of two correlated Gaussian sources Y_1 and Y_2 , as depicted in Fig. 15. Let $\sigma_{Y_1}^2$ and $\sigma_{Y_2}^2$ be the variances of Y_1 and Y_2 , respectively. Y_1 and Y_2 are transmitted at rates R_1 and R_2 and reconstructed with an average distortion D_1 and D_2 , respectively. Let $\rho = \frac{\text{cov}(Y_1, Y_2)}{\sqrt{\sigma_{Y_1}^2 \sigma_{Y_2}^2}}$ be the correlation coefficient between Y_1 and Y_2 , where $\text{cov}(Y_1, Y_2)$ is the covariance of these two variables. Considering a mean square error distortion measure, the inner bound for the rate region can be expressed as [143]

$$R_1 \geq \frac{1}{2} \log^+ \left[\frac{\sigma_{Y_1}^2}{D_1} (1 - \rho^2 + \rho^2 2^{-2R_2}) \right] \quad (15)$$

$$R_2 \geq \frac{1}{2} \log^+ \left[\frac{\sigma_{Y_2}^2}{D_2} (1 - \rho^2 + \rho^2 2^{-2R_1}) \right] \quad (16)$$

$$R_1 + R_2 \geq \frac{1}{2} \log^+ \left[(1 - \rho^2) \frac{\beta_{\max}}{2} \frac{\sigma_{Y_1}^2}{D_1} \frac{\sigma_{Y_2}^2}{D_2} \right] \quad (17)$$

where $\log^+ x = \max\{\log x, 0\}$ and $\beta_{\max} = 1 + \sqrt{1 + \frac{4\rho^2}{(1-\rho^2)^2} \frac{\sigma_{Y_1}^2}{D_1} \frac{\sigma_{Y_2}^2}{D_2}}$.

4.1.2 Robust distributed source coding

4.1.2.1 Distributed joint source-channel coding/decoding with side information

Following the idea of joint source-channel decoding, *e.g.*, for variable-length codes, joint source-channel decoding schemes for SW (mainly) and WZ encoded data aim at exploiting at decoder side, the redundancy present in the source or artificially introduced during the encoding process.

In this section, the side information is assumed to be available noiseless at decoder side. Figure 16 illustrates the problem of joint source-channel coding with side information at the decoder. The main idea

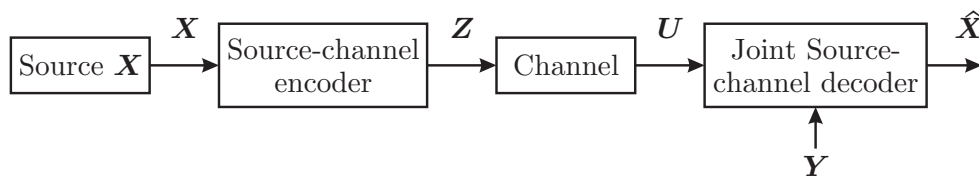


Figure 16: Joint source-channel coding of \mathbf{X} with side information \mathbf{Y} at the decoder

of distributed joint-source coding and decoding rely on the *parity-based* SW encoding, which will be recalled first, for more details, see [16, 183, 222].

Syndrome- and Parity-based SW schemes. In a conventional SW encoding scheme, to compress a sequence $\mathbf{x} \in \mathcal{X}^n$ of n source outcomes, a *syndrome-based encoder* employs an (n, k) channel code \mathcal{C} given by a generator matrix $\mathbf{G}_{k \times n} = [\mathbf{I}_k, \mathbf{P}_{k \times (n-k)}]$ ($\mathbf{G}_{k \times n}$ is assumed systematic for simplicity, but this is not mandatory). The associated parity-check matrix is then $\mathbf{H}_{(n-k) \times n} = [\mathbf{P}_{k \times (n-k)}^T, \mathbf{I}_{n-k}]$. The coder forms then a syndrome $\mathbf{z} = \mathbf{x}\mathbf{H}^T$ and sends it to the decoder. Then the decoder generates an n -length vector $\mathbf{t} = [\mathbf{0}_{1 \times k}, \mathbf{u}]$ from the received vector \mathbf{u} . Since $\mathbf{c} = \mathbf{x} \oplus \mathbf{t}$ is a valid codeword of \mathcal{C}^S , the idea is to decode $\mathbf{t} \oplus \mathbf{y}$ to get $\hat{\mathbf{c}} \in \mathcal{C}$ and to reconstruct the source as $\hat{\mathbf{x}} = \hat{\mathbf{c}} + \mathbf{t}$.

In the parity-based approach, an $(n+r, n)$ linear systematic code \mathcal{C}^P with generator matrix $\mathbf{G}_{k \times n}^P = [\mathbf{I}_n, \mathbf{P}_{n \times r}^P]$ is used. The encoder forms the r -length parity vector $\mathbf{z}^P = \mathbf{x}\mathbf{P}_{n \times r}^P$ and transmits it to the decoder, which gets \mathbf{u}^P . The decoder generates the length $(n+r)$ vector $\mathbf{t}^P = [\mathbf{y}_{1 \times n}, \mathbf{u}^P]$ by appending the side information $\mathbf{y}_{1 \times n}$ to \mathbf{u}^P . The decoding of \mathbf{t}^P in \mathcal{C}^P allows to obtain $\hat{\mathbf{c}}^P = \hat{\mathbf{x}}^P \mathbf{G}_{k \times n}^P$. The estimate $\hat{\mathbf{x}}^P$ is obtained by looking at the systematic part of $\hat{\mathbf{c}}^P$.

If $r = n - k$, both schemes are equivalent in terms of performance on a noiseless channel. The decoding using the syndrome-based encoder is simpler (the code is shorter), thus it is optimal for noiseless channels. Nevertheless, when the channel is noisy, the parity approach gives much more design flexibility. When $r \geq nH(X|Y)$, more data than strictly require are transmitted on the channel. The redundant information may easily be used to correct errors introduced by the channel.

Two-channel representation of the transmission scheme. An equivalent way to look at the joint source-channel decoding problem of data encoded with the parity-based encoder is represented in Figure 17. If

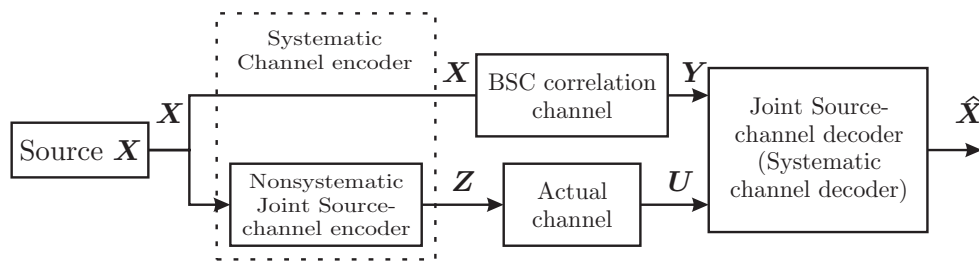


Figure 17: The two equivalent channels for nonsystematic joint source-channel coding of \mathbf{X} with side information \mathbf{Y} at the decoder

a (parity-based) joint source-channel code is used with \mathbf{X} as input and \mathbf{Z} as output, \mathbf{Z} passes through the conventional channel to obtain \mathbf{U} and \mathbf{X} passes through the correlation channel to get the side information \mathbf{Y} . Two parallel channels with different characteristics have thus to be considered by the decoder, see [183].

Joint decoding The channel code may be a turbo-code, as in [4, 117], or a Systematic Irregular Repeat-Accumulate (IRA) code, as in [119]. This latter class of LDPC codes may be optimized to account for the transmission of their systematic and parity parts over two different channels [92]. They are thus well-suited in this context, see Figure 18. Message-passing decoding algorithm may be easily used. Two initialisations of the log likelihood ratios have to be considered, depending on the channel (actual channel and correlation channel). For the IRA code, represented on Figure 18, the side information \mathbf{Y} serves as initialisation of the information nodes and the channel output \mathbf{U} is used to initialize the parity nodes. Classical iterative decoding algorithms may then be used.

In [80], the Slepian-Wolf encoding and the error protection are done separately, exploiting the fact that the optimisation of both codes can be performed in an optimal way separately [175]. Nevertheless, the decoding is performed jointly, in order to avoid error propagation, in the case of uncorrected channel errors. An LDPC code is used for the Slepian-Wolf coding and a turbo code is used to add some redundancy to the compressed data. Soft decoders for both codes are used iteratively in order to get the best decoding performance. To perform the iterative decoding, an *extended* parity-check matrix is used for

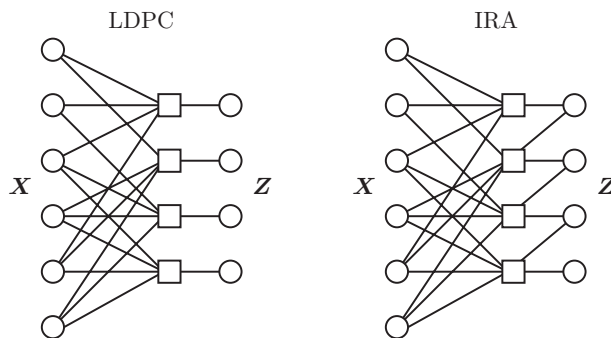


Figure 18: LDPC and IRA codes

the LDPC decoder. Such matrix is obtained by concatenating an identity matrix to the original parity-check matrix of the LDPC code. The extended parity-check matrix allows to get an estimate of the whole LDPC codeword (information and syndrome) from the noisy observation of a codeword. The easier rate allocation to the source and channel code is the main advantage of this scheme.

4.1.2.2 Sending correlated sources over noisy channels: theoretical background

Most of the theoretical results in multi-terminal scenarios are obtained by combining Wyner-Ziv and/or Gelfand-Pinsker coding, where the side information can be used to the non-relevant side resulting in lossless information rates. These coding schemes have been shown to be optimal for the single-user cases of source coding with side information at the decoder and channel coding with non-causal side information at the encoder [128]. However, in its full generality, the information theoretic understanding of multi-terminal scenarios touches upon the basic theoretic models not yet fully understood, involving a number of challenging questions. For example, there are several multi-terminal scenarios of interest for potential sensor networks and relay channels applications, where the separation theorem does not hold or, for which joint source-channel coding leads to simpler coding schemes. In addition to this, optimal encoders and decoders of multi-terminal systems are very sensitive to the knowledge of the time-varying parameters of source and channel statistics. Consequently, most of exciting results in this area may not hold or must be revised if only noisy estimates (perhaps very poor) of these parameters are available.

We now review the main interesting results in multi-terminal systems with separate and joint source-channel code design.

Multiple Access Channels and Sensor Networks. The multiple access channel (MAC) models the case where many independent senders transmit data over a common noisy channel. The MAC can therefore be seen as a generalization of the one sender case, whose properties are studied in the channel coding theorem. This theorem a.k.a. second Shannon's theorem considers the transmission of one sender $X \sim p(x)$ over a noisy channel with output Y and transition probability $p(y|x)$. This theorem shows that the maximal achievable rate (in bits per channel use) of the sender is the capacity of the channel defined as:

$$\max_{p(x)} I(X;Y)$$

where $I(X;Y)$ is the mutual information (MI) between the r.v. X and Y . The mutual information is a function of the input density $p(x)$ and of the channel transition probability $p(y|x)$. It measures the amount of information shared by the 2 r.v. X and Y . The set of achievable rates can therefore be rewritten as:

$$\bigcup_{X \sim p(x)} \{R \text{ s.t. } R \leq I(X;Y)\}$$

The 2-user MAC models the transmission of 2 independent senders $(X_1, X_2) \sim p(x_1)p(x_2)$ over the same channel with output Y and transition probability $p(y|x_1, x_2)$. The transmission rates are denoted R_1

and R_2 . The capacity region is the closure of the set of achievable (R_1, R_2) rate pairs and it is shown to be the closure of the following set:

$$\bigcup_{(X_1, X_2) \sim p(x_1)p(x_2)} \left\{ (R_1, R_2) \text{ s.t. } \begin{array}{l} R_1 \leq I(X_1; Y | X_2) \\ R_2 \leq I(X_2; Y | X_1) \\ R_1 + R_2 \leq I(X_1, X_2; Y) \end{array} \right\}$$

This result can be extended to any number of senders. Closed form expression of the capacity region exists in different examples, for instance the Gaussian MAC [44].

Distributed joint source-channel coding refers to the problem of sending correlated sources over a common noisy channel without communication between the senders. Note that cooperation among channel and source encoding of one sender is allowed but not between different senders. This problem occurs mostly in network, where the communication between the nodes is not possible or not desired due to its high energy cost (network video camera, sensor network...).

A major difference with the joint source-channel case is that the separation between source and channel coding does not always hold. This depends on the setup. If asymmetric source encoding is performed (one source can be recovered perfectly from the data sent by this source only), then separation holds [174]. Otherwise it depends on the channel. For independent channels, source channel separation holds [15] but for interfering channels (as the MAC) counterexamples can be found where joint source-channel scheme (but still distributed) performs better than the separated scheme [42]. Interestingly, the whole rate region is still unknown. More precisely, the converse does not provide any single letter characterization.

Broadcast Channels The Broadcast Channel was first devised and explored by Cover in 1972 [40]. This simply consists of a transmitter communicating information simultaneously to several receivers, where the receivers are not allowed to cooperate. Since then, a primary and fundamental problem has been to find the single-letter characterization of its capacity region in general, which has been extensively investigated. Up to the present, however, conclusive results have been established only for several special cases.

For the case of a broadcast channel with two users, where the sender wishes to transmit in n channel uses an integer $m_1 \in \{1, \dots, \exp(nR_1)\}$ to the receiver 1, an integer $m_2 \in \{1, \dots, \exp(nR_2)\}$ to the receiver 2 and also an integer $m_0 \in \{1, \dots, \exp(nR_0)\}$ to both receivers, where the variables m_1, m_2 and m_0 are assumed to be mutually independent. The largest region of achievable rate triple (R_0, R_1, R_2) for general discrete memoryless broadcast channel presently known is due to Marton [126], but the optimality of this region has not yet been established. In [76], this result has been extended for the case where the encoder wishes to send two correlated and identically distributed sources (S_1, S_2) over a general discrete memoryless broadcast channel. The authors established a new coding theorem for this joint source-channel problem. An apparently reasonable approach for this problem would be to compress the sources (S_1, S_2) into a triple of independent integers (m_0, m_1, m_2) and use Marton's encoding scheme to send it over the channel so that receiver 1 would obtain S_1 from its knowledge of (m_0, m_1) and receiver 2 would obtain S_2 from its knowledge of (m_0, m_2) . As is illustrated via an example this simple procedure happens to be not optimal even in cases where the Marton's region is in fact the capacity region of the broadcast channel (*e.g.*, the Blackwell channel). This implies that in this case the separation theorem fails, *i.e.*, separate source and channel codes are not optimal.

The broadcast channel with correlated sources has been also considered using a different setting in [37]. The system involves a source coding module that provides correlated indexes and then a channel coding module that uses binning. In the source coding module the sources are efficiently mapped into a nearly semi-regular bipartite graph, and in the channel coding module, the edges of this graph are reliably transmitted over the broadcast channel. The graphs are used to model the correlation between the messages. A partial characterization of the set of all graphs that can be used to represent a given pair of correlated sources is provided.

In [197] the author discusses the problem of reliable transmission of a discrete memoryless source over a discrete memoryless broadcast channel, where side information is available at each receiver (as

well as Slepian-Wolf) about the source and this side information is unknown to the sender. They show for the case of two users that the optimal coding strategy is based on separate source and channel codes. This means to build two independent binning structures, one for Wyner-Ziv source coding and the other for the broadcast channel coding. Interestingly enough, a generalization of this technique with arbitrarily number of users (more than two) does not fully resolve the case. Whereas joint source-channel coding allows for a much simpler coding strategy, without explicit binning, that yields for all channels a complete single letter characterization of the achievable rate region. The optimal strategy can be separated into source and channel components that operate independently. However, this does not imply *information separation* in the classical sense.

A more general problem consists of sending correlated sources through a broadcast channel network. This scenario arises, *e.g.*, when N transmitters desire to send N correlated sources through a noisy broadcast network with M receivers. Specifically, each transmitter is broadcasting a different source and each receiver wishes to estimate all sent sources from the corresponding channel outputs of the different broadcast channels. In this context, it is known that the source-channel separation theorem breaks down and the achievable rate region as well as the optimal coding scheme are unknown. Recently in [38], authors extended the above mentioned result for a single broadcast channel by showing that a type of source-channel separation is possible at the transmitters of broadcast channel networks, while joint source-channel decoding must be used at the receivers. The optimal rate region is not established, but the results illustrate that *operational separation*, in which at least joint source-channel decoding is performed, is general need for broadcast channels.

A network communication setup is considered in [14], where each node can exchange messages (the sources) with their neighbors over different discrete memoryless channels. These channels are assumed to be independent and after a finite number of communication rounds, each node must have received enough information to reproduce the sources sent. In this case, the general source-channel separation theorem holds, and Shannon information behaves as a classical network flow. The reason for this is that the considered model allows only for independent point-to-point channels between pairs of nodes, and not multi-user channels for which separation is well known not to hold. In the same context, authors in [46] have studied wireless erasure networks. The network model incorporates the broadcast nature of the wireless environment by requiring each node to send the same signal on all its neighbors. The capacity is obtained by assuming that all erasure localizations on all the links of the network are provided to the destinations.

Following a similar setting, but considering joint source-channel coding and decoding, in [156] authors have examined the issue of separation and code design for networks with finite alphabets. They establish that source-channel separation holds for several canonical network examples like the noisy multiple access channel and the erasure degraded broadcast channel. Whereas, the results show that separation may fail to achieve the optimal performance when additive noise is input-dependent. Though source-channel separation may not hold in general for broadcast channels, but they concluded that it does hold for the erasure broadcast channel for which linear codes are also optimal.

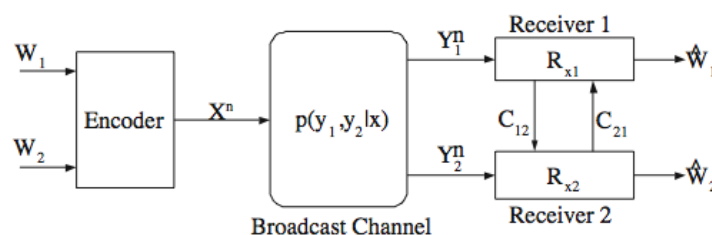


Figure 19: The broadcast channel with cooperating receivers.

