

Fountain Coding with Decoder Side Information

Dino Sejdinović, Robert J. Piechocki and Angela Doufexi

Centre for Communications Research

Department of Electrical & Electronic Engineering

University of Bristol, Bristol, UK

Email: {d.sejdinovic, r.j.piechocki, a.doufexi}@bristol.ac.uk

Mohamed Ismail

Toshiba Research Europe Ltd

Telecommunications Research Laboratory

32 Queen Square, Bristol, UK

Email: mohamed.ismail@toshiba-trel.com

Abstract—In this contribution, we consider the application of Digital Fountain (DF) codes to the problem of data transmission when side information is available at the decoder. The side information is modelled as a “virtual” channel output when original information sequence is the input. For two cases of the system model, which model both the virtual and the actual transmission channel either as a binary erasure channel or as a binary input additive white Gaussian noise (BIAWGN) channel, we propose methods of enhancing the design of standard non-systematic DF codes by optimizing their output degree distribution based on the side information assumption. In addition, a systematic Raptor design has been employed as a possible solution to the problem.

I. INTRODUCTION

Digital fountain (DF) codes are a universal, capacity-approaching Forward Error Correction (FEC) solution for data transmission over lossy networks. The first practical DF codes were Luby-Transform (LT) codes [1], whereas their extension, Raptor codes [2], represent a state-of-the-art DF solution for lossy transmission with excellent performance and linear encoding/decoding complexity. A significant amount of work has been done to extend the methods of DF codes for transmission over noisy channels [3], [4]. In this paper we consider another possible application of DF codes, which is set in the realm of distributed joint source-channel coding (DJSCC). We assume that the side information is available at the decoder, and that this side information is modelled as the output of a “virtual” channel when the original information sequence is its input. The DF encoding methods which incorporate the presence of decoder side information in their design are explored and studied. The aim of these encoding methods is to provide both the distributed source compression scheme and the channel coding scheme with a single DF code.

The distributed source coding part of the DJSCC problem, concerned with the separate encoding and joint decoding of two correlated sources, was originally studied by Slepian and Wolf [5], who produced a celebrated result that separate compression (Slepian-Wolf Coding - SWC) suffers no rate loss compared to the case of joint compression. The special case when one of the sources is fully known at the decoder (decoder side information) and its generalization, Wyner-Ziv Coding, have since attracted a lot of interest due to their practical significance. The usual means for constructing a binning scheme for SWC uses the syndrome-based encoder. The syndrome-based encoder for the SWC of the information

sequence \mathbf{x} having length k forms a $(k-r)$ -length syndrome vector $\mathbf{s} = \mathbf{x}\mathbf{H}^T$, based on a good linear (k, r) -channel code with parity-check matrix \mathbf{H} . Another equivalent approach uses *parity-based binning*, which employs a systematic $(k+t, k)$ -channel code and transmits the t -length parity vector, where $t = k-r$. In both cases, the Slepian-Wolf theorem requires that $t \geq kH(X|Y)$, where $H(X|Y)$ is the entropy of source X conditional on decoder side information Y . The syndrome-based approach is a better choice when the transmission channel is noiseless, due to an increase in complexity when parity-based binning is used. However, if the transmission channel is noisy, the parity-based binning may be beneficial in some cases. Namely, separate source-channel coding would utilize a syndrome-based encoder to create a syndrome vector and then concatenate a channel code to protect syndrome symbols. On the other hand, distributed joint source-channel design uses the parity-based approach to combine two channel codes, one for SWC and another for channel coding into a single channel code. The advantages of the joint design in comparison with the separate design when DF codes are used have been noted in [6] and [7].