The increasing interest in wireless networking motives the consideration of broadcast scenarios in which each node in the network, besides decoding its own information, tries to help other nodes in de-

coding. This problem comes up naturally in sensor networks, where a transmitter external to the sensor network wants to download data into the network, *e.g.*, to configure the sensor array. This general problem with correlated sources has not yet been considered. Whereas the transmission of discrete messages over discrete memoryless broadcast channels with cooperating receivers has been considered in [45]. In this setting, the receivers are able to exchange messages over noiseless conference links of finite capacities, prior to decoding the messages sent from the transmitter (see Figure 19). The capacity region for the case of degraded broadcast channels has been characterized, and an achievable rate region is provided for the case of general broadcast channels. Possible extensions are to consider joint channel-source coding in this cooperating scenario.

Relay Channels Network communication becomes more reliable and efficient when devices support each other to transmit data. Depending on the transmitters and receivers architectures and the information flow, there are many different configurations of networks. For example cooperative multiple access channels and cooperative broadcast channels, where relay links are incorporated into the transmitters or into the receivers, respectively, so that they are able to exchange their messages.

The three-terminal relay channel (source, relay and destination) was proposed by van der Meulen [204] and studied by Cover and El Gamal [41]. The relay node in this model only assists the transmission from the source to the destination, it is neither a source nor a sink of information (see Figure 20). A generalization of the classical relay channel has been proposed and investigated in [184]. In addition to the traditional communication from the source to the destination, the source has a private message for the relay, and the relay has a private message for the destination. Coding strategies based on decode-and-forward and compress-and-forward schemes are investigated. Achievable rate regions as well as outer bounds on the capacity region are obtained for the discrete memoryless relay channel with private messages. In this case the relay node is at the same time a receiver and a source of data.

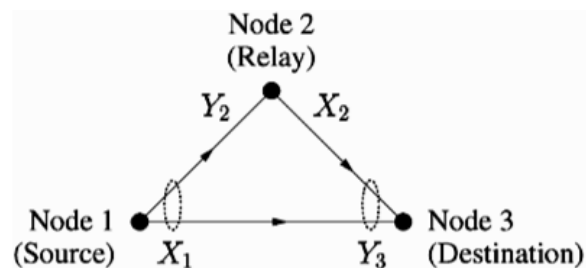


Figure 20: A one-relay channel.

Concerning relay network of communication, the capacity region of the two relay broadcast channel with partial and fully cooperation is investigated in [113]. Achievable rate regions are derived based on the relay using the decode-and-forward scheme. Also outer bounds on the capacity regions are derived and are shown to be tighter than the corresponding cut-set bounds. These results illustrate that relaying and user cooperation are powerful techniques for improving the capacity of broadcast channels.

More general relay networks that exploit node cooperation have been developed in [104]. The two basic schemes are studied in the context of broadcast and multiple access channels, where the relays decode-and-forward the source message to the destination, or they compress-and-forward their channel outputs to the destination. One relay network is the multiple access relay channel with two users, which has four nodes: nodes 1 and 2 transmit the independent messages m_1 and m_2 at rates R_1 and R_2 , respectively. While node 3 acts as a relay and node 4 is the destination for both messages. This model might fit a situation where sensors (the sources) are too weak to cooperate, but they can send their data to more powerful nodes that form a "backbone" network. This channel has three inputs X_1, X_2, X_3 and two outputs Y_3, Y_4 (see Fig. 21).

Another relay network is the broadcast relay channel with two users and three independent messages m_0, m_1, m_3 , which is depicted in Fig. 22. Node 1 transmits m_0 at rate R_0 to both nodes 3 and 4, m_1 at rate

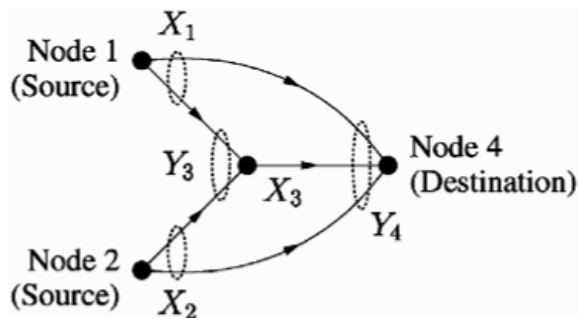


Figure 21: A multiple access relay channel with two sources. Node 3 is the relay.

R_1 to node 3, and m_2 at rate R_2 to node 4. Node 2 acts as a relay. Such a model might fit a scenario where a central node forwards instructions to a number of agents (the destinations) via a relay. More general networks with many relay nodes have been also proposed.

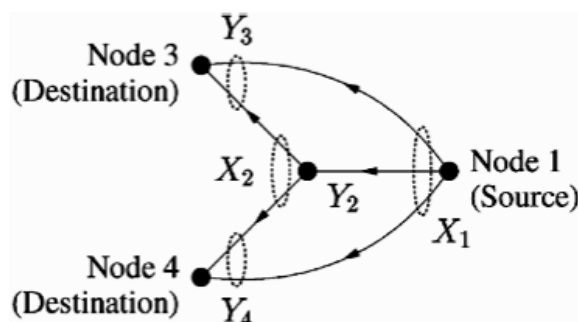


Figure 22: A broadcast relay channel with two destinations. Node 2 is the relay.

In the study of relay networks, the set-up considered through these models is based on the separation principle, which consists in first to compress the respective acquired signals, potentially applying the concepts of distributed source coding. Then, the source communicate the compressed version to the interested party (the destination) assisted by one or many relay nodes. However, for this multi-terminal scenario, the separation theorem does not extend. Consequently, cooperative joint source-channel coding and decoding strategies are needed.

Despite of recent progress, developing an unified theory for network information flow remains an elusive task. In [59] the capacity of a particular large Gaussian relay network is determined in the limit as the number of relays tends to infinity. It is shown that uncoded transmission achieves optimal performance. Joint source-channel coding strategies for this scenario are proposed in [58].

In [108] the authors have been considered, perhaps, the most simplified scenario of wireless networks. A network composed of three nodes each observing a different correlated source (see Fig. 23). Nodes are interested in obtaining a subset or all the source variables at the other nodes. To archive this goal, the nodes are allowed to coordinate and exchange information over the wireless channel. A new cooperation scheme that exploits the wireless feedback is proposed. This scheme combines the beneficence of the decode-and-forward and compress-and-forward strategies. The derived achievable rates show the value of noisy feedback in relay channels and the need of joint source-channel decoding to efficiently exploit receiver side information in the wireless setting.

In the context of sensor networks, by using a similar set up, a novel cooperative source-channel coding scheme that exploit correlation between the sources and the wireless channel is proposed in [135]. The basic challenge in such networks is the reliable transmission of the correlated observations available at the different sensors to one, or more, collector nodes. In this case the separation of source and channel coding entails significant performance losses. The cooperation schemes strive to map the correlation between

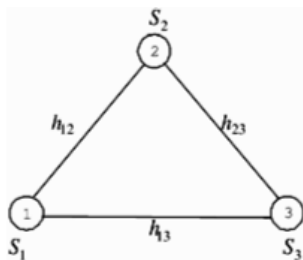


Figure 23: The three-node wireless network. Each node may be interested in a subset or all the observation variables distributed across the network.

the source observations into correlation between transmitted signals. This is the first attempt to construct explicit coding schemes for statistical cooperation. Achievable rates associated to the proposed coding schemes establish the gain allowed by the proposed techniques that efficiently exploit the correlation between the sources.

Throughout this brief review of the main results in wireless networks with cooperative nodes, it turns out that we have a limited understanding of general source-channel schemes for such multi-terminal systems. Furthermore, even for the more basic scenarios, optimal joint source-channel schemes for cooperative communications and sufficient conditions for separation are not known.

The interest in this research area has been rising sharply in recent years. In particular, Gunduz and Erkip [71] has been considered reliable transmission of a discrete memoryless source over a cooperative relay broadcast channel, where both the relay and the destination terminals want to obtain a lossless reconstruction of the source. This model assumes that both the relay and the destination have correlated side information. The achievable rates are based on the operational separation (*i.e.*, separate source and channel codes that interact at the decoding stage). Note that this coding scheme does not provide information separation in Shannon sense. The relay channel decoder uses the information provided by the relay source decoder from the previous block, similarly the destination channel decoder uses the output of destination source decoder from the next block. Also, the relay uses the decoded source codeword from the previous block to find its transmitted channel codeword. Finally, the optimality of operational separation is proved for a physically degraded relay channel with degraded side information and for an arbitrary relay channel with perfect destination relay feedback and degraded side information.

This problem has been also considered in the context of Gaussian correlated sources in [72]. Several cooperative joint source-channel strategies are proposed. In general, coding schemes depend on the correlation among the source and the relay signals, and the average link qualities. Results show that source cooperation performs well when correlation is high and the source relay link has low quality, while channel cooperation performs well for low correlation cases. Hybrid schemes extend the benefits to a wide range of correlation and channel conditions, and for most cases perform very close to the lower bound.

4.1.2.3 Sending correlated sources over noisy channels: practical aspects

An example of degraded side information is the case of encoding two correlated sources and decoding them jointly in an iterative manner. For the decoding of the first source, the other plays the role of side information, which may be corrupted. The roles are exchanged for the decoding of the second source. Several joint source-channel coding schemes have been proposed, extending schemes designed for the noiseless case.

Punctured turbo codes have been proposed in [56] for the noiseless compression of two correlated binary sources. Any point in the Slepian-Wolf achievable rate region is obtained without time-sharing.

For both sources \mathcal{S}_x and \mathcal{S}_y , blocks of n bits are formed. For \mathcal{S}_x , ℓ bits among n are encoded using a standard source coder, the $n - \ell$ remaining bits are then encoded using a systematic channel code, the systematic bits and some parity bits being eliminated after encoding. Similarly, for \mathcal{S}_y , $n - \ell$ bits

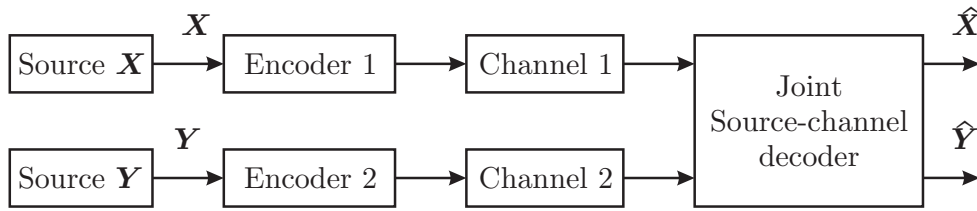


Figure 24: Joint source-channel coding of two correlated sources transmitted through two different noisy channels

are encoded using a standard source coder, the ℓ remaining bits are encoded using a systematic channel coder, the systematic bits and some parity bits being again removed after encoding. At the decoder, the ℓ and $n - \ell$ bits encoded using standard source coders are decoded first and used as side information for the recovery of the remaining bits. Assuming that $R_x < H(X)$ and $R_y < H(Y)$, one gets the proportion of bits which have to be encoded using a standard source coder

$$\frac{\ell}{n} = \frac{H(Y) - R_y}{H(X) + H(Y) - R}$$

$$\frac{n - \ell}{n} = \frac{H(X) - R_x}{H(X) + H(Y) - R}$$

with $R > H(X, Y)$.

Simulation results show that the joint probability distribution $p(x, y)$ between sources does not need to be known at the decoder site, since it can be estimated jointly with the decoding process.

A variant of this scheme has been used in [224] for the compression of two correlated sources, where the correlation is assumed to be described by a hidden Markov model. Each source is encoded independently of the other using a turbo code. Both sequences are divided into blocks of length n , and for each block, the information bits generated from the second source are interleaved using a spread interleaver. The desired transmission rate (in joint source-channel coding) is achieved by puncturing. The puncturing scheme uses half of the systematic bits for each turbo encoder (the punctured bits should not overlap), and uniformly puncture the parity bits to achieve the desired rate, see [57] for more details. In order to facilitate the estimation of the state of the hidden Markov model, a percentage of overlapping systematic bits are not punctured. At decoder side, two turbo decoders are involved. They exchange extrinsic information processed by the hidden Markov model, which state may also be estimated using the Baum-Welch algorithm [19, 214].

When each encoded source is sent over an independent channel, see Figure 24, the problem has been show to be separable, see [13]. Nevertheless, using a joint approach consisting in reducing the puncturing, as described in [224], is much simpler. Assuming that R_{c_x} and R_{c_y} denotes the rate at which the turbo codes encode independently X and Y , then, the following conditions have to be satisfied to obtain a lossless compression

$$H(X|Y) \leq \frac{C_x}{R_{c_x}}$$

$$H(Y|X) \leq \frac{C_y}{R_{c_y}}$$

$$H(X, Y) \leq \frac{C_x}{R_{c_x}} + \frac{C_y}{R_{c_y}}$$

where C_x and C_y are the capacities of the channels on which the compressed X and Y are transmitted respectively. The convergence is slower when the parameters of the hidden Markov model have to be estimated at decoder side, but the performance remain those fo the case of know correlation parameters.

In [225, 226], a similar setup has been employed using systematic linear codes with low-density generator matrix (LDGM codes) in place of the turbo codes. Such codes, which can be viewed as a

particular class of LDPC codes present a coding complexity similar to that of turbo codes and a decoding complexity comparable to that of LDPC codes. Nevertheless, concatenated structures are necessary to reduce the error floor of the LDGM codes.

4.2 Source coding with side information of time-varying quality at the decoder

Imperfect knowledge of the statistics of the noise on the correlation channel generally induces heavy degradation of the performance in classical schemes for source coding with side information at the decoder. In real scenarios, the correlation noise varies with time, without the decoder nor the encoder being informed. With respect to the application of video coding, for example, the decoder estimation of the side information may result, occasionally, in very poor prediction, whereas a good quality is guaranteed for the remaining time. Furthermore, motion compensation techniques may result, i.e. in the case of objects appearing for the first time in the frame, in unavailable side information samples. A first attempt to cope with this problem is the introduction in [17, 213] of two new models for the correlation channel, motivated by the assumption that is possible to estimate the fraction of time for which the quality of the side information will be (significantly) degraded.

Let $\{X_j, Y_j\}_{j=1}^{\infty}$ be a sequence of independent drawings of a pair of correlated continuous random variables (X, Y) . Assume $X \sim \mathcal{N}(0, \sigma_x^2)$. The correlation is defined assuming a memoryless channel between X and Y :

$$Y = X + N_b + B \cdot N_i, \quad (18)$$

where N_b and N_i are drawings from random processes of known (similar) statistics, with $\sigma_b^2 \ll \sigma_i^2$. The random Bernoulli process B with $\Pr(B = 1) = p$ models the occurrence of impulsive noise on the side information sample. Neither the encoder nor the decoder are informed of the value of B . Consider now the case of side information that is occasionally missing at the decoder side. The correlation channel in this case is defined by

$$Y = (1 - B) \cdot (X + N_b) + B\Delta \quad (19)$$

where the Bernoulli process B models the occurrence of the erasure of the side information sample, and Δ is a special erasure symbol. It is assumed that the encoder is not aware of the state on the correlation channel, whereas the decoder is able to detect the erased positions in a sequence.

4.2.1 Practical design

Practical schemes for time-varying correlation must deal with the problem of estimating the quality of the side information, since neither the encoder nor the decoder have access to the extra information about the state of the correlation channel. The problem is simpler for the case of erased side information, since the decoder (only) is aware of the positions of the missing samples. Practical schemes for both models can be obtained exploiting a refinement principle: a first layer, which induces a rate cost dependent only on the parameter p , has the aim to output a first (rough) estimate of the source sequence, compensating for the positions where the quality of the side information is poor; the refinement layer regards this first estimate as constant-quality side information, and thus can be implemented using classical Wyner-Ziv coding techniques. A practical solution for the problem in case of Gaussian correlation noise with intermittent side information is proposed in [17]; the scheme can be adapted to the case of degraded quality of the side information as discussed in the following.

4.2.1.1 Practical Wyner-Ziv coding for intermittent side information

Consider the following correlation channel between the source sample $X \sim \mathcal{N}(0, \sigma_x^2)$ and the side information sample Y

$$Y = \begin{cases} X + N_b & \text{with probability } (1 - p) \\ \Delta & \text{with probability } p \end{cases} \quad (20)$$

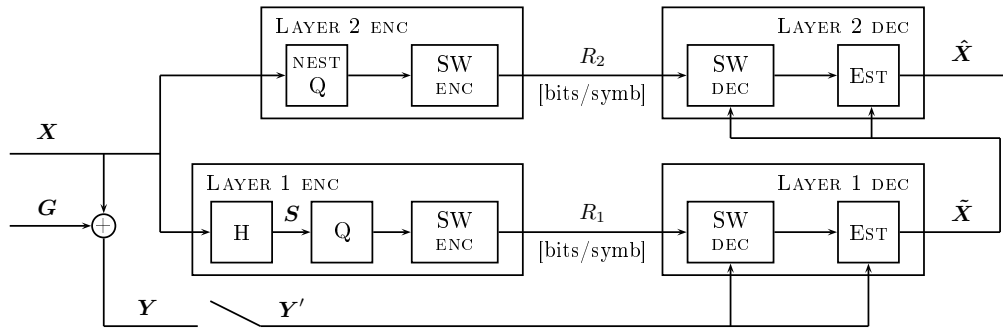


Figure 25: Practical two-layer coding scheme.

with $N_b \sim \mathcal{N}(0, \sigma_b^2)$ correlation noise. The two-layer encoding scheme proposed in [17] is depicted in figure 25. Both layers operate on blocks of N i.i.d. source samples. Layer 1 aims to produce a coarse estimation of the source, \tilde{X} , compensating for the erased entries. Layer 2 regards \tilde{X} as side information, employing standard Wyner-Ziv coding techniques to provide the finer estimate \hat{X} within target distortion per source symbol D . The rate required by the system is then $R = R_1 + R_2$ bits per source symbol. Layer 1 encoder consists of a precoding block followed by a quantizer. The aim of precoding is to reduce the dimension of the sequence to be conveyed by a factor $\frac{N-K}{N}$, while preserving the information needed to estimate the erased entries. Precoding consists in applying some linear transform matrix $H_{(N-K) \times N}$, whose rows are required to be orthonormal. Each sample is then quantized using a uniform quantizer. For values $R_1 > \frac{(N-K)}{N} \cdot 2.5$ bits, the rate on Layer 1 for target distortion D_q per source sample is well approximated by the high-rate expression

$$R_1(D_q) = \frac{N-K}{N} \left(\frac{1}{2} \log_2(2\pi e \sigma_x^2) - \frac{1}{2} \log_2(12D_q) \right) \quad (21)$$

bits per source sample. The quantization indexes are then Slepian-Wolf encoded. The same transformation and quantization are applied to the side information at the decoder to reconstruct the quantized transformed samples. Using the transmitted bits it is possible to recover an estimate of S . The decoder exploits all the information available on the source sequence to produce the MAP estimate \tilde{X} . The observation vector is expressed as

$$A = \begin{bmatrix} S \\ V \cdot Y \end{bmatrix} = \begin{bmatrix} H \\ V \end{bmatrix} \cdot X + \begin{bmatrix} Q \\ V \cdot G \end{bmatrix} = F \cdot X + Z, \quad (22)$$

where $V_{(N-v) \times N}$ is obtained from I_N by eliminating the v lines corresponding to the erased positions in Y' , and Q is the quantization noise vector. The distortion per source symbol on Layer 1 output is

$$D_1 = \frac{1}{N} E[\|X - \tilde{X}\|^2] = \frac{1}{N} \text{trace}(\Gamma_{x-\tilde{x}}), \quad (23)$$

where $\Gamma_{x-\tilde{x}} = E[(X - \tilde{X})(X - \tilde{X})^T]$ is the covariance matrix of the estimation error. Layer 2 may then use any Wyner-Ziv coding scheme for Gaussian correlation. Here we consider nested scalar quantization followed by a Slepian-Wolf coder, described in [116]. The output of the nested quantizer is the coset leader of the source X , whose correlation with the side information \tilde{X} is exploited to achieve compression. As in [116] we assume theoretical Slepian-Wolf rate. In order to match the distortion constraint D , Layer 2 requires rate

$$R_2 = \frac{1}{2} \log_2(2\pi e D_1) - \frac{1}{2} \log_2(12D) \quad (24)$$

bits per source sample. The overall distortion per source symbol of the estimate \hat{X} is derived from (21), (24), and (27) as

$$D(R_1, R_2) = \frac{1}{N} \frac{2\pi e}{12} \text{trace}(\Gamma_{x-\tilde{x}}(R_1)) 2^{-2R_2}. \quad (25)$$

The MAP estimator allows the estimation of \tilde{X} even when more than $N - K$ samples are erased. Its expression is

$$\tilde{X} = \tilde{X}_{\text{MAP}} = (F^T \Gamma_z^{-1} F + K_x^{-1})^{-1} F^T \cdot A; \quad (26)$$

the covariance matrix of the estimation error is

$$\begin{aligned} \Gamma_{\text{MAP}} &= E[(\tilde{X}_{\text{MAP}} - X)(\tilde{X}_{\text{MAP}} - X)^T] \\ &= (D_q^{-1}(R_1) H^T H + \sigma_g^{-2} V^T V + K_x^{-1})^{-1}. \end{aligned} \quad (27)$$

4.2.1.2 Results

Consider the model for $p = 0.05$, $\sigma_x^2 = 1$, $\sigma_b^2 = 0.04$. H is obtained by extracting $N - K$ rows of a size N discrete Fourier transform matrix. The size N and the $N - K$ rows are chosen such that row i and row $N - K + 1 - i$ are hermitian symmetric. Thus, for a real-valued vector, after a multiplication by H , only $N - K$ real entries need to be transmitted. Figure 26 shows the performance of Layer 1 of the coding scheme, for codes with fixed $N - K$, and varying N . For $p = 0.05$ the choice of $N - K = 4$ provides enough information to estimate the erased entries of the side information (the probability of having more than 4 erased entries per block is about 10^{-4}). The influence of the increase of R_1 becomes negligible for $R_1 > 0.7$ bits per source sample. Layer 2 is now appended to Layer 1 (R_1 is fixed at 0.75 bits per

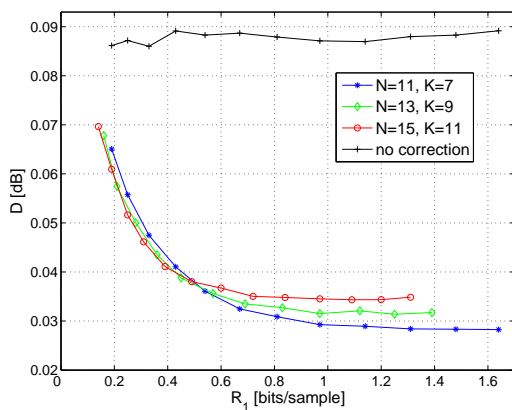


Figure 26: Layer 1 performance

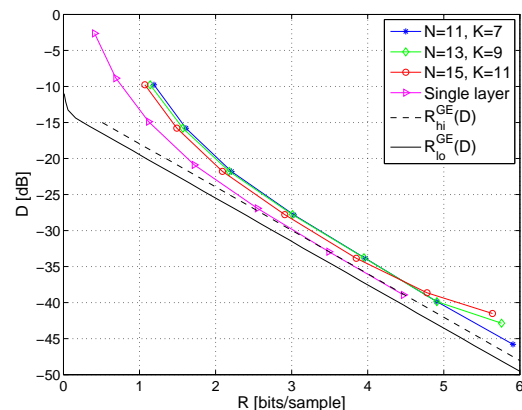


Figure 27: Two-layer scheme performance.

source sample). The performance of the entire coding scheme is represented in Figure 27. The width of the coarse quantization cell is taken as 4, in order to guarantee a vanishing probability of Wyner-Ziv decoding error. The rate is the sum of R_1 and the rate at the output of the ideal Slepian-Wolf coder of Layer 2. Figure 27 also shows the performance of a theoretical single-layer coding scheme (quantization of the source sequence followed by ideal Slepian-Wolf encoding). The two-layer scheme performs about 1.5 dB away from the upper bound of the rate-distortion curve (code $N = 15, K = 11$). A floor can be observed at high rates, for longer codes ($N = 13, N = 15$). This is due to the increased probability of more than $N - K$ erasures per sequence.

4.3 Perspectives

4.3.1 Distributed coding with joint sparsity models

The scheme for intermittent side information in [17] can be adapted for the case of correlation channel as in Eq. 18. For this purpose only Layer 1 needs to be modified, introducing a tool for the estimation of the impulse noise pattern to enable correction. The precoding matrix $H_{[N \times (N-K)]}$ is chosen as a parity check matrix of a (N, K) channel code over the real field [157]; BCH codes over the reals [55] are good candidates. If the correlation scheme is $Y = X + B N_i$, the sequence S becomes a syndrome that can be used by the decoder to correct the side information sequence Y , allowing compensation of up to $t = (N - K)/2$ impulse errors per sequence. In the correlation model Eq. 18, however, the decoded sequence \hat{S} is corrupted by the transformed background noise component $H N_b$, as well as by the quantization noise. A modified version of the algorithm used to estimate the impulse pattern is necessary to take this noise into account.

A different practical coding approach, based on LDPC codes, has been introduced in [213] for sparse correlation models with (Uniform-)Bernoulli-Uniform noise distribution. This approach will be generalized to other sparse correlation noise distributions and to distributed compressed sensing scenarios.

4.3.2 Robust DSC

DSC has the interesting property to allow the same compression rate as joint compression schemes while performing independently the compression (but not the decoding) of several correlated sources.

A majority of works available in the literature assume that perfect channel coding is performed at the edge of the system. In practical situations, codes are however imperfect and the outputs of the SW coders may not necessarily be recovered perfectly at the decoder side. The degradation introduced by the transmission channel can therefore compromise the efficiency of the DSC mechanism. Hence, robust DSC taking these imperfections into account must be designed.

Note moreover that in some situations, the separation theorem is no longer valid: considering separate source and channel coding is then no longer optimal. This is for example the case for interfering channel for which counterexamples can be found where (distributed) joint source-channel coding performed better than separated coding scheme.

The goal will therefore be the design and the optimization of distributed joint source-channel coding scheme. A few contributions have already considered the problem of robust distributed source coding, see [163, 183]. However, these contributions only consider an asymmetric scenario and assume that the side information is perfectly transmitted to the decoder. We intend to extend the tools and the methodologies which have been developed in the context of (non-distributed) joint source-channel coding and distributed source coding to more general problem of distributed joint source-channel coding, such as the non-asymmetric case or different types of channel degradations.

4.3.3 Non asymmetric DSC: robust estimate of the sources, when the difference pattern (between the sources) has error

First attempts to construct DSC codes have considered the so-called asymmetric DSC problem, i.e., the problem of coding with side information, corresponding to a corner of the SW region. Unfortunately, the asymmetric problem does not fit with realistic scenarios when dealing with practical communication problems. In particular, it has been shown in [3] that the optimal rate allocation depends on the transmission conditions and can therefore be any point of the SW region. Hence, methodologies which can operate in the so-called symmetric and non-asymmetric modes have been designed. Both syndrome-based and parity-based approaches have been considered, see e.g. [56, 60, 166, 200].

Here, we intend to focus on the syndrome-based method proposed in [60]. This method proceeds as follows. The first k' bits of source \mathbf{x} , the next $k - k'$ bits of source \mathbf{y} and the syndrome of both sources are transmitted. At the receiver, one first proceeds by estimating the "error pattern" $\mathbf{z} = \mathbf{x} \oplus \mathbf{y}$. Once the error

pattern is found, the subsequences of information bits $\mathbf{x}_{k'+1}^k$ and $\mathbf{y}_1^{k'}$ can be retrieved from the error pattern $\mathbf{z} = \mathbf{x} \oplus \mathbf{y}$ as

$$\hat{\mathbf{x}}_{k'+1}^k = \mathbf{y}_{k'+1}^k \oplus \hat{\mathbf{z}}_{k'+1}^k \text{ and } \hat{\mathbf{y}}_1^{k'} = \mathbf{x}_1^{k'} \oplus \hat{\mathbf{z}}_1^{k'}. \quad (28)$$

Subsequences of $n - k$ bits, i.e. \mathbf{x}_{k+1}^n and \mathbf{y}_{k+1}^n , remain to be computed for both sequences. Let us assume that $H = (A \ B)$, where B is an invertible square matrix of dimension $(n - k) \times (n - k)$. Note that for a rate k/n channel code, the parity check matrix H has rank $n - k$. Therefore, one can always find a permutation to be applied on the columns of H so that the resulting parity check matrix has the right form. Thus, $\mathbf{s}_x = H\mathbf{x} = (A \ B)\mathbf{x} = A \mathbf{x}_1^k \oplus B \mathbf{x}_{k+1}^n$, and the remaining $n - k$ unknown bits of the sequence \mathbf{x} (and similarly for the sequence \mathbf{y}) can be computed as

$$\hat{\mathbf{x}}_{k+1}^n = B^{-1} (\mathbf{s}_x \oplus A \hat{\mathbf{x}}_1^k), \quad (29)$$

where B^{-1} denotes the inverse of the matrix B .

In the case the error pattern is not perfectly estimated, the estimation of the remaining symbols \mathbf{x}_{k+1}^n in equation (29) may yield error propagation. Indeed, although B may be designed to be sparse, its inverse is not necessarily so. As a consequence, a few errors in the estimation of \mathbf{z} may lead to dramatic degradations in the estimation of \mathbf{x} and \mathbf{y} .

A future direction of research will therefore consist in taking this kind of behavior into account and proposing procedures to design matrix B so that error propagation is limited.

4.3.4 Robust DSC employing distributed arithmetic codes

DSC coders typically represent the source using its syndrome (or the parity bits) of a suitable channel code of given rate. Although near-capacity DSC coders have been designed for simple ideal sources, their application to realistic signals typically incurs several problems. Capacity-achieving channel codes require very large data blocks (typically in excess of 10^5 symbols), while in many practical applications the basic units to be encoded are much smaller. Real-world sources are not stationary and would benefit from context-based modeling, which is not provided by existing coders. Furthermore, upgrading an existing compression algorithm like JPEG 2000 or H.264/AVC in order to provide DSC functionalities requires at least to redesign the entropy coding stage, adopting one of the existing DSC schemes. Among these issues, the block length is particularly important. While the performance of some practical DSC coders on ideal sources with very large block length can be as close as 0.09 bits per symbol (bps) to the theoretical limit [221], so far DSC of real-world data has fallen short of its expectations, one reason being the necessity to employ much smaller blocks; e.g., block lengths between 200 and 2000 symbols are typically employed in multimedia applications.

To address these issues, in [8, 66] the Distributed Arithmetic Coding (DAC) concept has been proposed. Binary DAC is a generalization of arithmetic coding, in which, at each coding step, the intervals corresponding to symbols "0" and "1" are allowed to overlap. The overlapping introduces an ambiguity in the source description, much in the same way as the syndrome coders; the larger the overlap, the more the ambiguity, and the less the coding rate. The decoder employs sequential decoding of arithmetic codes, aided by a correlated side information signal. In [66, 68] it has been shown that DAC has similar or better performance than turbo and LDPC codes at small and medium block lengths. A similar scheme has been developed in [8] using quasi-arithmetic codes. Quasi-arithmetic codes are a low-complexity approximation to arithmetic codes, providing smaller encoding and decoding complexity, and slightly poorer performance [78]. These codes can be modeled as finite state machines, simplifying the decoding process. In [69] a rate-compatible version of DAC has been proposed, showing that the decoding can be made error-free by incrementally sending refinement bits upon the decoder's request.

Moreover, DAC can also be employed for symmetric DSC, i.e., when both sources are coded with ambiguity. In particular, in [67, 68] it has been shown that two sources can be encoded using separated DACs and decoded jointly, with performance similar to that achieved in the asymmetric case. The two (or more) DACs must be set up in such a way that, at each input symbol, ambiguity is inserted in only one

sequence; this keeps the complexity of the decoding problem manageable. Interestingly, this amounts to perform time-sharing among the sources. The joint decoder constructs a decoding tree where each node stores the state of the two (or more) arithmetic decoders, and the most likely output sequence is selected at the end of the arithmetic decoding process.

Syndrome-based DSC coders are clearly amenable to distributed joint-source channel coding/decoding, as the amount of redundancy can be easily increased to take into account the presence of channel errors. The DAC coder is also amenable to such extensions, which will be subject of future work. In particular, it is known that binary arithmetic coders can be used for error correction using a ternary alphabet with a forbidden symbol. The same concept can be applied to the DAC, by considering a decoder branch metric that also models the channel behaviour, and using a ternary alphabet.

5 TASK TR7.4 TOOLS FOR EVALUATING THE EFFICIENCY OF JSCC/D METHODS

The focus of Task 7.4 is on theoretical tools and methods that can be used to predict the performance of JSCC/D systems, or to benchmark the performance of such systems. Examples of the former are methods to obtain code distance spectra, which lead to performance prediction via the union bound, and can be applied in code optimization procedures. Examples of the latter are any kind of lower or upper bounds on the theoretically or practically attainable performance, often obtained by information-theoretic methods or by analysis of ideal systems.

5.1 State of the art

5.1.1 Performance characterization of arithmetic-coding-based JSCC/D using distance spectra

Arithmetic coding (AC) [216] is currently being deployed in a growing number of compression standards, e.g. H.264 and JPEG2000, as it yields higher compression performance when compared to other compression methods. However, its efficiency makes AC particularly vulnerable to transmission errors. This issue has motivated the recent development of joint source-channel techniques for AC-encoded data. Improving the robustness of AC against transmission errors is usually achieved by introducing redundancy in the compressed bitstream. This redundancy is often introduced by means of a forbidden symbol (FS). Many approaches for introducing a forbidden symbol have been developed during the last years [51], [79], [70]. The comparison between these different techniques is usually experimental and is restricted to simulations. To overcome this limitation, [20] proposed first analytical tools that allow to characterize and objectively compare these techniques. The approach in [20] has been strongly inspired, first, by the classical results on the error correction properties of convolutional codes, and more generally of linear codes [114], and second, by the extension of these results to Variable Length Codes (VLC) and Variable Length Error-correcting Codes (VLEC) [32]. In [20], they consider a practical integer-based implementation of AC for a memoryless source. Then, they develop a specific Finite State Machine (FSM) and trellis representation which is suited to efficient asymptotic error rate evaluation, unlike the trellis representations of [32] and [51], which serve other purposes.

The error correcting properties of a code can be characterized by a set of parameters known as *distance spectrum*, with the single most important parameter being the *free distance* (d_{free}). These *distance properties* of the code can be obtained from the same trellis representation of the code which is used in a maximum likelihood decoding algorithm. The *free distance* is the minimum Hamming distance between all pairs of paths of the same length on the trellis, with the same start and end points. It determines the dominant error event affecting a code. The *distance spectrum* $\{A_d\}$, with $d = d_{free}, d_{free} + 1 \dots \infty$, counts the average number of pairs of paths of the same length at Hamming distance d . It allows to obtain an upper bound on both bit and symbol error probability through the *union bound*. The *distance spectrum* can be compactly represented by the *generating function* $T(D)$:

$$T(D) = \sum_{d=d_{free}}^{+\infty} A_d D^d. \quad (30)$$

The first error event is said to occur at an arbitrary bit position r , corresponding to state s_r in the trellis representation, if the correct path is eliminated for the first time at bit position r in favor of a competitor (the incorrect path). The incorrect path must be some path that had previously diverged from the correct one at some bit position q , $q < r$, and is now converging back for the first time at bit position r . Suppose that the Hamming distance between the correct path and the selected path is d . Let P_d be the probability of selecting the wrong path. For a binary symmetric channel with cross-over probability p , P_d is given by

$$P_d = \begin{cases} \sum_{k=(d+1)/2}^d \binom{d}{k} p^k (1-p)^{d-k}, & d \text{ odd} \\ \frac{1}{2} \binom{d}{d/2} p^{d/2} (1-p)^{d/2} + \sum_{k=d/2+1}^d p^k (1-p)^{d-k}, & d \text{ even.} \end{cases} \quad (31)$$

For BPSK signaling used over an AWGN channel P_d is given by

$$P_d = \frac{1}{2} \operatorname{erfc} \left(\sqrt{d \frac{E_b}{N_0}} \right), \quad (32)$$

where $\frac{N_0}{2}$ is the variance of the zero-mean Gaussian channel noise, and the bit 1 is mapped onto $+\sqrt{E_b}$ and the bit 0 onto $-\sqrt{E_b}$.