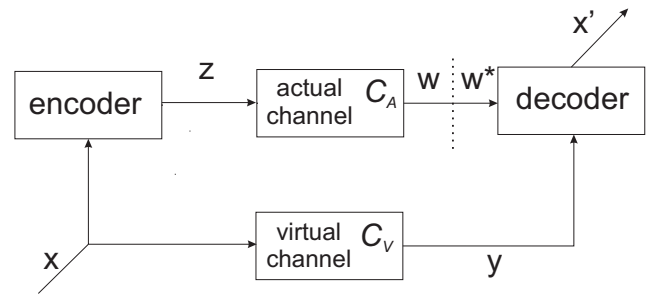


Fig. 1. System model

The system model we are considering is presented in figure 1. The binary information source X is correlated with decoder side information Y via a “virtual” correlation channel C_V . The encoder processes an information sequence $\mathbf{x} = (x_1, \dots, x_k)$ of length k at a time, produces the potentially infinite binary stream $\mathbf{z} = (z_1, z_2, \dots)$, $\mathbf{z} = f_{enc}(\mathbf{x})$, of the encoding symbols and transmits it through an “actual” transmission channel C_A . The channel outputs are depicted as the new “noisy” stream $\mathbf{w} = (w_1, w_2, \dots)$. The decoder picks up any t channel outputs $\mathbf{w}^* = (w_{i_1}, w_{i_2}, \dots, w_{i_t})$, aware of their coordinates

vector $\mathbf{i} = (i_1, i_2, \dots, i_t)$, where $t \geq kH(X|Y)/Cap(\mathcal{C}_A)$. By taking advantage of the side information sequence $\mathbf{y} = (y_1, \dots, y_k)$ corresponding to \mathbf{x} , the information sequence $\mathbf{x}' = f_{dec}(\mathbf{w}^*, \mathbf{i}, \mathbf{y})$ is decoded. The objective is to devise the encoding strategy such that it is possible to have the true rate t/k close to the optimal value equal to the Slepian-Wolf limit in the noisy channel case, $H(X|Y)/Cap(\mathcal{C}_A)$, and to still allow for the high probability of successful decoding, i.e., of $\mathbf{x}' = \mathbf{x}$.

We consider two approaches of application of DF codes for the proposed problem. The first one deals with the standard non-systematic DF code as a single code for source and channel coding and its enhancement based on the assumption that the decoder side information is present. The second one employs a systematic Raptor design as the possible solution. Both approaches can also be used as the basis for the Combined Incremental Redundancy Hybrid ARQ (Automatic Repeat reQuest) - IR-HARQ schemes which are studied in [8].

II. A SIMPLE "ERASURES-ONLY" SCENARIO

One of the simplest nontrivial cases described by our model is the one where both the actual and the virtual channel are binary erasure channels. This is the case where the side information available at the decoder is given by a certain portion of the original information sequence, i.e., the side information is the output of a binary erasure channel with a fixed erasure probability p . One can think of this case as of a scenario in which the receiver has successfully decoded only a part of the data prior to the failure of previous transmission. The previous transmission may have used a classical DF code for erasure protection, and the obvious solution would be just to increment redundancy by transmitting additional DF encoding symbols. However, the failure of decoding may have fully erased all the supplementary data, e.g., the decoding graph and previously received encoding symbols, due to the limited buffer size, and thus continuation of the same encoding procedure is by no means advisable. On the other side, the transmitter has no knowledge about *which* part of the data has already been decoded, but it is able to estimate *how much* of the data has already been decoded and needs to adapt its encoding procedure to take the advantage of the portion of data available at the decoder, i.e., the decoder side information.

In the remainder of this section we investigate the possible solutions for the nearly optimal transmission based on the DF code design which makes use of these assumptions. This scenario also provides great insight into more realistic case when both the actual and the virtual channels are, in fact, noisy, which we consider in section III.

A. Universal systematic Raptor solution

The "erasures-only" scenario is not different from the transmission over binary erasure channel, assuming that the systematic symbols are sent first and the erasure rate during the transmission of systematic symbols can be estimated at the encoder. Hence, a universal systematic rateless code for transmission over an erasure channel would be sufficient to

optimally solve the proposed problem of DJSCC. Systematic Raptor design was discussed in [2] and it has been adopted in practical applications like Multimedia Broadcast/Multicast Services (MBMS) within 3GPP [9], amongst others. In this design, the decoder does not decode the original information sequence \mathbf{x} but the *intermediate* symbols $\bar{\mathbf{x}}$ instead, related to the information sequence by