The first event error probability P_E is bounded (via the *union bound*) by

$$P_E < \sum_{d=d_{free}}^{+\infty} A_d P_d. \quad (33)$$

The Chernoff bound allows to upper bound the probabilities in Eq. 31 by the following:

$$P_d < 2^d (p(1-p))^{d/2}. \quad (34)$$

Finally, the first even error probability in the BSC case can be more loosely bounded by

$$P_E < T(D) \Big|_{D=2\sqrt{p(1-p)}}. \quad (35)$$

The first event error probability represents the word error probability in the code domain, and Eq. 35 is an upper bound on this probability. We can also obtain in the same manner an upper bound on the symbol error probability in the symbol domain by a modification of the generating function. It means that in the generating function, the spectrum $\{A_d\}$ is replaced by another spectrum $\{B_d\}$ and the equations 33, 34 and 35 are applied to obtain an upper bound on the symbol error probabilities.

We saw that with the *distance properties* of a code, one can assess its error correcting performances, i.e. its correction capability, as well as obtain upper bounds on both bit and symbol error probability. The following sections recall how the *distance properties* can be obtained for convolutional codes, VLEC and joint source-channel arithmetic coding (JSCAC), respectively.

5.1.1.1 Convolutional Codes

Convolutional codes were first introduced by Elias and consist of a linear finite state machine made of K -stages shift registers and n linear algebraic function generators. The input data, which is usually, though not necessarily binary, is shifted along the register b bits at a time. Convolutional codes have been applied as channel coding to increase the efficiency of numerous communications systems, where they invariably outperform block codes of the same order of complexity [207]. There are many representations of convolutional codes: the convolutional encoder, the tree code representation, the trellis code representation and the state diagram representation. The first two representations are suited for the encoding process. The trellis is effective for the decoding process and for calculating both *free distance* and *distance spectrum*, whereas the state diagram is used to calculate the *distance spectrum*. Due their linearity, it is easy to evaluate the *free distance* and the *distance spectrum*. The *free distance* is the minimum Hamming weight of all paths in the trellis, excluded the all-zero path. A path which has minimum Hamming weight diverges from the all zeros path at some node i and merges back with the all zeros path for the first time at some arbitrary node j , with the minimum number of branches. The *distance spectrum* $\{A_d\}$ is identically to the code's *weight spectrum*, i.e. the average number of paths with Hamming weight d . The spectrum $\{B_d\}$ is obtained as $B_d = K_d \times A_d$, for $d = d_{free}, d_{free} + 1, \dots, \infty$, where K_d is the weight of the input sequence which allow to generate the encoded message with the Hamming weight d .

5.1.1.2 Variable-Length Error-Correcting Codes

Variable length codes (VLC) are normally used for source coding. But they are very sensitive against transmission errors on noisy channels. That motivated the development of variable-length joint source-channel codes named VLEC. In [32] the authors present a way of decoding VLEC codes by treating them

as trellis codes. Consequently, it is shown that in many respects they behave very much like convolutional codes and exhibit a form of “memory”, although in the case of VLEC the memory is not related to an encoding shift register, but is due to the different codeword lengths.

Using this representation, the authors give a maximum likelihood decoding algorithm based on the Viterbi algorithm and also derive a maximum a-posteriori (MAP) metric for use with this algorithm. There are other properties of VLEC codes similar to those of convolutional codes like their *free distance*, *distance spectrum*, constraint length and catastrophic behavior. The remainder of this section is devoted to the characterization of the *free distance* and the *distance spectrum* of VLEC.

Free distance. As we saw previously, the free distance is one of the most important parameters for both convolutional codes and VLEC. In the case of convolutional codes, the free distance is easy to evaluate due to linearity. In the case of VLEC, it becomes slightly more difficult to evaluate the free distance, since these codes are not linear.

Denote \mathcal{C} a VLEC which has s codewords : $\mathcal{C} = \{c_1, c_2, \dots, c_s\}$. Let $l_i = |c_i|$, $i = 1, 2, \dots, s$ be the length of codeword c_i . Without loss of generality, assume that $l_1 \leq l_2 \leq \dots \leq l_s$. Further, let σ denote the number of different codewords lengths in the code \mathcal{C} and let these lengths be $L_1, L_2, \dots, L_\sigma$ where $L_1 < L_2 < \dots < L_\sigma$. If $\sigma = 1$, then \mathcal{C} is a fixed length code. Hence we shall define a variable length code \mathcal{C} to be a code with $\sigma > 1$. Denote F_N the set of all extended codes of code \mathcal{C} with length N . As we saw previously the authors in [32] have defined a VLEC using a trellis. Then they define the *free distance* of this code as *the minimum Hamming distance in the set of all arbitrary long paths that diverge from some common state S_i and converge again in another common state S_j , $j > i$* . More formally,

$$d_{free} = \min \{d_H^{out}(f_i, f_j) : f_i, f_j \in F_N : N = 1, 2, \dots, \infty\}, \quad (36)$$

where $d_H^{out}(f_i, f_j)$ is the Hamming distance between the extended codeword f_i and f_j .

Computing the free distance with Eq. 36 is computationally expensive, hence, in [32], the authors introduce a lower bound of the free distance. To find this lower bound, they define three parameters: the *overall minimum block distance*, b_{min} of the VLEC \mathcal{C} , the *minimum diverging distance*, d_{min} , and the *minimum converging distance*, c_{min} .

The overall minimum block distance of the VLEC code. The minimum block distance for length L_k , b_k , of the VLEC \mathcal{C} is defined as the minimum Hamming distance between all codewords with the same length L_k , i.e.

$$b_k = \min \{d_H^{out}(c_i, c_j) : c_i, c_j \in \mathcal{C}, i \neq j : |c_i| = |c_j| = L_k\}, \quad (37)$$

The overall minimum block distance, b_{min} , of a VLEC code \mathcal{C} is defined as the minimum value of b_k over all $k = 1, 2, \dots, \sigma$.

The minimum diverging distance of the VLEC code. The diverging distance between two codewords $c_i = c_{i1}c_{i2} \dots c_{il_i}$ and $c_j = c_{j1}c_{j2} \dots c_{jl_j}$, $D(c_i, c_j)$, where $(c_i, c_j) \in \mathcal{C}^2$, with $l_i > l_j$, is defined as

$$D(c_i, c_j) = d_H^{out}(c_{i1}c_{i2} \dots c_{il_i}, c_{j1}c_{j2} \dots c_{jl_j}). \quad (38)$$

Note that $D(c_i, c_j) = D(c_j, c_i)$. The minimum diverging distance d_{min} is the minimum value of all the diverging distances between all possible pairs of unequal length codewords of \mathcal{C} , i.e.

$$d_{min} = \min \{D(c_i, c_j) : c_i, c_j \in \mathcal{C}, l_i \neq l_j\}. \quad (39)$$

The minimum converging distance of the VLEC code. The converging distance between two codewords $c_i = c_{i1}c_{i2} \dots c_{il_i}$ and $c_j = c_{j1}c_{j2} \dots c_{jl_j}$, $\mathcal{C}(c_i, c_j)$, where $c_i, c_j \in \mathcal{C}$ with $l_i > l_j$, is defined as

$$\mathcal{C}(c_i, c_j) = d_H^{out}(c_{il_i-l_j+1}c_{il_i-l_j+2} \dots c_{il_i}, c_{j1}c_{j2} \dots c_{jl_j}). \quad (40)$$

Again note that $\mathcal{C}(c_i, c_j) = \mathcal{C}(c_j, c_i)$. The minimum converging Hamming distance c_{min} is the minimum value of the converging distance between all possible pairs of unequal length codewords of \mathcal{C} , i.e.

$$c_{min} = \min \{\mathcal{C}(c_i, c_j) : c_i, c_j \in \mathcal{C}, l_i \neq l_j\}. \quad (41)$$

After defining these three parameters, the free distance can be lower bounded by

$$d_{free} \geq \min(b_{min}, d_{min} + c_{min}). \quad (42)$$

Distance spectrum. In [32], two methods have been proposed to evaluate the *distance spectrum* of VLEC. However, these will be not treated here, we just give the results. Before we give the results, some notations have to be introduced. Let $p_{0,r}^i$ be the correct path segment through the trellis with initial and final state S_0 and S_r . The probability of occurrence of $p_{0,r}^i$ is $P(p_{0,r}^i)$. $C_{0,r}$ is the set of all pairs of path segment indices corresponding to path segments which diverge at state S_0 and merge again for the first time at state S_r . The *distance spectrum* $\{A_d\}$ is given by

$$A_d = \sum_{r=1}^{\infty} \sum_{\substack{(i,j) \in C_{0,r} \\ d_H^{out}(p_{0,r}^i, p_{0,r}^j) = d}} P(p_{0,r}^i). \quad (43)$$

B_d is the average Levenshtein distance of all converging pairs of paths whose encoded messages are at Hamming distance d from each other. Let $in(p_{0,r}^i)$ be the input sequence which has generated $p_{0,r}^i$. $\{B_d\}$ is given by

$$B_d = \sum_{r=1}^{\infty} \sum_{\substack{(i,j) \in C_{0,r} \\ d_H^{out}(p_{0,r}^i, p_{0,r}^j) = d}} d_L(in(p_{0,r}^i), in(p_{0,r}^j)) \times P(in(p_{0,r}^i)), \quad (44)$$

where $d_L(a, b)$ is the Levenshtein distance between a and b , i.e. the minimum number of deletions, insertions or substitutions needed to transform a into b or conversely. Note that a and b need not have the same length. Note that $P(in(p_{0,r}^i)) = P(p_{0,r}^i)$.

5.1.1.3 Joint Source-Channel Arithmetic Coding

Let X be a source that generates symbols from alphabet $A = \{a_1, a_2, \dots, a_N\}$ with occurrence probabilities $P_i = P(X = a_i)$, $i = 1 \dots N$. The basic idea of arithmetic coding (AC) is to assign to any symbol sequence generated by X a unique interval I_{coding} belonging to $[0, 1)$. This is done by a recursive partitioning of the interval $[0, 1[$ according to the symbol probabilities. Then assign to I_{coding} a single binary representation which will be transmitted [167]. In [20], the alphabet A is reduced to $\{0, 1\}$, hence the AC becomes binary AC.

Binary Arithmetic Coding. In the case of binary arithmetic coding, the current interval $[low, high)$ is partitioned into two subintervals I_0 and I_1 , the widths of which are proportional to the probabilities P_0 and P_1 of the source symbols 0 and 1, respectively. One of these interval is selected as the new source interval, according to the value of the current symbol. Once the last symbol is encoded, the encoder selects the widest dyadic *code interval* entirely contained in the current interval and assigns its binary representation to the symbol sequence. For sources with skewed and for long symbol sequences, subintervals may get too small to be accurately handled by a finite-precision computer. This problem is solved by integer binary AC.

Integer binary AC. Finite precision arithmetic coding was first introduced by Pasco [148] and Rissanen [162] in 1976. Howard *et al.* showed in [79] that a slight modification of the symbol probabilities, such that the interval bounds become rational numbers, decreases the computational cost of AC, without significantly degrading compression performance. Therefore, integer AC works like the ideal AC presented above, but using the interval $[0, T)$ of integers, where $T = 2^P$, $P \geq 2$ is the bit size of the initial interval, and rounding all interval boundaries to integers. Partition and selection are carried out every time a source symbol is encoded. Renormalization of the current interval $[low, high[$ by doubling the size is performed if one the following conditions holds:

1. if $high \leq 2^{P-1}$, low and $high$ are doubled.
2. if $2^{P-1} \leq low$, low and $high$ are doubled after subtracting 2^{P-1} .
3. if $2^{P-2} \leq low$ and $high \leq 3 \times 2^{P-2}$, low and $high$ are doubled after subtracting 2^{P-2} .

If the current interval (before renormalization) overlaps the midpoint of $[0, 2^P)$, no bit is output. The number of consecutive times this occurs is stored in a variable called *follow*. If the current interval (before renormalization) lies entirely in the upper or lower half interval of $[0, 2^P)$, the encoder emits the leading bit of low (1 or 0) and *follow* opposite bits. This is called *follow-on* procedure [216]. At the decoder side, the same operations are carried out analogously.

Trellis-based arithmetic coding. Integer binary AC can be considered as an automaton represented by a finite number of states and transitions, except for the possibly unbounded *follow* counter. This approach was first proposed by [79], where arithmetic operations were replaced by a table lookup. Integer arithmetic decoding was represented by a three-dimensional trellis taking into account the presence of FS in the source alphabet [51].

When performing integer AC, the number of possible subintervals of $[0, T)$ is finite. Considering that after renormalization, $range = high - low$ is always greater than $\frac{T}{4}$ and that $[low, high)$ cannot be a proper subinterval of $[\frac{T}{4}, \frac{3T}{4})$, the number of possible intervals $[low, high)$ is $\frac{T}{2} \times \frac{T}{2} - \frac{T}{4} \times \frac{T}{4} = \frac{3T^2}{16}$.

On the other hand, for a memoryless source and known symbol probabilities, the encoder is entirely characterized by the current interval $[low, high)$ and the value of *follow*. Hence, the encoder state can be represented by $(low, high, follow)$ as defined in [70]. For an order M Markov source, the encoder state has to be extended to include the last M encoded symbols, to properly take account the memory of the source. In [20], for the sake of simplicity, only memoryless sources have been considered.

The idea is thus to precompute all the states of the arithmetic encoder such that any source stream may be encoded using table lookups rather than arithmetic operations. However, as the variable *follow* might grow without bound, the number of states $(low, high, follow)$ could be infinite. In order to cope with this uncontrolled growth of *follow*, it can be bounded by a given threshold F_{max} as in [51]. To this end, renormalization incrementing *follow* are only performed as long as $follow < F_{max}$. Whenever $follow = F_{max}$ and the current source interval is such that *follow* could be further incremented, the symbol probabilities are modified in order to force the follow-on procedure after encoding the current symbol.

In [20], three FSMs describing the AC operations with a bound on *follow* were proposed, namely a *symbol-clock FSM*, which is well suited for encoding, a *reduced FSM* leading to a compact trellis better suited for decoding, and a *bit-clock FSM* suited for distance evaluation.

Symbol-clock FSM The first proposed FSM is such that each transition corresponds to a single encoded symbol. Hence it is called a *symbol-clock FSM*. Starting at the initial state $(0, T, 0)$, the encoder is fed by the two possible input symbols, and may either reach a new state or return to $(0, T, 0)$. This defines the two starting transitions of the FSM describing the AC. Every arrival state is then considered as a starting state of two new transitions leading either to new states or to already known ones. When no bits are output by the encoder, the transition is called mute transition. This exhaustive search stops when no new states are found. The set of states and the set of transition between states of this symbol-clock FSM are denoted by S_{FSM}^s and τ_{FSM}^s , respectively.

Reduced FSM When dealing with noisy transmissions, a trellis-based soft-input decoder may be implemented as in [51], [70]. In [20] a Viterbi decoder which keeps only the best path among all those converging in the same state at a given depth of the trellis is used. The saved path is called survivor. If the decoder uses a likelihood-based metric to compare converging paths, the associated bit sequences must be of equal length. Therefore, the symbol-clock trellis derived from the FSM described will not be appropriate for such a decoder, as merging paths would have equal length in symbols, but variable length in bits. A bit-clock trellis is needed instead. The trellis adopted in [20]

is derived from the symbol-clock FSM proposed, which is modified in order to no mute transition. This allow to reduce the number of states and thus to reduce the memory required by the Viterbi Algorithm (VA), which will work without any constraint on the evaluation order of survivors. To obtain the reduced FSM, the transitions are allowed to have more than one input symbol. In fact, every mute transition is extended until at least one bit is output.

Bit-clock FSM An efficient recursive evaluation of the free distance properties of trellis-based AC is possible if transitions have exactly one output bit. Such a trellis may be obtained from the reduced FSM introducing additional intermediate states, such that every state transition outputs exactly one bit [20]. When a given transition outputs two bit, it is split into two transitions, the first of which inherits the input symbols and the first output bit, while the second has no input symbol and outputs the second bit. The set of states in the bit-clock FSM is denoted by S_{FSM}^b , and the set of transitions by τ_{FSM}^b .

Performance analysis of trellis-based AC. In joint source-channel coding schemes, redundancy is introduced in order to allow error detection and/or correction at the decoder side. According to [79], when considering integer probability spaces in $[0, T)$, the additional redundancy due to the integer approximation is at most $\frac{0.497}{T} + O\left(\frac{1}{T^2}\right)$ bits/symbol if correct probability estimates are used by the encoder. Limiting the value of *follow* may be another source of redundancy. Nevertheless, as the probability of having $follow = F_{max}$ decreases exponentially with F_{max} , also this additional redundancy remains small. A popular joint source channel technique for AC is based on the introduction of a forbidden symbol (FS) in the source alphabet. The FS is never emitted by the source, although a positive probability P_ϵ is assigned to it. In that way, decoding the FS indicates the occurrence of a transmission error. It can be shown that introducing a FS of probability P_ϵ adds a redundancy of $-\log(1 - P_\epsilon)$ bits/symbol to the coded bitstream [20]. All the cited techniques for adding redundancy may be applied to integer AC jointly with a bound on *follow*. The resulting operations will thus be described by a FSM. Consequently, JSCAC operations may entirely rely on the trellis representation of AC. The effects of the additional redundancy in terms of error detection and correction capability appear in the trellis representation of the associated FSM. The effectiveness of this redundancy may be characterized by the distance properties derived from the trellis.

Free distance The free distance d_{free} , is the minimum Hamming distance between all paths of the same length in bits, diverging in some state and converging in another state. If the encoded messages are long enough to avoid boundary effects, d_{free} plays a similar role as the minimum distance for a block code, in that all error patterns of weight t can be corrected if $d_{free} > 2t$. For convolutional codes d_{free} is the smallest Hamming weight over all paths diverging from and converging to the all-zero path due to linearity. In the case of non linear code such as VLEC, comparing paths to the all zero path may be not sufficient to determine d_{free} , since such codes are generally not geometrically uniform. But in [32], Buttigieg deduces a lower bound on d_{free} . Extending this technique to AC is not possible, as the transition output do not satisfy the prefix condition in general. Therefore in [20], the authors propose an algorithm for computing d_{free} for trellis based AC. This algorithm relies on an iterative computation of the smallest distance between all different paths of equal length starting from a common state. Its iterative structure allows an evaluation of d_{free} with polynomial complexity in the number of states of S_{FSM}^b and in the number of transitions between these states. The iterations for computing d_{free} will be performed on the length in bits of the paths on the trellis, it is thus advantageous to use the bit-clock trellis. Before presenting the algorithm for computing d_{free} in [20], some notations have to be introduced.