$$\mathbf{G}_{LT}^{\{1:k\}} \bar{\mathbf{x}}^T = \mathbf{x}, \quad (1)$$

where $\mathbf{G}_{LT}^{\{1:k\}}$ represents the first k rows of the LT generator matrix. The encoder needs to calculate intermediate symbols $\bar{\mathbf{x}}$ for each information sequence via Gaussian elimination, whereas the decoder has an additional *encoding* step upon the successful decoding in order to calculate the actual information sequence by multiplying intermediate symbols $\bar{\mathbf{x}}$ with the LT generator matrix. The universality of Raptor codes for erasure channels implies that application of the systematic Raptor codes to the proposed problem will bring nearly optimal design. However, our aim in the rest of this section is to study the limitations of the simpler, i.e., non-systematic DF codes when applied to the proposed "erasures-only" problem of DJSCC. The advantage of the possible applicable solution based on standard non-systematic DF codes is the lower complexity and simplicity of design, since systematic Raptor codes require significant amount of preprocessing as well as Gaussian elimination to be performed at the transmitter for each transmitted block [9]. Another advantage of the solution based on standard non-systematic DF codes may arise in the practical scenario of incrementing redundancy, which we discussed above. Namely, the encoder needs only to modify its DF output degree distribution in order to adapt to the failure of decoding at the receiver side.

B. Encoding with the non-systematic fountain code

Let us consider the case when the encoding method employs a simple LT code with an output degree distribution Φ . During the transmission, the encoder samples the distribution Φ in order to generate as many encoding symbols as necessary until the decoder picks up $t \geq pk$ correctly received symbols, where p is the channel erasure probability. Note that pk symbols correspond to the optimal compression rate, since $H(X|Y) = p$. The receiver forms the decoding graph based on the received encoding symbols and removes the known, previously decoded input nodes from the graph, based on the symbols available from the side information, correspondingly updating the output nodes. A belief propagation decoding for the erasure channel, i.e., graph pruning procedure, can then be performed.

Once removal of the already known input nodes has occurred, the output degree distribution changes considerably, which influences erasure correction performance. However, since the source symbols are chosen uniformly after the degree of an output node has been selected, one can relate the "starting" degree distribution Φ to the degree distribution after removal of the known source nodes from the decoding graph. Let $\Phi(x) = \sum_{d=0}^n \Phi_d x^d$ and $\Omega(x) = \sum_{d=0}^n \Omega_d x^d$ be the generating polynomials of the output degree distribution used

at the encoder (*incoming* degree distribution) and the output degree distribution after removal of the known source nodes from the decoding graph (*resulting* degree distribution). The probability that an arbitrary output node has degree i after removal of the known source nodes conditioned on its degree before removal being $j \geq i$ is clearly $\binom{j}{i}(1-p)^{j-i}p^i$. Thus, the relation between the distributions Φ and Ω is given by the following set of linear equations

$$\begin{aligned}\Omega_0 &= \Phi_0 + \Phi_1(1-p) + \Phi_2(1-p)^2 + \dots + \Phi_k(1-p)^k \\ \Omega_1 &= \Phi_1 p + \Phi_2 \binom{2}{1}(1-p)p + \dots + \Phi_k \binom{k}{1}(1-p)^{k-1}p \\ \Omega_2 &= \Phi_2 p^2 + \dots + \Phi_n \binom{k}{2}(1-p)^{k-2}p^2 \\ &\dots = \dots \\ \Omega_k &= \Phi_k p^k.\end{aligned}\quad (2)$$

This set of linear equations need not have a positive solution, i.e., the ensemble of distributions $\{\Phi^{(i)}\}_{i \in \mathbb{N}}$, such that the expected ensemble of distributions $\{\Omega^{(i)}\}_{i \in \mathbb{N}}$ after removal of the known source nodes is the one that achieves the capacity in the sense of DF codes, need not exist. However one can try to look for incoming distributions Φ whose resulting distributions are “close” to the asymptotically optimal output degree distributions in some sense, for example via minimization of the sum of the squares of the differences between coordinates of distribution (nonnegative least squares problem).

Another approach would be to devise methods of optimizing the incoming distributions through its relation with the resulting distribution. Namely, in generating polynomial notation, (2) gives the relation

$$\Omega(x) = \Phi(1-p+px), \quad (3)$$

which allows for the simple analysis of the performance of distribution Φ in terms of the and-or lemma and density evolution [10]. The corresponding density evolution of asymptotic bit error rate (BER) in terms of the distribution Φ can be calculated by using the relation (3) and is given by

$$\begin{aligned}y_0 &= 1 \\ y_l &= e^{-(1+\varepsilon)p\Phi'(1-py_{l-1})}, \quad l \geq 1,\end{aligned}\quad (4)$$

where ε is the code overhead.

It should be noted that this density evolution gives the asymptotic BER across the input symbols unknown at the decoder, and the actual bit error rate across the entire information sequence is given by $\hat{y}_\infty = p \cdot y_\infty$, where

$$y_\infty = \lim_{l \rightarrow \infty} y_l. \quad (5)$$

C. Construction and asymptotic analysis of incoming distributions

Asymptotically good Soliton-like DF output degree distribution [2] is given by

$$\Omega_\varepsilon(x) = \frac{1}{\mu+1} \left(\mu x + \sum_{d=2}^D \frac{x^d}{(d-1)d} + \frac{x^{D+1}}{D} \right). \quad (6)$$

For any fixed overhead $\varepsilon > 0$, this distribution provides for recovering at least $(1-\delta)k$ input symbols from any set of $(1+\varepsilon)k$ output symbols, where k is the length of the information sequence, $\delta = (\varepsilon/2)/(1+2\varepsilon)$, $\mu = \varepsilon + \varepsilon^2$ and $D = \lceil 2(1+2\varepsilon)/\varepsilon \rceil$.

The system (2) cannot be solved for Φ in case of the distribution from (6), for any values of ε and p of interest. For example, $\Phi_D \geq 0$ would imply $1-p \leq \frac{1}{D^2}$, which does not have any practical sense. However, one can look for such an output degree distribution Φ , for which $\Phi(1-p+px)$ closely resembles $\Omega_\varepsilon(x)$, using the nonnegative least squares method. In case of $\varepsilon = 0.03$ and $p = 0.3$, the custom degree distribution Φ_ε was obtained by using the nonnegative least squares method. This distribution is quite different from Soliton-like distributions and the sampling of this distribution produces more than 80% of encoding symbols with degree 8. Figure 2 shows how well the resulting distribution $\Phi_\varepsilon(0.7+0.3x)$ approximates Ω_ε in terms of the density evolution estimates of the bit error rate. We can see that the imperfection of approximation introduces a significant additional delay, i.e., the knee in the BER curve occurs at much higher overhead than that of the distribution (6).

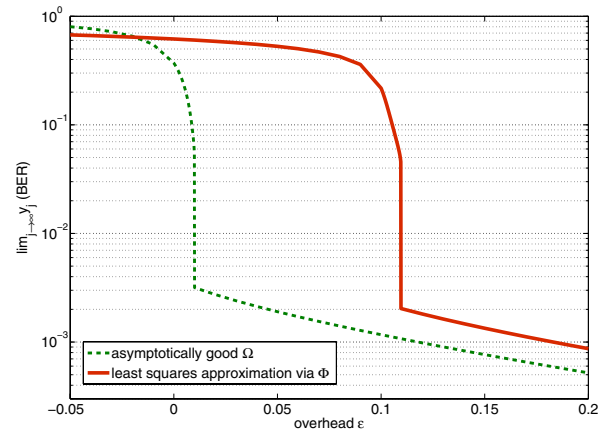


Fig. 2. The approximation of the asymptotically good distribution (6) by the nonnegative least squares method.