A path of n bits on a bit-clock trellis, starting from state x and ending in y is denoted by p_{xy}^n . The Hamming distance between the output bits of two paths p^n and q^n of the same output length n is denoted by $d_H^{out}(p^n, q^n)$. When a path p_{xy}^n is extended by a transition tyz , one obtains a new path $p_{xz}^{n'} = p_{xy}^n \circ t_{yz}$ of length $n' = n + l(out(t_{yz}))$, l being the length function. The set of all pairs of paths diverging from the same starting state x and converging for the first time n bits later in state y is denoted by $C_n(x, y)$. The set of all pairs of paths of length n bits, diverging from the starting state x and never converging is $D_n(x, y, z)$.

Finally, let $\tau_{xy}^b \subset \tau_{FSM}^b$ be the set of transitions starting from x and ending in y in the bit-clock FSM.

To evaluate d_{free} in [20], a three-dimensional array Δ_n defined as follows : for $y \neq z$, $\Delta_n(x, y, z)$ is the minimum Hamming distance between all pairs of paths of length n , starting from the state x , ending respectively in y and z , and never converging. For $z = y$, $\Delta_n(x, y, y)$ is defined as the minimum distance between pairs of paths of most at n bits, diverging at state x and converging for the first time in state y . Using this notation, one may write

$$\Delta_n(x, y, z) = \min_{(p_{xy}^n, q_{xz}^n) \in D_n(x, y, z)} d(p_{xy}^n, q_{xz}^n), \quad (45)$$

$$\Delta_n(x, y, y) = \min_{n' \leq n} \min_{(p_{xy}^{n'}, q_{xy}^{n'}) \in C_{n'}(x, y)} d_H^{out}(p_{xy}^{n'}, q_{xy}^{n'}), \quad (46)$$

and

$$d_{free} = \min_{n, x, y} \Delta_n(x, y, y). \quad (47)$$

The evaluation of $\Delta_n(x, y, z)$ seems to require the evaluation of distance between an exponentially growing number of paths pairs. [20] provides the following iterative technique for the evaluation of $\Delta_n(x, y, z)$ with polynomial complexity. For any initial state $x \in S_{FSM}$, the Eq. 45 and Eq.46 can be evaluated recursively starting from

$$\Delta_1(x, y, z) = d_H^{out}(t_{xy}, t_{xz}), \quad (48)$$

with $y, z \in S_{FSM}^b$ and $t_{xy}, t_{xz} \in \tau_{FSM}^b$, using

$$\Delta_n(x, y, z) = \min_{y' \neq z'} \{ \Delta_{n-1}(x, y', z') + d_H^{out}(t_{y'y}, t_{z'z}) \}, y \neq z, \quad (49)$$

$$\Delta_n(x, y, y) = \min \left\{ \Delta_{n-1}(x, y, y), \min_{y' \neq z'} \{ \Delta_{n-1}(x, y', z') + d_H^{out}(t_{y'y}, t_{z'y}) \} \right\}, \quad (50)$$

where the minimization is intended over all distinct $y', z' \in S_{FSM}^b$.

For a given $x \in S_{FSM}$ assume that there is exists some N_x such that

$$\min_{y', z' \in S_{FSM}^b} \Delta_{N_x}(x, y', z') \geq \Delta_{N_x}(x, y, y). \quad (51)$$

Then the recursion Eq. 50 can be stopped, since it will be no longer possible to reduce $\Delta_n(x, y, y)$ by merging paths of paths ending in states y' and z' , respectively and 51 will be satisfied for all $n \geq N_x$. Moreover, let $\bar{N} = \max_{x \in S_{FSM}} N_x$. Then d_{free} may be evaluated as

$$d_{free} = \min_{n \leq \bar{N}, x, y} \Delta_n(x, y, y). \quad (52)$$

If the AC is non-catastrophic, \bar{N} will be finite and the evaluation of d_{free} is of polynomial complexity.

Distance spectrum. As with convolutional codes and VLEC, the distance spectrum of AC yields an upper bound on error probability. The spectrum $\{A_d\}$ and $\{B_d\}$ can be obtained as :

$$A_d = \sum_{n=d}^{\infty} \sum_{\substack{(p^n, q^n) \in C_n: \\ d_H^{out}(p^n, q^n) = d}} P(p^n), \quad (53)$$

$$B_d = \sum_{n=d}^{\infty} \sum_{\substack{(p^n, q^n) \in C_n: \\ d_H^{out}(p^n, q^n) = d}} d_L(in(p^n), in(q^n)) P(in(p^n)), \quad (54)$$

where \mathcal{C}_n is the set of all path pairs on the reduced FSM trellis diverging from the same starting state and converging for the first time n bits later (thus $\mathcal{C}_n = \cup_{x, y \in FSM} \mathcal{C}_n(x, y)$). For the evaluation of $\{A_d\}$ and $\{B_d\}$ for VLECs in [32], only paths beginning at a single initial state (corresponding to depth $n = 0$) had to be considered. In the case of AC in [20], two paths may diverge at any state of the reduced trellis and so each state has to be considered as possible diverging state, with its associated stationary probability.

5.2 Bounds on the Wyner-Ziv rate distortion function for GE and GBG correlation models

In realistic Wyner-Ziv coding scenarios the correlation of the side information changes with time. The aim of this section is to provide theoretical rate distortion bounds for the new correlation models introduced in Section 4.2.

Let $\{X_j, Y_j\}_{j=1}^{\infty}$ be a sequence of independent drawings of a pair of correlated continuous random variables (X, Y) . Assume $X \sim \mathcal{N}(0, \sigma_x^2)$. The correlation is defined assuming a memoryless channel between X and Y .

In the Bernoulli Gaussian (BG) correlation model, the side information might or might not be affected by Gaussian noise $N_i \sim \mathcal{N}(0, \sigma_i^2)$. The occurrence of a noisy side information sample is modeled by a Bernoulli process B with $\Pr(B = 1) = p$. The BG correlation channel is defined by

$$Y_{\text{BG}} = \begin{cases} X & \text{with probability } (1 - p), \\ X + N_i & \text{with probability } p. \end{cases} \quad (55)$$

The Gaussian Bernoulli Gaussian (GBG) correlation extends the BG model assuming some additional background Gaussian noise $N_b \sim \mathcal{N}(0, \sigma_b)$, with $\sigma_b^2 \ll \sigma_i^2$. The GBG correlation channel is defined by

$$Y_{\text{GBG}} = \begin{cases} X + N_b & \text{with probability } (1 - p), \\ X + N_b + N_i & \text{with probability } p. \end{cases} \quad (56)$$

In the Gaussian Erasure (GE) model the correlation channel is still affected by Gaussian background noise G_b . The Bernoulli process B with $\Pr(B = 1) = p$ models the occurrence of erasure of the whole side information sample Y ; the GE correlation channel is defined by

$$Y_{\text{GE}} = \begin{cases} X + N_b & \text{with probability } (1 - p), \\ \Delta & \text{with probability } p. \end{cases} \quad (57)$$

Notice that in the GE case the decoder is aware of the erasure occurred over the correlation channel.

Bounds on the rate distortion function. Consider the MSE distortion measure, and let D be the target distortion between the source X and its reconstruction \hat{X} . We want to determine lower and upper bounds to the rate distortion function for the proposed correlation models.

Intuitively, the rate required by the system can be lower-bounded by assuming that the encoder knows the locations of the samples affected by impulse noise (BG and GBG models) or erased (GE model). This results in a setup similar to [54], with partial side information B (the positions of the erasures or impulses) available at both encoder and decoder. The encoder works in a time-division regime governed by B , employing in each fraction of the time the minimum rate required to match the global distortion constraint D .

For the BG correlation model the optimal encoder operates on the WZ rate-distortion function $R_{\text{WZ}}^{(i)}(D/p)$ for Gaussian noise of variance σ_i^2 during a fraction p of the time, and does not send information for the remaining:

$$R_{\text{WZ}}^{\text{BG}}(D) \geq R_{\text{WZ}_{i_o}}^{\text{BG}}(D) = p \cdot \frac{1}{2} \log_2 \left(\frac{p \cdot \sigma_x^2 \sigma_i^2}{(\sigma_x^2 + \sigma_i^2) D} \right). \quad (58)$$

For the GBG and GE correlation models when $D \leq \sigma_b^2$ the optimal rate allocation results in equal distortions for both fractions of time. Thus the optimal encoder for the GBG model operates on the WZ rate-distortion function $R_{\text{WZ}}^{(b,i)}(D)$ for Gaussian noise of variance $\sigma_b^2 + \sigma_i^2$ during a fraction p of the time, and on the WZ rate distortion function $R_{\text{WZ}}^i(D)$ for the remaining:

$$R_{\text{WZ}}^{\text{GBG}}(D) \geq R_{\text{WZ}_{i_o}}^{\text{GBG}}(D) = \frac{1}{2} \log_2 \left(\frac{\sigma_x^2 \sigma_b^2}{(\sigma_x^2 + \sigma_b^2) D} \right) + p \cdot \frac{1}{2} \log_2 \left(\left(\frac{\sigma_x^2}{\sigma_b^2} + 1 \right) \left(\frac{\sigma_x^2}{(\sigma_b^2 + \sigma_i^2)} + 1 \right)^{-1} \right). \quad (59)$$

Similarly the optimal encoder for the GE correlation model sends full rate for the description of X during a fraction p of the time, while it operates on the Gaussian WZ rate-distortion function $R_{WZ}^{(b)}(D)$ for Gaussian noise of variance σ_b^2 for the remaining time:

$$R_{WZ}^{GE}(D) \geq R_{WZ_{lo}}^{GE}(D) = \frac{1}{2} \log_2 \left(\frac{\sigma_x^2 \sigma_b^2}{(\sigma_x^2 + \sigma_b^2) D} \right) + p \cdot \frac{1}{2} \log_2 \left(\frac{\sigma_x^2}{\sigma_b^2} + 1 \right). \quad (60)$$

For $D > \sigma_b^2$ some rate needs to be spent only where impulses or erasures are present in the correlation, as for the BG model. The minimum rate required is then $R_{WZ_{lo}}^{GBG} = p \cdot R_{WZ}^{(b,i)}(D_i)$ for the GBG model, and $R_{WZ_{lo}}^{GE} = p \cdot R_{WZ}^{(i)}(D_i)$ for the GE model, where D_i has to satisfy the constraint $D = pD_i + (1-p)\sigma_b^2$. For target distortion $D > \sigma_b^2$ the lower bound in figure 28 was determined numerically using reverse water-filling.

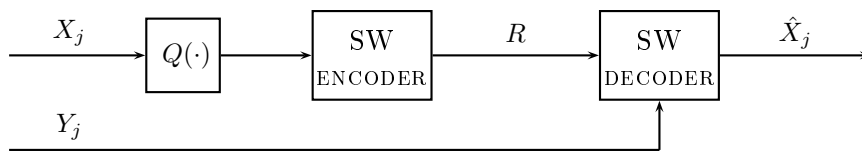


Figure 28: Theoretical scheme.

In order to obtain an upper bound on the rate-distortion function for a given target distortion $D \leq \sigma_b^2$ one may consider the theoretical scheme in Figure 28 and assume high-rate regime. The source is scalar-quantized with step ψ , then Slepian-Wolf encoded [44, Sec. 14.8] and sent to the receiver. Assume ideal Slepian-Wolf encoding scheme with vanishing probability of decoding error. In the BG case the decoder can detect the presence of impulses by comparing the recovered quantized source with the quantized side information. When no impulse is detected the decoder reconstructs $\hat{X} = Y$. The resulting overall distortion is smaller than $D = \psi^2/12$, and the (achievable) upper bound is

$$R_{WZ}^{BG}(D) \leq R_{WZ_{hi}}^{BG}(D) = H(B) + R_{WZ_{lo}}^{BG}(D \cdot p) + p \cdot \frac{1}{2} \log_2 \left(\frac{2\pi e}{12} \right), \quad (61)$$

with $H(B) = p \cdot \log_2(p) + (1-p) \cdot \log_2(1-p)$. Similarly for the GBG model the (achievable) upper bound is

$$R_{WZ}^{GBG}(D) \leq R_{WZ_{hi}}^{GBG}(D) = H(B) + R_{WZ_{lo}}^{GBG}(D) + \frac{1}{2} \log_2 \left(\frac{2\pi e}{12} \right), \quad (62)$$

and for the GE correlation model is

$$R_{WZ}^{GE}(D) \leq R_{WZ_{hi}}^{GE}(D) = R_{WZ_{lo}}^{GE}(D) + \frac{1}{2} \log_2 \left(\frac{2\pi e}{12} \right). \quad (63)$$

In all cases the term $1/2 \log_2(2\pi e/12) = 0.255$ bits is the loss per sample due to scalar quantization. In the GE correlation model it is assumed that the decoder can detect an erasure, whenever it occurs on the correlation channel; in the BG and GBG models, however, the decoder does not know whether the received side information sample is affected by impulsive noise. The term $H(B)$ represents thus the (theoretical) rate per sample needed to estimate the side information quality.

Figure 29 shows upper and lower bounds obtained for the GBG and GE correlation models for $p = 0.05$, $\sigma_x = 1$, $\sigma_g = 0.2$, $\sigma_i = 1$. The figure also shows that for a GBG or a GE correlation model with small p a performance comparable to standard Gaussian correlation may be obtained.

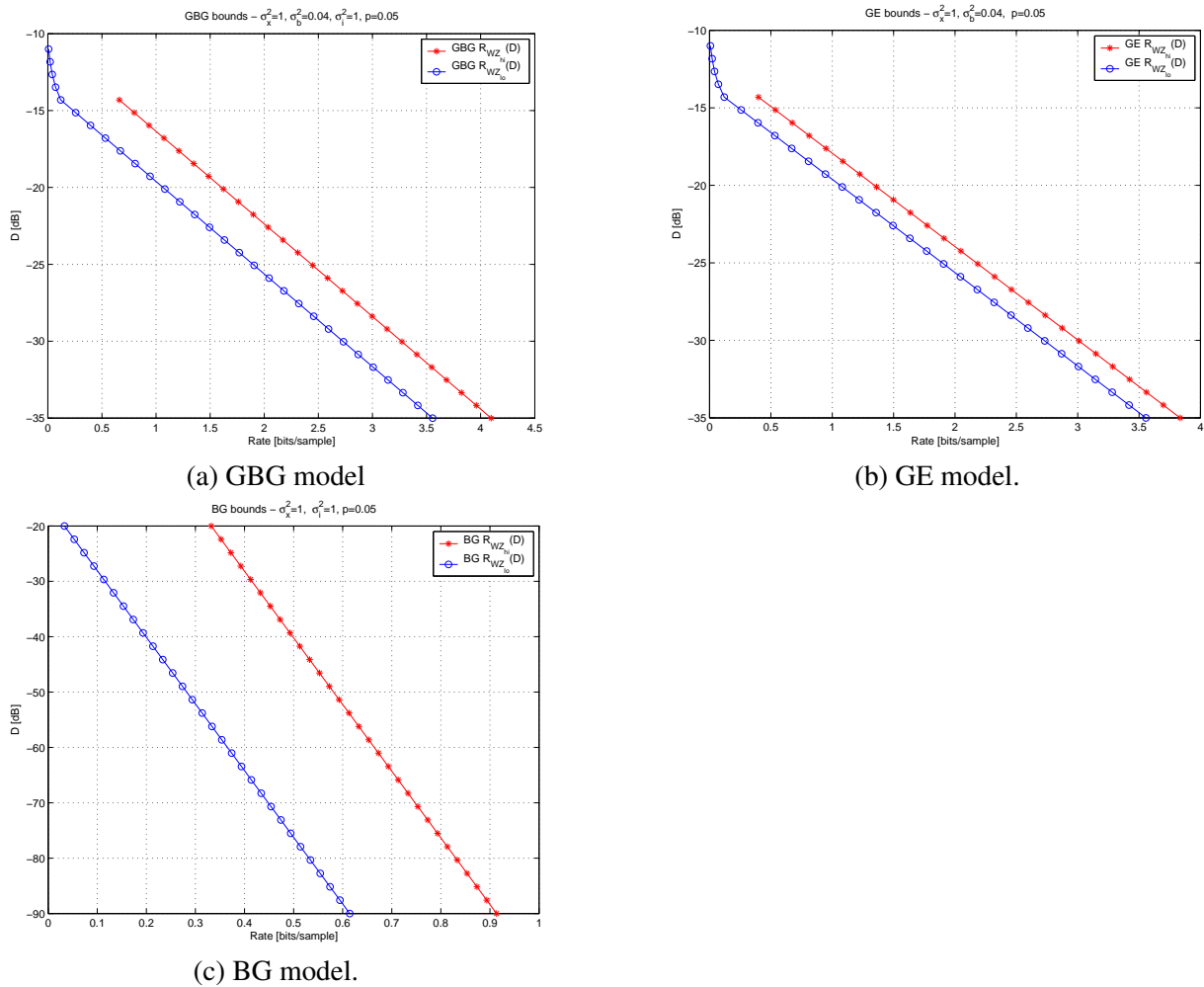


Figure 29: Lower and upper bounds of the rate-distortion functions

5.3 Perspectives

5.3.1 Distance spectra of AC-based and general JSC codes

The fairly complex techniques proposed in [20] to obtain the A_d spectrum will be simplified and generalized to FSMs which do not necessarily correspond to AC implementations. This will allow to extend the method also to sources with memory, as well as finite-state adaptive encoders. New techniques to improve the spectral code properties will be studied. The goal is to obtain a set of bounds and construction guidelines that provide a simple way to design arithmetic-coding based JSCC systems.

5.3.2 Distance properties of distributed arithmetic codes

The paper [66] proposed a novel Slepian-Wolf coding method termed *distributed arithmetic coding*, which encodes a source at a rate smaller than its entropy by allowing the coding intervals to overlap. A soft joint decoder then exploits knowledge of the side information to decode the ambiguous codeword. So far, the design of the interval overlap has not been tackled systematically; therefore, we plan to study the applicability of the distance spectrum techniques outlined above to this problem.

5.3.3 Models and bounds for distributed source coding with sparse correlation structures

The proposed models for distributed source coding with variable-quality side information (see sections 4.2 and 5.2) will be extended to sources and/or correlation with memory. The guiding principle for

extending these models will be their appropriateness to applications such as distributed video coding.