A different approach yields asymptotically better distributions Φ . From the generalized density evolution for the side information LT codes (4), we obtain the condition

$$\Phi'(x) \geq -\frac{1}{p(1+\varepsilon)} \ln\left(\frac{1-x}{p}\right), \quad x \in [1-p, 1-p\delta], \quad (7)$$

which needs to be satisfied if Φ should provide for the decoding $(1-\delta)t$ portion of data after reception of $(1+\varepsilon)t$ output symbols, when $k \rightarrow \infty$, where $t \sim pk$ is the number of unknown source symbols. When $p = 1$, one can find such a distribution $\Phi = \Omega$, for any pair $\varepsilon > 0$, $\delta > 0$. However, we conjecture that this is not the case for $p < 1$, since we have found by extensive simulation that for the fixed value of ε there is a certain threshold value of δ_0 , such that the

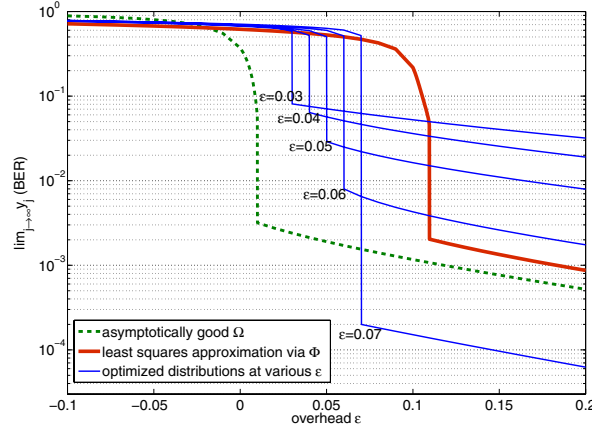


Fig. 3. The asymptotic performance of the optimized distributions for various overheads 3-7%

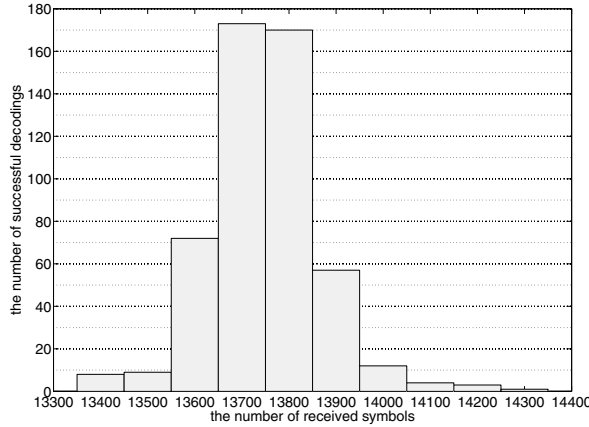


Fig. 4. The precoded DF code (Raptor) with modified incoming distribution

problem (7) stays unsolvable when $\delta < \delta_0$, no matter how large the maximum degree of the distribution Φ is. Figure 3 shows the asymptotic analysis of the bit error rate for the best distributions found at various overheads. We have fixed the overhead value, and then searched for the lowest possible δ such that (7) becomes feasible. For that δ , by means of linear programming (discretization of the segment), we have found the distributions Φ with minimum average degree. This method can be naturally extended to the finite length design, similarly as in [3], based on setting lower bounds on the expected size of the input ripple [11]. Table I shows four distributions calculated in this way.

By concatenating an LT code with a custom distribution derived by the optimization procedure as described above to a very high-rate hybrid LDPC-Half precode, constructed as in [9], we have implemented a non-systematic solution to the 'erasures-only' problem. The length of the information sequence was set to $k = 40000$ and virtual channel erasure

TABLE I
OPTIMIZED DEGREE DISTRIBUTIONS FOR VARIOUS BLOCK LENGTHS

d	Φ_d	d	Φ_d	d	Φ_d	d	Φ_d
9	0.7509	9	0.6711	9	0.6989	9	0.6689
10	0.1471	10	0.2305	10	0.2016	10	0.2330
63	0.0722	67	0.0647	63	0.0735	63	0.0782
141	0.0243	118	0.0074	158	0.0240	212	0.0032
142	0.0055	119	0.0263	159	0.0020	213	0.0168
k	ε	k	ε	k	ε	k	ε
10000	0.1	20000	0.09	40000	0.08	∞	0.06

rate was $p = 0.3$. Note that the precoding somewhat changes the optimization procedure since equation (3) now reads

$$\Omega(x) = \Phi(s(1-p) + (1-s+sp)x), \quad (8)$$

where s is the precode rate and density evolution-based constraints (7) are updated correspondingly. Figure 4 depicts the histogram of the numbers of received symbols necessary for successful decoding for 500 transmission trials. On average, 13750 symbols were required, which is about 34.4% of the length of information sequence, compared to the optimal 30%, i.e., 12000 symbols.