5.3.4 *Distributed joint source channel coding: comparison of source coding (quasi-arithmetic) and channel coding based methods*

The positioning of the sequence length constraint leading to the best decoding performance has been studied in [124] for quasi-arithmetic codes. It is shown that the optimum position of the length constraint is not the last time instant of the decoding process. The best position actually depends on the channel signal to noise ratio and on the error recovery properties of the code. This observation led to the introduction of a new approach to add redundancy to QA codes in order to improve their decoding performance. The redundancy is added under the form of length constraints exploited at different time instants of the decoding process on the aggregate state model. In comparison with other methods based on side information such as markers which help the resynchronization of the decoding process, the strategy proposed here does not lead to any modification of the compressed bit-stream. The side information can be transmitted separately from the compressed stream. The approach turns out to outperform widely used techniques based on the introduction of a forbidden symbol.

We will then pursue this study by carrying out a comparative assessment between a JSCC strategy based on quasi-arithmetic codes with redundancy under the form of length constraints exploited at different time instants of the decoding process, and a JSCC strategy based on turbo codes.

5.3.5 *Analysis of joint source channel coding: analysis of the puncturing of quasi-arithmetic coding*

Variable length coding offers great performance in terms of compression, but it is however very sensitive to channel noise. Indeed, a single error in the bit-stream may result in the de-synchronization of the decoder leading to dramatic symbol error rates (SERs). Synchronization recovery properties of variable length codes (VLC) have been first studied in [127]. A model to derive the error span following a single bit error (i.e. the expected number of symbols on which the error propagates inside the bit-stream) is proposed. It can be shown that VLC strictly satisfying the Kraft inequality statistically resynchronize with a probability of 1. However, they do not always resynchronize in the strict sense (i.e. the symbol length of the decoded bit-stream may differ from the one of the encoded bit-stream). In [181], the authors have extended the model of [127] to calculate the probability mass function (PMF) of the so-called “gain/loss”. The “gain/loss” represents the difference between the number of encoded and decoded symbols. It is shown in [227], that the error span reflects the performance of a VLC with hard decoding. In [123], the measure of the “gain/loss” is shown to reflect the performance of VLC when soft-input decoding with a termination constraint on the number of encoded symbols is applied at the decoder side, and similarly for quasi-arithmetic codes [124]. The term termination constraint denotes here a constraint that enforces the symbol length of the decoded sequence to be equal to a given number of symbols (known by the decoder).

The error recovery capability of VLC and quasi-arithmetic codes can thus be predicted with the help of a so-called error state diagram. Transfer functions defined on this error state diagram allow us to compute the expected error span (i.e., the average number of symbols on which the bit error propagates) and the gain/loss (i.e., the difference between the number of encoded and decoded symbols) following a single bit error. The computation of these quantities can be extended to the case where the bit-stream is transmitted over a BSC. The entropy of the gain/loss turns out to be a powerful tool to analyze the performance of soft-input decoding of VLC and QA codes with termination constraints. It allows us for example to quantify the amount of information brought by a length constraint on both optimal and aggregated state models. For a given code and a given channel signal to noise ratio, the entropy of the gain/loss allows us for example to determine the value of the aggregation parameter for which close-to-optimum decoding performance will be obtained.

The goal will be to extend the tools and methodology to the case of an erasure channel.

6 TASK TR7.5: STUDY OF PRACTICAL, LONG-TERM SOLUTIONS PLUS TEST-BEDS

This section summarizes the status of JSCC/D with respect to practical applications. First, a few situations, sometimes dating many years ago, are recalled, where Shannon's separation principle was not really applied. Then, it is shown that the concept already serves as a publicity for a few products, and finally, we aim at selecting some situations in which JSCC/D could really have some impact in a near future, even if many points remain to be studied.

6.1 Actual impact of JSCC/D in standards

6.1.1 *Hidden cases where separation theorem is not really used*

The separation theorem states that an optimal situation (in terms of global rate) is attained **(i)** if the signal to be transmitted is compressed at a maximum, for a given admissible distortion, while **(ii)** the transmission is protected by a very powerful channel coder, so that the error probability is almost zero. In such a case, all distortion is introduced by the source encoder, while no transmission error is admitted.

However, wireless channels have a wide variability, and some communications may encounter very poor channel situations. Hence, it has soon been recognized that it was not possible to tune any transmission system for the worst possible channel, and that some mechanisms were needed to allow better transmission rates. Such mechanisms can be found at all layers of the communication chain. On the transmission side, the mechanisms that are strongly linked to the variability of channels are

- Adaptive modulation and coding, allowing to adapt the transmitted bit rate to the quantity the channel may carry
- All mechanisms making use of diversity, which are using the variability of the channels to improve the communication

On the network side

- Packetization, which prevents an error to propagate too much
- Acknowledgement procedures, allowing re-transmission when a packet is erroneous

On the source coding side

- Natural error resilience of source coders
- Error concealment

the last two items are practically used in cases where the separation theorem is clearly not valid : In GSM, for example, the bit error rate after channel coding/decoding has been tuned not to exceed 10^{-3} because speech coders were still understandable after going through a channel with this BER. Moreover, error concealment is an important part of the recent video CODECS, and allows to play an H264 encoded video without too much damage even when few blocks are lost. Finally, even the notions of **(i)** Unequal Error Protection (UEP), **(ii)** of using various Qualities of Service for various types of signals, **(iii)** of scalable video bitstreams (to be used jointly with UEP) also shows that communication systems do not stick to a strict interpretation of the separation theorem.

The list above does not intend to be exhaustive, and rather tries to demonstrate that many mechanisms already exist, and that the proposed tools have to be more efficient than those already existing. In various cases they should also be compatible with the use of some of them. This motivates the choice of some situations of interest which could serve as testbeds. The intention here is that the testbeds could serve altogether to an evaluation of the performance as well as to the emergence of new problems coming from a more global vision of the multimedia communication paradigm.

6.1.2 Recent cases where JSCC/D has contributed to standards

Even if JSCC/D does not appear explicitly in standards, there are a few cases where there is a strong connection with options available in some of them.

1. Error correcting arithmetic codes [65], contributed by CERCOM (member of CNIT) to JPEG2000 part 11 (JPWL), [81]
2. DVB-H [53] where hierarchical modulation has been normalized (I do not think that any practical use has been proposed yet, but it is one option of the standard).
3. Redundant slices have been proposed in conjunction with Flexible Macroblock Ordering in H264 [84]: it is actually a very simple spatial Multiple Description Coding technique, and is an explicit insertion of redundancy at the source coding level.

These few examples are by no means (with the exception of the first one) pure JSCC/D techniques. However, they are getting closer and closer to what is studied under this acronym.

6.2 Emergence of the concept in products

Recent announcements on products based on JSCC can be found. You will find below an extract of one of them (AMIMON Ltd.)

Amimon Ships Chip Set for Short Range, High Data Rate, Wireless HDTV

TV stations usually concentrate on transmitting HDTV content to distant locations using as little data bandwidth as possible while maintaining quality. Amimon's new WHDI chip set transmits HDTV to nearby devices using as much data bandwidth as possible.

By using Joint Source-Channel Coding (JSCC) instead of conventional data transmission methods that treat each bit as equally important, Amimon's WHDI system is able to transmit data rates between 250 and 800 Mbps using a 20 MHz wide 5 GHz channel over short distances inside a house. 8-bit per color HDTV samples can be compressed to about half their original data rate, 1.5 Gbps for 720p and 1080i HDTV or 3.0 Gbps for 1080p HDTV. On a good 20 MHz wide channel (or two for 1080p), Amimon's WHDI system is able to transmit a near perfect picture.

What happens when the channel is degraded and the data rate drops below half the uncompressed rate? That's where JSCC makes the difference. Video components are prioritized according to their importance. For example, most significant bits are more important than least significant bits, lower spatial frequencies are more important than higher frequencies and the luminance component is more important than the chrominance components. Higher priority components receive more error correction than less important components. High priority components are also transmitted using a less complex modulation method—a coarser constellation, for example, transmitted on frequencies with less noise. As the channel SNR deteriorates, lower priority components are lost first.

This does not mean that the corresponding product is really making use of all the power of JSCC/D (this seems more like Unequal Error Protection) but it does mean that the idea progresses, and that it may become a discriminating factor for new innovative products, just like MIMO has been a few years ago for Wireless LAN modems (even if in some cases, MIMO was mostly a selling argument...)

6.3 Possible impact in the near future depending on the context

This section intends to select a situation in which JSCC/D could be checked against traditional solutions by taking the whole context into account, including the network layers and a realistic model of the Physical layer. Obviously, the constraints can be very different in various settings, and this choice is very important.

A first, general, consideration is that source coders are so efficient that getting rid of them and working on the original signal or designing a specific encoder which would not involve image prediction nor Variable Length Codes would not be realistic, since they would result in a much larger bit stream. However,

it is obviously possible to tune them in a different way, for example in order to leave some redundancy at the cost of a slight to moderate increase of the bit rate. This consideration seems to apply to all situations described below.

6.3.1 Point to point communications

Considering point to point transmission, one is usually in a communication situation, in which interaction is important, and the delay is a very strong constraint. In fact, when the delay is larger than 0.5 second the interaction between the two parties becomes very difficult. This delay constraint also prevents the ACK/NACK procedures to be applied. However, the classical solutions that are used (source coding using block prediction, etc... strong channel codes,...) already come close to the limit in terms of delay. Hence, any modification introduced by JSCC should not increase the delay, and the situation becomes very difficult. It does not mean that JSCC/D would not be able to improve the situation, but point to point transmission is not the preferred case...

6.3.2 Broadcasting

In a broadcasting context, the situations at first sight seems to be more comfortable : the transmitter addresses many terminals, and the communication is (mainly) unidirectional. Hence, the delay constraints seems to vanish, and the ACK/NACK procedure is not feasible, since the number of terminals prevents the transmitter to wait for all of them.

This has to be somewhat balanced by more practical constraints, limiting for example the delay to an order of magnitude of 1 sec, for example in TV broadcasting, since the "zapping" would become unrealistically uncomfortable. The ACK/NACK procedure at block level can also be mimicked by introducing redundancy at packet level (such as MPE-FEC in DVB-H [53]).

Altogether, this seems to be the easiest case, and will be addressed as a first setting. However, the existence of the above mentioned constraints will be taken into account, since they will state new problems.

6.3.3 Multicasting

A slightly more complex situation is "multicasting". by multicasting here, we mean a service in which a TV show, for example, would be delivered to a selected set of receivers in a given cell. This number of receivers being small to moderate. This situation is intermediate between broadcasting (very close to it) and point to point communication. The advantage is that the delay constraint is not too strong, and that it is essentially a one way communication, even if a return channel exists. In this case, some kind of ACK/NACK procedure can be used, but needs to be re-defined to be compatible with JSCC/D.

6.4 Longer term situation : Video streaming over peer to peer networks

Finally, one situation seems to be of an increasing interest : video streaming through peer to peer networks. Obviously, this application is much more demanding, and has been addressed currently by tools that are closer to cross-layer design in a traditional sense than to real JSCC/D. However, it is possible (likely) that the techniques that are studied in this task could prove to be useful in this context. This is the first step that we will be studying during the course of NEWCOM++.

A very good research group of Stanford, led by Andrea Goldsmith and Bernd Girod is already addressing this problem. (see the link http://www.stanford.edu/~zhuxq/adhoc_project/adhoc_project.html) They selected the following topics as being important for solving this issue:

1. Adaptive link layer techniques
2. Joint capacity and flow assignment
3. Distributed search for capacity region

4. Congestion-minimized stream routing
5. Intelligent packet scheduling
6. Congestion-aware and delay-constrained rate allocation
7. Distributed rate allocation over heterogeneous networks
8. Distributed routing for heterogeneous data
9. Video distortion modeling
10. Rate allocation among multiple streams
11. Joint optimization of source/channel coding
12. Error-resilient source coding
13. Peer-to-peer multicast video streaming

While some of the topics are clearly in the PHY layer (the first three), some of them are clearly in the network domain (items 4-8), the last ones (items 9-13) could clearly benefit from a JSCC/D approach. However, the topic must be addressed jointly with the networking layers to be relevant. During the course of NEWCOM++, the feasibility of this study will be addressed within deliverable DR7.3. If the conclusions are positive (and this will be discussed with researchers from the networking are, in WPR11), the topic will begin to be addressed during the last year of NEWCOM++, together with them.

6.5 Testbeds

During the course of NEWCOM++, in order to facilitate the comparison of various techniques proposed by the various teams, and to promote exchange of ideas and software building blocks, two software simulators, based on IT++ and OMNET++ will be made available. The first one is concerned with video broadcasting based on DVB H, corresponding to section 6.3.2 and the other one to a multicast situation based on WIMAX (corresponding to section 6.3.3). These simulators will not cover the whole standards, (e.g. will suppose that the WIMAX communication is already established) but will anyway allow to evaluate the efficiency of our algorithms in realistic contexts.

7 CONCLUSION

This deliverable has described the status of the work in the various teams contributing to this workpackage, with respect to the global state of the art in the domain of Joint Source and Channel Coding/Decoding. The specificity of this work is that, altogether, this group is contributing on theoretical aspects of the domain, as well as on more practical ones. In fact, some of the results described as "state of the art" is already the result of contributions between various partners in the framework of the former NoE NEWCOM.

This document also describes section by section, the joint work that will be worked out in the next period.

REFERENCES

- [1] Minimum expected distortion in gaussian layered broadcast coding with successive refinement. In *Proc. IEEE International Symposium on Information Theory*, June 2007.
- [2] 3GPP TS 26.346 V6.1.0. *Technical specification group services and system aspects; multimedia broadcast/multicast service; protocols and codecs*, Jun 2005.
- [3] R. A and D. Gesbert. Optimal matching in wireless sensor networks. *IEEE Journal on Selected Topics in Signal Processing, Special Issue on Convex Optimization Methods in Signal Processing*, 1(4):725–735, December 2007.
- [4] A. Aaron and B. Girod. Compression with side information using turbo codes. *Data Compression Conference*, pages 252–261, Apr. 2002.
- [5] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, July 2000.
- [6] J. B. Anderson and S. Mohan. *Source and Channel Coding: An Algorithmic Approach*. Kluwer, 1991.
- [7] ANSI/IEEE. 802.11, part 11 : Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. Technical report, 1999.
- [8] X. Artigas, S. Malinowski, C. Guillemot, and L. Torres. Overlapped quasi-arithmetic codes for distributed video coding. In *Proceedings of IEEE ICIP*, 2007.
- [9] E. Baccaglioni, G. Barrenetxea, and B. Beferull-Lozano. Interaction between multiple description coding and sensor networks with finite buffers. In *Proc. of IEEE International Conference on Multimedia and Expo*, pages 1460–1463, Jul 2005.
- [10] E. Baccaglioni, T. Tillo, and G. Olmo. Network adaptive multiple description coding for JPEG2000. In *Proc. of IEEE International Conference on Image Processing*, volume 3, pages 936 – 939, Sept. 2005.
- [11] E. Baccaglioni, T. Tillo, and G. Olmo. A flexible r-d based multiple description scheme for JPEG 2000. *IEEE Signal Processing Letters.*, Mar. 2007.
- [12] G. Barrenechea, B. Beferull-Lozano, V. Abhishek, P. L. Dragotti, and M. Vetterli. Multiple description source coding and diversity routing: A joint source channel coding approach to real-time services over dense networks. *International Packet Video Workshop*, 2003.
- [13] J. Barros and S. D. Servetto. The sensor reachback problem. *IEEE Transactions on Information Theory*, Nov. 2003.
- [14] J. Barros and S. D. Servetto. Network information flow with correlated sources. *IEEE Transactions on Information Theory*, Oct. 2005.
- [15] J. Barros and S. D. Servetto. Network information flow with correlated sources. *IEEE Transactions on Information Theory*, 52(1):155–170, Jan. 2006.
- [16] F. Bassi, C. Guillemot, M. Kieffer, A. Roumy, and V. Toto-Zarasoia. Dr 1.1: State of art techniques in distributed source coding. Technical report, 2007.
- [17] F. Bassi, M. Kieffer, and C. Weidmann. Source coding with intermittent and degraded side information at the decoder. In *Proc. ICASSP*, pages 2941–2944, April 2008.

- [18] R. Bauer and J. Hagenauer. On variable length codes for iterative source/channel decoding. In *Proc. IEEE Data Compression Conference*, pages 272–282, Snowbird, UT, 1998.
- [19] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *Annals of Mathematical Statistics*, 41(1):164–171, 1970.
- [20] S. Ben-Jamaa, C. Weidmann, and M. Kieffer. Analytical tools for optimizing the error correction performance of arithmetic codes. *IEEE Trans. Commun.*, 2008. to appear.
- [21] T. Berger. *Multiterminal source coding*, volume 229 of *The Information Theory Approach to Communications*. 1977.
- [22] T. Y. Berger-Wolf and M. A. Harris. Sharp bounds for bandwidth of clique products. Submitted to *SIAM Journal of Discrete Mathematics*, October 2002.
- [23] C. Bergeron and C. Lamy-Bergot. Soft-input decoding of variable-length codes applied to the H.264 standard. In *Proc. IEEE 6th Workshop on Multimedia Signal Processing*, pages 87–90, 29 Sept.-1 Oct. 2004.
- [24] C. Bergeron and C. Lamy-Bergot. Modelling h.264/avc sensitivity for error protection in wireless transmissions. In *Proc. of IEEE MMSP*, Oct 2006.
- [25] K. Bhattad, K. R. Narayanan, and G. Caire. On the distortion snr exponent of some layered transmission. *IEEE Trans. on Information Theory* (submitted - available at <http://arxiv.org/abs/cs/0703035>), 2008.
- [26] R. E. Blahut. *Theory and Practice of Error Control Codes*. Addison-Wesley, Reading, MA, 1984.
- [27] M. Bogino, P. Cataldi, M. Grangetto, E. Magli, and G. Olmo. Sliding-window digital fountain codes for streaming of multimedia contents. In *IEEE International Symposium on Circuits and Systems (ISCAS 2007)*, pages 3467 – 3470, May 2007.
- [28] C. Bormann, C. Burmeister, M. Degermark, H. Fukushima, H. Hannu, L.-E. Jonsson, R. Hakenberg, T. Koren, K. Le, Z. Liu, A. Martensson, A. Miyazaki, K. Svanbro, T. Wiebke, T. Yoshimura, and H. Zheng. Robust header compression (ROHC): Framework and four profiles. Technical Report RFC 3095, 2001.
- [29] A. Bouabdallah, M. Kieffer, J. Lacan, G. Sabeva, G. Scot, C. Bazile, and P. Duhamel. Evaluation of cross-layer reliability mechanisms for satellite digital multimedia broadcast. *IEEE Transactions on Broadcasting*, 53(1):391–404, 2007.
- [30] D. Burshtein and G. Miller. An efficient maximum likelihood decoding of LDPC codes over the binary erasure channel. *IEEE Trans. Inform. Theory*, 50(11), Nov. 2004.
- [31] V. Buttigieg and P. Farrell. A MAP decoding algorithm for variable-length error-correcting codes. In *Codes and Cyphers: Cryptography and Coding IV*, pages 103–119, Essex, England, 1995. The Inst. of Mathematics and its Appl.
- [32] V. Buttigieg and P. Farrell. Variable-length error-correcting codes. *IEE Proceedings on Communications*, 147(4):211–215, Aug. 2000.
- [33] J. Byers, M. Luby, and M. Mitzenmacher. A digital fountain approach to reliable distribution of bulk data. *IEEE J. Select. Areas Commun.*, 20(8):1528–1540, Oct. 2002.
- [34] J. B. Cain, G. C. Clark Jr., and J. M. Geist. Punctured convolutional codes of rate $(n - 1)/n$ and simplified maximum likelihood decoding. *IEEE trans. Information Theory*, 25(1):97–100, 1979.

- [35] G. Caire and K. Narayanan. On the snr exponent of hybrid digital-analog space time coding. *IEEE Trans. on Information Theory*, 53(8):2867–2878, Aug. 2007.
- [36] G. Caire and K. Narayanan. On the snr exponent of hybrid digital-analog space time coding. page 2005, Sept. Proc. Allerton Conf. Commun., Contr. and Comp.
- [37] S. Choi and S. Pradhan. A graph-based framework for transmission of correlated sources over broadcast channels. *IEEE Transactions on Information Theory*, 2006.
- [38] T. Coleman, E. Martinian, and E. Ordentlich. Joint source-channel decoding for transmitting correlated sources over broadcast networks. *IEEE International Symposium on Information Theory*, pages 2144–2147, July 2006.
- [39] T. Cover. Broadcast channels. *IEEE Trans. on Information Theory*, 18(1):2–14, Jan. 1972.
- [40] T. Cover. Broadcast channels. *IEEE Transactions on Information Theory*, 18(1):2–14, Jan. 1972.
- [41] T. Cover and A. Gamal. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory*, 25(5):572–584, Sept. 1979.
- [42] T. Cover, A. Gamal, and M. Salehi. Multiple access channels with arbitrarily correlated sources. *IEEE Transactions on Information Theory*, 26(6):648–657, Nov. 1980.
- [43] T. Cover and J. Thomas. *Elements of Information Theory*. 1991.
- [44] T. Cover and J. Thomas. *Elements of Information Theory*. 1991.
- [45] R. Dabora and S. D. Servetto. Broadcast channels with cooperating decoders. *IEEE Transactions on Information Theory*, 52(12):5438–5454, Dec. 2006.
- [46] A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros. Capacity of wireless erasure networks. *IEEE Transactions on Information Theory*, 52(3):789–804, Mar. 2006.
- [47] S. N. Diggavi and V. A. Vaishampayan. On multiple description source coding with decoder side information. *IEEE Information Theory Workshop*, Oct. 2004.
- [48] Digital Video Broadcasting (DVB). *Framing structure, channel coding and modulation for Satellite Services to Handheld devices (SH) below 3GHz*, 2007. Blue Book.
- [49] A. Dimakis, J. Wang, and K. Ramchandran. Unequal growth codes: Intermediate performance and unequal error protection for video streaming. In *IEEE 9th Workshop on Multimedia Signal Processing, 2007 (MMSP 2007)*, pages 107 – 110, Oct. 2007.
- [50] C. Doe-Man Chung and W. Yao. Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms. *IEEE Trans. on Circuits and Systems for Video Technology*, 9(6):895–908, Sept. 1999.
- [51] B. Dongsheng, W. Hoffman, and K. Sayood. State machine interpretation of arithmetic codes for joint source and channel coding. *Proc. of DCC, Snowbird, Utah, USA.*, pages 143–152, 2006.
- [52] F. Etemadi and H. Jafarkhani. Optimal layered transmission over quasi-static fading channels. In *Proc. IEEE Symposium Information Theory*, Jul 2006.
- [53] ETSI. Digital video broadcasting (DVB); transmission system for handheld terminals (DVB-h). Technical report, ETSI EN 302 304 v1.1.1, nov. 2004.
- [54] M. Fleming and M. Effros. Rate-distortion with mixed types of side information. *IEEE Trans. Inform. Theory*, 52(4):1698–1705, April 2006.

- [55] A. Gabay, M. Kieffer, and P. Duhamel. Joint source-channel coding using real bch codes for robust image transmission. *IEEE Transactions on Image Processing*, december 2005. Submitted.
- [56] J. Garcia-Frias and F. Cabarcas. Approaching the Slepian–Wolf boundary using practical channel codes. *Signal Processing*, 86(11):3096–3101, 2006.
- [57] J. Garcia-Frias and Y. Zhao. Compression of correlated binary sources using turbo codes. *IEEE Communications Letters*, 5(10):417–419, Oct. 2001.
- [58] M. Gastpar and M. Vetterli. Source–channel communication in sensor networks. *Lecture Notes in Computer Science*, 2634, 2003.
- [59] M. Gastpar and M. Vetterli. On the capacity of large gaussian relay networks. *IEEE Transactions on Information Theory*, 51(3):765–779, Mar. 2005.
- [60] N. Gehrig and P. Dragotti. Symmetric and asymmetric slepian-wolf codes with systematic and non-systematic linear codes. *IEEE Communications Letters*, 9(1):61–63, Jan. 2005.
- [61] I. M. Gelfand and S. V. Fomin. *Calculus of variations*. 1962.
- [62] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero. Distributed video coding. *Proceedings of the IEEE*, 93(1):71–3, Jan. 2005.
- [63] V. K. Goyal and J. Kovacević. Generalized multiple description coding with correlating transforms. *IEEE Trans. on Information Theory*, 47(6):2199–2224, Sept. 2001.
- [64] V. K. Goyal, J. Kovacevic, and J. Kelner. Quantized frame expansions with erasures. *Journal of Appl. and Comput. Harmonic Analysis*, 10(3):203–233, 2001.
- [65] M. Grangetto, P. Cosman, and G. Olmo. Joint source/channel coding and map decoding of arithmetic codes. *IEEE Transactions on Communications*, 53(6):1007–1016, June 2005.
- [66] M. Grangetto, E. Magli, and G. Olmo. Distributed arithmetic coding. *IEEE Communications Letters*, 11(11):883–885, Nov 2007.
- [67] M. Grangetto, E. Magli, and G. Olmo. Symmetric distributed arithmetic coding of correlated sources. In *Proceedings of IEEE MMSP*, 2007.
- [68] M. Grangetto, E. Magli, and G. Olmo. Distributed arithmetic coding for the asymmetric and symmetric Slepian-Wolf problems. *IEEE Transactions on Signal Processing*, 2008. preprint available at URL: www1.tlc.polito.it/sas-ipl/Magli/index.php.
- [69] M. Grangetto, E. Magli, R. Tron, and G. Olmo. Rate-compatible distributed arithmetic coding. *IEEE Communications Letters*, 2008. preprint available at URL: www1.tlc.polito.it/sas-ipl/Magli/index.php.
- [70] T. Guionnet and C. Guillemot. Soft decoding and synchronization of arithmetic codes: Application to image transmission over noisy channels. *IEEE Trans. on Image Processing*, 12(12):1599–1609, 2003.
- [71] D. Gunduz and E. Erkip. Reliable cooperative source transmission with side information. *2007 IEEE Information Theory Workshop on Information Theory for Wireless Networks*, pages 1–5, July 2007.
- [72] D. Gunduz, C. Ng, E. Erkip, and A. Goldsmith. Source transmission over relay channel with correlated relay side information. *IEEE International Symposium on Information Theory*, June 2007.

- [73] H.264/AVC Software Coordination. Joint model software. available in <http://iphome.hhi.de/suehring/tml/>.
- [74] J. Hagenauer. Source-controlled channel decoding. *IEEE trans. on Communications*, 43(9):2449–2457, 1995.
- [75] R. Hamzaoui, V. Stankovic, and Z. Xiong. Optimized error protection of scalable image bit streams [advances in joint source-channel coding for images]. *IEEE Signal Processing Magazine*, 22(6):91–107, November 2005.
- [76] T. Han and M. Costa. Broadcast channels with arbitrarily correlated sources. *IEEE Transactions on Information Theory*, 33(5):641–650, Sept. 1987.
- [77] M. C. Hong, H. Schwab, L. P. Kondi, and A. K. Katsaggelos. Error concealment algorithms for compressed video. *Signal Processing: Image Communication*, 14:473–492, 1999.
- [78] P. Howard and J. Vitter. Practical implementations of arithmetic coding. In *Image and Text Compression*. 1992.
- [79] P. G. Howard and J. S. Vitter. Practical implementations of arithmetic coding. *Image and Text Compression*, 13(7):85–112, 1992.
- [80] R. Hu, R. Viswanathan, and J. Li. A new coding scheme for the noisy-channel Slepian–Wolf problem: separate design and joint decoding. *Global Telecommunications Conference*, 1:51–55, Nov. 2004.
- [81] ISO/IEC. JPEG 2000 part 11: JPEG 2000 image coding system: JPWL - wireless (ISO/IEC 15444-11). Technical report, ISO/IEC, 2007.
- [82] ITU. Recommendation IT-81, information technology - digital compression and coding of continuous-tone still images - requirements and guidelines (JPEG). Technical report, ITU - CCITT, 1992.
- [83] ITU-T. H263 - video coding for low bitrate communications. Technical report, ITU-T Rec. H.263 (02/98), 1998.
- [84] ITU-T and ISO/IEC JTC 1. Advanced video coding for generic audiovisual services. Technical report, ITU-T Rec. H.264, and ISO/IEC 14496-10 AVC, nov. 2003.
- [85] ITU-T/SG16/VCEG, ISO/JTC1/SC29/WG11. *H264 Joint Committee Draft, Document JVT-C167*. 3rd meeting: Fairfax, Virginia, USA, 6-10 may 2002.
- [86] A. Jagmohan, A. Sehgal, and N. Ahuja. Two-channel predictive multiple description coding. In *Proc. of IEEE International Conference on Image Processing*, volume 2, pages 670–673, Sept 2005.
- [87] X. Jaspard and L. Vandendorpe. Design and performance analysis of joint source-channel turbo schemes with variable length codes. In *Proc. of ICC*, Seoul, 2005.
- [88] H. Jégou, S. Malinowski, and C. Guillemot. Trellis stage aggregation for soft decoding of variable-length codes. In *Proceedings of SPIS*, 2005.
- [89] H. Jenkac, T. Stockhammer, and W. Xu. Permeable-layer receiver for reliable multicast transmission in wireless systems. In *Proc. IEEE Wireless Communications and Networking Conference*, volume 3, pages 1805–1811, 13-17 March 2005.

- [90] H. Jenkac, T. Stockhammer, and W. Xu. Cross-Layer assisted reliability design for wireless multimedia broadcast. *EURASIP Signal Processing Journal, Special Issue on Advances in Signal Processing-assisted Cross-layer Designs*, 86(8):1933–1949, 2006.
- [91] W. Jiang and A. Ortega. Multiple description coding via polyphase transform and selective quantization. In *Proc. SPIE Intl. Conference on Visual Communication and Image Processing*, pages 998–1008, Jan. 1999.
- [92] H. Jin, A. Khandekar, and R. McEliece. Irregular repeat–accumulate codes. *Proceedings of the 2nd International Symposium on Turbo codes and Related Topics*, pages 1–8, 2000.
- [93] M. Kac. On the notion of recurrence in discrete stochastic processes. *Bulletin of the American Mathematical Society*, 53:1002–1010, 1947.
- [94] S. Kaiser and M. Bystrom. Soft decoding of variable-length codes. In *Proc. IEEE ICC*, volume 3, pages 1203–1207, New Orleans, 2000.
- [95] M. Kalman and B. Girod. Rate-distortion optimized video streaming using conditional packet delay distributions. In *IEEE MMSP*, Siena, Italy, Sep 2004.
- [96] S. Kang and A. Zakhor. Packet scheduling algorithm for wireless video streaming. In *Proc. of 12th Packetvideo Workshop 2002*, Pittsburgh PA, USA, 2002.
- [97] S. Kanumuri, P. Cosman, A. Reibman, and V. Vayshampayan. Modeling packet-loss visibility in mpeg-2 video. *IEEE Trans. on Multimedia*, 8(2), Apr 2006.
- [98] W. Karner, O. Nemethova, and M. Rupp. Link error prediction based cross-layer scheduling for video streaming over umts. In *15th IST Mobile & Wireless Communications Summit*, Myconos, Greece, Jun 2006.
- [99] W. Karner, O. Nemethova, and M. Rupp. Link error prediction in wireless communication systems with quality based power control. In *IEEE Int. Conf. on Communications (ICC)*, Glasgow, UK, June 2007.
- [100] W. Karner and M. Rupp. Measurement based analysis and modeling of umts dch error characteristics for static scenarios. In *Proc. of 8th Int. Symp. on DSP and Comm. Systems (DSPCS)*, Sunshine Coast, Australia, Dec 2005.
- [101] S. Katti, H. Rahul, W. Hu, D. Katabi, M. M'edard, and J. Crowcroft. Xors in the air: practical wireless network coding. In *SIGCOMM Computation Communication Review*, volume 36, pages 243–254, 2006.
- [102] N. Koshlev. *Hierarchical coding of discrete sources*. 1980.
- [103] P. Koutsakis. Scheduling and call admission control for burst-error wireless channels. In *10th IEEE Symp. on Computers and Communications (ISCC)*, 2005.
- [104] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, Sept. 2005.
- [105] W.-Y. KUNG, C.-S. KIM, and C.-C. J. KUO. Spatial and temporal error concealment techniques for video transmission over noisy channels. *IEEE transactions on circuits and systems for video technology*, 16:789–802, 2006.
- [106] J. F. Kurose and K. W. Ross. *Computer Networking: A Top-Down Approach Featuring the Internet*. Addison Wesley, Boston, third edition, 2005.

- [107] H. Kushwaha, Y. Xing, R. Chandramouli, and H. Heffes. Reliable multimedia transmission over cognitive radio networks using fountain codes. *Proceedings of the IEEE*, 96(1):155 – 165, Jan. 2008.
- [108] L. Lai, K. Liu, and H. El Gamal. The three–node wireless network: achievable rates and cooperation strategies. *IEEE Transactions on Information Theory*, 52(3):805–828, Mar. 2006.
- [109] C. Lamy and O. Pothier. Reduced complexity maximum a posteriori decoding of variable-length codes. In *Proceedings of Globecom'01*, pages II/1410–II/1413, november 2001.
- [110] L. A. Larzon, M. Degermark, L. E. Jonsson, and G. Fairhurst. The lightweight user datagram protocol (UDP-Lite). Technical Report RFC 3828, The Internet Society, 2004.
- [111] C. Lee, M. Kieffer, and P. Duhamel. Soft decoding of vlc encoded data for robust transmission of packetized video. In *Proceedings of ICASSP*, pages 737–740, 2005.
- [112] C. Lee, J. Kim, Y. Altunbasak, and R. M. Mersereau. Layered coded vs. multiple description coded video over error-prone networks. *Signal Processing: Image Communication*, May 2003.
- [113] Y. Liang and V. Veeravalli. Cooperative relay broadcast channels. *IEEE Transactions on Information Theory*, 53(3):900–928, Mar. 2007.
- [114] S. Lin and D. J. Costello. *Error Control Coding: Fundamentals and applications*. Prentice-Hall, Englewood Cliffs, 1983.
- [115] Z. Liu, S. Cheng, A. Liveris, and Z. Xiong. Slepian–Wolf coded nested lattice quantization for Wyner–Ziv coding: High-rate performance analysis and code design. *IEEE Transactions on Information Theory*, 52(10):4358–4379, Oct. 2006.
- [116] Z. Liu, S. Cheng, A. Liveris, and Z. Xiong. Slepian-wolf coded nested lattice quantization for wyner-ziv coding: High-rate performance analysis and code design. *Information Theory, IEEE Transactions on*, 52(10):4358–4379, Oct. 2006.
- [117] A. Liveris, Z. Xiong, and C. Georghiades. A distributed source coding technique for correlated images using turbo–codes. *IEEE Communications Letters*, 6(9):379–381, Sept. 2002.
- [118] A. Liveris, Z. Xiong, and C. Georghiades. Compression of binary sources with side information at the decoder using LDPC codes. *IEEE Communications Letters*, 6(10):440–442, Oct. 2002.
- [119] A. Liveris, Z. Xiong, and C. Georghiades. Joint source–channel coding of binary sources with side information at the decoder using ira codes. *IEEE Workshop on Multimedia Signal Processing*, pages 53–56, Dec. 2002.
- [120] M. Luby. LT codes. In *Proc. of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 271–282, Vancouver, Canada, Nov. 2002.
- [121] M. Luby, M. Watson, T. Gasiba, T. Stockhammer, and W. Xu. Raptor codes for reliable download delivery in wireless broadcast systems. In *Proc. of 2006 IEEE Consumer Communications and Networking Conf.*, volume 1, pages 192–197, Jan 2006.
- [122] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, 2003.
- [123] S. Malinowski, H. Jégou, and C. Guillemot. Synchronization recovery and state model reduction for soft decoding of variable length codes. *IEEE Transactions on Information Theory*, 53(1):368–377, Jan. 2007.