III. BIAWGN CASE - SOFT INFORMATION DECODING

In this section we assume that the binary information source X over the alphabet $\{-1, 1\}$ (standard BPSK mapping $0 \rightarrow 1, 1 \rightarrow -1$ is used) and the soft side information Y are correlated via $Y = X + N$, where N is a Gaussian random variable of zero mean and variance σ_V^2 . This means that $y = \text{BIAWGN}_{\sigma_V}(x)$, where BIAWGN stands for the binary input additive white Gaussian noise channel (this is the virtual channel). Note that in this case we have $H(X|Y) = 1 - \text{Cap}(\text{BIAWGN}_{\sigma_V})$, where the capacity of BIAWGN channel of noise variance σ^2 [3] is given by

$$\text{Cap}(\text{BIAWGN}_{\sigma}) = 1 - \frac{1}{2\sqrt{\pi m}} \int_{-\infty}^{\infty} \log_2(1 + e^{-x}) e^{-\frac{(x-m)^2}{4m}} dx, \quad (9)$$

where $m = 2/\sigma^2$.

A. Encoding with the non-systematic fountain code

Similarly as in the "erasures-only" case, when using a simple non-systematic fountain code, e.g., an LT code with distribution Ω , the decoder embeds the side information directly into the decoding graph. Here, side information is embedded into the decoding graph directly at the input nodes via *intrinsic soft information*, i.e., log-likelihood ratios based on the output of the virtual BIAWGN channel. Therefore, even if the transmission channel is assumed to be a noiseless or an erasure channel, belief propagation decoding for this scenario is no longer a simple graph pruning procedure, due to the presence of soft information at the input nodes. We may further assume that the transmission occurs through another BIAWGN channel, with a noise variance σ_A^2 and in addition, output nodes may contain soft information based on the actual

channel output. The more general sum-product algorithm may be employed for this kind of construction, with the soft information present at both sides of the decoding graph. In the case when a precode is used (Raptor-like scenario), parity symbols are initialized to have zero log-likelihood ratios, since no side information is directly available about the parity symbols and parity checks are initialized to have log-likelihood ratios equal to $+\infty$, since they are deterministically equal to zero. In [6], precisely this kind of design with irregular repeat accumulate (IRA) [12] precoding was employed for the joint-source channel coding scenario using non-systematic Raptor codes with a classical Soliton-like output degree distribution. In the rest of this subsection, we will show how to enhance the design of the output degree distribution under the assumption that decoder side information is available.

The optimization of fountain code output degree distribution for the BIAWGN channel when soft information is not present at the input nodes was treated in [3] by using a refined Gaussian approximation [13]. This optimization is based on the simple rationale that the means of the messages transmitted from the input nodes to the output nodes should keep on increasing under the usual all-zeroes information sequence assumption. The linear program used for optimization carries the constraints

$$\alpha \sum_d \omega_d f_d(\mu) > \mu, \mu \in (0, \mu_{max}) \quad (10)$$

where $\omega(x) = \frac{\Omega'(x)}{\Omega'(1)}$ is the output edge degree distribution (the proportion of incoming messages which carry the mean $f_d(\mu)$), α is the average input degree and $f_d(\mu)$ is the refined mean [3] of the messages passed from the output node of degree d when the mean of the incoming messages is μ .

Note that the absence of the soft information implies that the starting mean of the messages sent from the input nodes to the output nodes has to be zero. However, in the case when there is soft information available also at the input nodes, the constraints imposed by (10) are more strict than it is necessary. If we assume that the information sequence is an all-zeroes message, then soft side-information available (modelled by a virtual correlation BIAWGN channel) induces that the mean of the messages sent from the input nodes to the output nodes at the first iteration of the message-passing algorithm is $\mu_V = 2/\sigma_V^2$, where σ_V^2 is the noise variance of the virtual channel. Hence, means should keep on increasing only on the interval (μ_V, μ_{max}) . Also, the input node-update is now different since it needs to take intrinsic soft information into account. The new “mean-increase” condition is given by:

$$\mu_V + \alpha \sum_d \omega_d f_d(\mu) > \mu, \mu \in (\mu_V, \mu_{max}). \quad (11)$$

We have adopted this new set of linear constraints in order to optimize output degree distributions under the assumption that side information is present at the decoder, similarly as in the “erasures-only” case.