- [124] S. Malinowski, H. Jégou, and C. Guillemot. Error recovery and new state models for quasi-arithmetic codes. *Eurasip journal on applied signal processing*, 2008.
- [125] C. Marin, P. Duhamel, K. Bouchireb, and M. Kieffer. Robust video decoding through simultaneous usage of residual source information and mac layer crc redundancy. In *Proceedings of Globecom 07*, pages 2070–2074, 2007. Submitted.
- [126] K. Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Transactions on Information Theory*, 25(3):306–311, May 1979.
- [127] J. Maxted and J. Robinson. Error recovery for variable length codes. *IEEE Transactions on Information Theory*, 31(6):794–801, Nov. 1985.
- [128] N. Merhav and S. hamai. On joint source–channel coding for the Wyner–Ziv source and the Gelfand–Pinsker channel. *IEEE Transactions on Information Theory*, 49(11):2844–2855, Nov. 2003.
- [129] A. Miguel, A. Mohr, and E. Riskin. SPIHT for generalized multiple description coding. In *Proc. of IEEE International Conference on Image Processing*, volume 3, pages 842–846, Oct 1999.
- [130] U. Mittal and N. Phamdo. Hybrid digital-analog (hda) joint source-channel codes for broadcasting and robust communications. *IEEE Trans. on Information Theory*, 48(5):1082–1102, May 2002.
- [131] G. R. Mohammad-Khani, M. Kieffer, and P. Duhamel. Simplification of VLC tables with application to ML and MAP decoding algorithms. *IEEE Transactions on Communications*, 2005. à paraître.
- [132] G. R. Mohammad-Khani, C. M. Lee, M. Kieffer, and P. Duhamel. Simplification of vlc tables with application to ml and map decoding algorithms. *IEEE Transactions on Communications*, 2005. submitted.
- [133] A. Mohr, E. Riskin, and R. Ladner. Generalized multiple description coding through unequal loss protection. In I. I. C. on Image Processing, editor, *EEE International Conference on Image Processing*, volume 1, pages 411–415, October 1999.
- [134] A. E. Mohr, E. A. Riskin, and R. E. Ladner. Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction. *IEEE Journal on Selected Areas in Commun.*, 18(6):819–828, June 2000.
- [135] A. Murugan, P. Gopala, and H. Gamal. Correlated sources over wireless channels: cooperative source–channel coding. *IEEE Journal on Selected Areas in Communications*, 22(6):988–998, Aug. 2004.
- [136] O. Nemethova, W. Karner, A. Al-Moghrabi, and M. Rupp. Cross-layer error detection for h.264 video over umts. In *Proc. of Wireless Personal Mobile Communications, Int. Wireless Summit*, Aalborg, Denmark, Sept 2005.
- [137] O. Nemethova, W. Karner, C. Weidmann, and M. Rupp. Distortion-minimizing network-aware scheduling for umts video streaming. In *Proc. EUSIPCO*, September 2007.
- [138] C. Ng, D. Gunduz, A. Goldsmith, and E. Erkip. Recursive power allocation in gaussian layered broadcast coding with successive refinement. In *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2007.
- [139] H. Nguyen and P. Duhamel. Compressed image and video redundancy for joint source-channel decoding. In *Proc. Globecom*, 2003.

- [140] H. Nguyen and P. Duhamel. Iterative joint source–channel decoding of variable length encoded video sequences exploiting source semantics. In *Proceedings of ICIP*, 2004.
- [141] H. Nguyen, P. Duhamel, J. Brouet, and D. Rouffet. Robust vlc sequence decoding exploiting additional video stream properties with reduced complexity. In *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, pages 375–378, June 2004. Taipei, Taiwan.
- [142] J. W. Nieto and W. N. Furman. Cyclic redundancy check (CRC) based error method and device. US Patent US 2007/0192667 A1, Aug. 16 2007.
- [143] Y. Oohama. Gaussian multiterminal source coding. *IEEE Transactions on Information Theory*, 43(6):1912–1923, Nov. 1997.
- [144] Y. Oohama. The rate–distortion function for the quadratic gaussian CEO problem. *IEEE Transactions on Information Theory*, 44(3):1057–1070, May 1998.
- [145] V. N. Padmanabhan, H. J. Wang, and P. A. Chou. Resilient peer–to–peer streaming. *IEEE International Conference on Network Protocols*, pages 16–27, Nov. 2003.
- [146] E. Paolini, M. Chiani, G. Liva, and B. Matuz. Generalized IRA erasure correcting codes for hybrid iterative / maximum likelihood decoding. *IEEE Communications Letters (to be issued)*, 2008.
- [147] E. Paolini, M. Varrella, M. Chiani, G. Liva, and B. Matuz. Low-complexity LDPC codes with near-optimum performance over the BEC. In *Proc. of the 4-th Advanced Satellite Mobile Systems Conference*, 2008.
- [148] R. C. Pasco. *Source Coding Algorithms for Fast Data Compression*. Ph.D. Thesis Dept. of EE, Stanford University, CA, 1976.
- [149] L. Perros-Meilhac and C. Lamy. Huffman tree based metric derivation for a low-complexity soft VLC decoding. In *Proc. IEEE ICC*, volume 2, pages 783–787, 2002.
- [150] V. Prabhakaran, D. Tse, and K. Ramchandran. Rate region of the quadratic gaussian CEO problem. *Proceedings of the International Symposium on Information Theory*, page 119, June 2004.
- [151] S. Pradhan, J. Chou, and K. Ramchandran. Duality between source coding and channel coding with side information. *IEEE Transactions on Information Theory*, 49(5):1181–1203, May 2003.
- [152] S. Pradhan, J. Kusuma, and K. Ramchandran. Distributed compression in a dense microsensor network. *IEEE Signal Processing Magazine*, 19(2):51–60, Mar. 2002.
- [153] S. Pradhan and K. Ramchandran. Distributed source coding using syndromes (DISCUS): Design and construction. *Proceedings of the IEEE Data Compression Conference*, pages 158 – 167, Mar. 1999.
- [154] R. Puri and K. Ramchandran. Multiple description source coding using forward error correction codes. In *Proc. of Thirty-Third Asilomar Conference on signals, systems and computers*, volume 1, pages 342–346, October 1999.
- [155] N. Rahnavard, B. Vellambi, and F. Fekri. Rateless codes with unequal error protection property. *IEEE Transactions on Information Theory*, 53(4):1521 – 1532, April 2007.
- [156] S. Ray, M. Effros, M. Medard, T. Ho, D. Karger, R. Koetter, and A. J. On separation, randomness, and linearity for network codes over finite fields. *IEEE Transactions on Information Theory*, 2006.
- [157] G. R. Redinbo. Decoding real block codes: Activity detection, wiener estimation. *IEEE trans. on Information Theory*, 46(2):609–623, 2000.

- [158] Z. Reznic, M. Feder, and R. Zamir. Distortion bounds for broadcasting with bandwidth expansion. *IEEE Transactions on Information Theory*, 52(8):3778–3788, Aug. 2006.
- [159] T. Richardson and U. Urbanke. *Modern Coding Theory*. Cambridge University Press, 2008.
- [160] B. Rimoldi. Successive refinement of information: characterization of achievable rates. *IEEE Trans. Information Theory*, 40(1):253–259, Jan. 1994.
- [161] O. Rioul and M. Vetterli. Wavelet and signal processing. *IEEE SP Magazine*, pages 14–37, Oct. 1991.
- [162] J. Rissanen. Generalized Kraft inequality and arithmetic coding. *IBM Journal of Research and Development*, 20(3):198, 1976.
- [163] A. Roumy, K. Lajnef, and C. Guillemot. Rate-adaptive turbo-syndrome scheme for Slepian–Wolf coding. *IEEE Asilomar Conference on Signals, Systems, and Computers*, pages 545–549, Nov. 2007.
- [164] T. Ruster, M. Spiertz, and J. Ohm. H.264/avc compatible scalable multiple description video coding with rd optimization. In *ISPACS '06: International Symposium on Intelligent Signal Processing and Communications*, pages 131 – 134, Dec. 2006.
- [165] G. Sabeva, S. Ben-Jamaa, M. . Kieffer, and P. Duhamel. Robust decoding of h.264 encoded video transmitted over wireless channels. In *Proceedings of MMSP*, Victoria, Canada, 2006.
- [166] M. Sartipi and F. Fekri. Distributed source coding in wireless sensor networks using ldpc coding: the entire slepian-wolf rate region. In *IEEE Wireless Communications and Networking Conference*, March 2005.
- [167] K. Sayood. *Introduction to Data Compression, Second Edition*. Morgan Kaufmann, San Francisco, 2000.
- [168] D. Sejdinovic, D. Vukobratovic, A. Doufexi, V. Senk, and R. J. Piechocki. Expanding window fountain codes for unequal error protection. In *Forty-First Asilomar Conference on Signals, Systems and Computers*, pages 1020 – 1024, Nov. 2007.
- [169] S. Servetto. Lattice quantization with side information. *Proceedings of the IEEE Data Compression Conference*, pages 510–519, Mar. 2000.
- [170] S. D. Servetto, K. Ramchandran, V. A. Vaishampayan, and K. Nahrstedt. Multiple description wavelet based image coding. *IEEE Trans. on Image Processing*, 9(5):813–826, May 2000.
- [171] S. Sesia, G. Caire, and G. Vivier. Lossy transmission over slow-fading awgn channels: a comparison of progressive, superposition and hybrid approaches. In *Proc. IEEE International Symposium on Information Theory*, Sept. 2005.
- [172] S. Shamai. A broadcast strategy for the gaussian slowly fading channel. In *Proc. 1997 International Symposium on Information Theory*, page 150, June 1997.
- [173] S. Shamai and A. Steiner. A broadcast approach for single-user slowly fading mimo channel. *IEEE Trans. Information Theory*, 49(10):2617–2635, Oct. 2003.
- [174] S. Shamai and S. Verdú. Capacity of channels with side information. *European Transactions on Telecommunications*, 6(5):587–600, Sept.–Oct. 1995.
- [175] S. Shamai, S. Verdú, and R. Zamir. Systematic lossy source/channel coding. *IEEE Transactions on Information Theory*, 44(2):564–579, Mar. 1998.

- [176] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 1948.
- [177] M. Shokrollahi. Raptor codes. *IEEE Trans. Inform. Theory*, 52(6):2551–2567, June 2006.
- [178] D. Slepian and J. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 19(4):471–480, July 1973.
- [179] K. Stuhlmüller, N. Färber, M. Link, and B. Girod. Analysis of video transmission over lossy channels. *IEEE Jour. on Selected Areas in Communications*, 18(6), June 2000.
- [180] K. Subbalakshmi and S. Somasundaram. Multiple description image coding framework for EBCOT. In *Proc. of IEEE International Conference on Image Processing*, volume 3, pages 541 – 544, June 2002.
- [181] P. F. Swaszek and P. DiCicco. More on the error recovery for variable-length codes. *IEEE Transactions on Information Theory*, 41(6):2064–2071, Nov. 1995.
- [182] A. S. Tan, A. Aksay, C. Bilen, G. Bozdagi Akar, and E. Arikan. Rate-distortion optimized layered stereoscopic video streaming with raptor codes. In *Packet Video 2007*, pages 98 – 104, Nov. 2007.
- [183] P. Tan and J. Li. Enhancing the robustness of distributed compression using ideas from channel coding. *Global Telecommunications Conference*, 4:5pp, Nov. 2005.
- [184] R. Tannious and A. Nosratinia. Relay channel with private messages. *IEEE Transactions on Information Theory*, 53(10):3777–3785, Oct. 2007.
- [185] D. Taubman. High performance scalable image compression with EBCOT. *IEEE Trans. on Image Processing*, 9(7), July 2000.
- [186] D. Taubman and M. Marcellin. *JPEG-2000 Image Compression Fundamentals, Standards and Practice*. Norwell, MA: Kluwer Academic Publishers., 2002.
- [187] S. Ten Brink. Convergence behavior of iteratively decoded parallel concatenated codes. *IEEE Trans. Comm.*, 49:1727–1737, 2001.
- [188] R. Thobaben and J. Kliewer. On iterative source-channel decoding for variable-length encoded markov sources using a bit-level trellis. In *Proc. IV IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC'03)*, Rome, 2003.
- [189] R. Thobanen and J. Kliewer. Robust decoding of variable-length encoded markov sources using a three-dimensional trellis. *IEEE Communications Letters*, 7(7):320–322, 2003.
- [190] T. Tillo, E. Baccaglini, and G. Olmo. A flexible multi-rate allocation scheme for balanced multiple description coding applications. In *Proc. IEEE International Workshop on Multimedia Signal Processing (MMSP 2005)*, Nov. 2005.
- [191] T. Tillo, M. Grangetto, and G. Olmo. A flexible error resilient scheme for jpeg 2000. In *Multimedia Signal Processing, 2004 IEEE 6th Workshop on*, pages 295–298, 29 Sept.-1 Oct. 2004.
- [192] T. Tillo, M. Grangetto, and G. Olmo. Redundant slice optimal allocation for h.264 multiple description coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(1):59 – 70, Jan. 2008.
- [193] T. Tillo and G. Olmo. A novel multiple description coding scheme compatible with the JPEG 2000 decoder. *IEEE Signal Processing Letter*, 11(11):908–911, Nov. 2004.

- [194] D. N. C. Tse. Optimal power allocation over parallel gaussian broadcast channels. In *U.C. Berkeley technical report UCB/ERL M99/7*. <http://www.eecs.berkeley.edu/Pubs/TechRpts/1999/3578.html>, 1999.
- [195] Z. Tu, J. Li, and R. Blum. Compression of a binary source with side information using parallel concatenated convolutional codes. *IEEE Global Telecommunications Conference*, 1:46–50, Nov. 2004.
- [196] Z. Tu, J. Li, and R. Blum. An efficient SF–ISF approach for the Slepian–Wolf source coding problem. *Eurasip Journal on Applied Signal Processing*, 6:961–971, May 2005.
- [197] E. Tuncel. Slepian–Wolf coding over broadcast channels. *IEEE Transactions on Information Theory*, 52(4):1469–1482, Apr. 2006.
- [198] S. Tung. *Multiterminal Source Coding*. PhD thesis, May 1978.
- [199] R. Tupelly, J. Zhang, and E. Chong. Opportunistic scheduling for streaming video in wireless networks. In *Proc. of Int. Conf. on Inf. Sciences and Systems*, Johns Hopkins University, Baltimore, MD, March 2003.
- [200] Z. X. V. Stankovic, A.D. Liveris and C. Georghiades. Design of slepian-wolf codes by channel code partitioning. In *Proceedings of the IEEE data compression conference*, 2004.
- [201] V. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Trans. on Information Theory*, 39(3):821–834, May 1993.
- [202] V. Vaishampayan and J. Domaszewicz. Design of entropy-constrained multiple-description scalar quantizers. *IEEE Trans. on Information Theory*, 40(1):245–250, Jan. 1994.
- [203] V. Vaishampayan, N. J. A. Sloane, and S. D. Servetto. Multiple-description vector quantization with lattice codebooks: design and analysis. *IEEE Trans. on Information Theory*, 47(5):1718–1734, July 2001.
- [204] E. Van Der Meulen. Three-terminal communication channels. *Advances in Applied Probability*, 3(1):120–154, 1971.
- [205] M. Van der Schaar and S. Shankar. Cross-layer wireless multimedia transmission: Challenges, principles, and new paradigms. *IEEE Wireless Communications Magazine*, 12(4):50–58, 2005.
- [206] K. Vermeirsch, Y. Dhondt, S. Mys, and R. Van de Walle. Low complexity multiple description coding for h.264/avc. In *WIAMIS '07. Eighth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 61–64, June 2007.
- [207] A. J. Viterbi and J. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, New-York, 1979.
- [208] F. D. Vito, D. Quaglia, and J. D. Martin. Model-based distortion estimation for perceptual classification of video packets. In *Proc. of IEEE MMSP*, Sep 2004.
- [209] J. Wagner, J. Chakaresk, and P. Frossard. Streaming of scalable video from multiple servers using rateless codes. In *IEEE International Conference on Multimedia and Expo*, pages 1501 – 1504, July 2006.
- [210] T. Wang, M. T. Orchard, V. Vaishampayan, and A. R. Reibman. Multiple description coding using pairwise correlating transforms. *IEEE Transactions on Image Processing*, 10(3):351–366, June 2001.

- [211] X. Wang and M. Orchard. Design of trellis codes for source coding with side information at the decoder. *Proceedings of the IEEE Data Compression Conference*, pages 361–370, Mar. 2001.
- [212] C. Weidmann. Constant rate loss strategies for the gaussian broadcast channel. In *Proc. NEWCOM-ACoRN Joint Workshop*, September 2006.
- [213] C. Weidmann, F. Bassi, and M. Kieffer. Practical distributed source coding with impulse-noise degraded side information at the decoder. In *Proc. EUSIPCO*, August 2008. to appear.
- [214] L. Welch. Hidden Markov models and the Baum–Welch algorithm. *IEEE Information Theory Society Newsletter*, 53(4):10–13, 2003.
- [215] J. Widmer and J. Y. Le Boudec. Network coding for efficient communication in extreme networks. In *WDTN05: Proceeding of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, pages 284–291, Aug. 2005.
- [216] I. H. Witten, R. M. Neal, and J. G. Cleary. Arithmetic coding for data compression. *Communications of the ACM*, 30(6):520–540, 1987.
- [217] J. K. Wolf. Efficient maximum-likelihood decoding of linear block codes using a trellis. *IEEE Trans. Inform. Theory*, 24(1):76–80, 1978.
- [218] K. K. Wong. *The Soft-Output M-Algorithm And Its Applications*. Phd thesis, Queen’s University, Kingston, Canada, 2006.
- [219] A. Wyner. Recent results in the Shannon theory. *IEEE Transactions on Information Theory*, 20:2–10, 1974.
- [220] A. Wyner and J. Ziv. The rate–distortion function for source coding with side information at the receiver. *IEEE Transactions on Information Theory*, 22:1–11, Jan. 1976.
- [221] Z. Xiong, A. Liveris, and S. Cheng. Distributed source coding for sensor networks. *IEEE Signal Processing Magazine*, 21(5):80–94, Sept. 2004.
- [222] Q. Xu, V. Stankovic, , and Z. Xiong. Layered Wyner–Ziv video coding for transmission over unreliable channels. *IEEE Signal Processing Magazine*, 86(11):3212–3225, 2006.
- [223] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao. Wyner-Ziv coding based on TCQ and LDPC codes. *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, 1:825–829, Nov. 2003.
- [224] Y. Zhao and J. Garcia-Frias. Turbo compression/joint source–channel coding of correlated binary sources with hidden Markov correlation. *Signal Processing*, 86(11):3115–3122, Nov. 2006.
- [225] W. Zhong and J. Garcia-Frias. LDGM codes for channel coding and joint source–channel coding of correlated sources. *EURASIP Journal on Applied Signal Processing*, 6:942–953, 2005.
- [226] W. Zhong, H. Lou, and J. Garcia-Frias. LDGM codes for joint source–channel coding of correlated sources. *International Conference on Image Processing*, 1:14–17, Sept. 2003.
- [227] G. Zhou and Z. Zhang. Synchronization recovery of variable length codes. *IEEE Transactions on Information Theory*, 48(1):219–227, 2002.