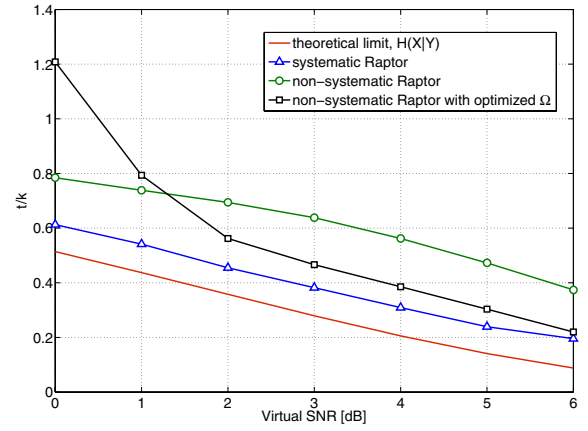


Fig. 5. The comparison of different DJSC Raptor schemes, noiseless transmission.

B. Encoding with a Systematic Raptor

One can also employ the systematic Raptor codes [2], [9] for the DJSC design in the case when virtual channel is modelled as a BIAWGN channel. The stream z in this case consists only of the non-systematic Raptor encoding symbols and the side information is embedded into the decoding graph at the output nodes corresponding to the systematic Raptor encoding symbols, since soft information about the systematic Raptor encoding symbols is already available due to decoder side information. The nodes corresponding to the non-systematic Raptor encoding symbols may contain soft information based on the actual channel output. Although the application of the systematic Raptor codes to this kind of DJSC problems seems straightforward, practical system proposals, as in [6], use a non-systematic version of Raptor codes, arguably due to the simpler design, and enhance the code design differently, by introducing *bias* towards choosing parity symbols when forming the Raptor encoding symbols.

C. Simulation results

Three different Raptor codes for DJSC, non-systematic Raptor code with a classical Soliton-like output degree distribution, non-systematic Raptor code with an output degree distribution optimized as prescribed in subsection A, and systematic Raptor code, were simulated under the assumption that the virtual channel is modelled as the BIAWGN channel, and the results are presented in figures 5 and 6. The horizontal axis represents the signal-to-noise ratio (SNR) of the virtual channel, which is related with the channel noise variance by $\text{SNR} = 10 \log_{10} \frac{1}{\sigma^2}$. The vertical axis represents the average joint source-channel code rate necessary for successful decoding, i.e., t/k , where t is the average number of received encoding symbols at the decoder. Figure 5 shows the comparison of different Raptor schemes for the case of distributed compression, i.e., when the actual transmission channel is noiseless, whereas in figure 6 the actual channel

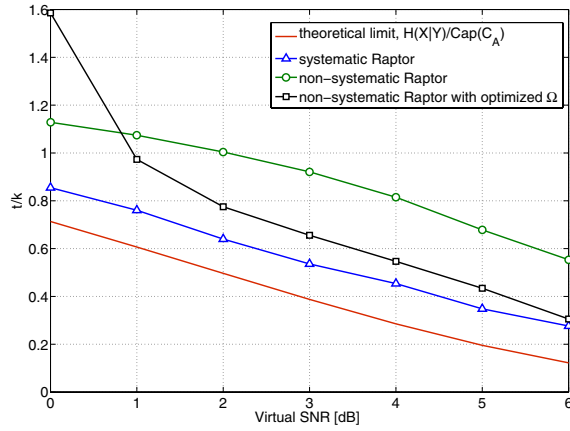


Fig. 6. The comparison of different DJSCC Raptor schemes, transmission SNR is 3 dB.

is also modelled as the BIAWGN channel with SNR set to 3 dB. The systematic Raptor design was adopted from application layer FEC scheme described in [9]. The output degree distribution Ω we obtained as prescribed above is

$$\begin{aligned} \Omega(x) = & 0.0954x^5 + 0.1192x^6 + 0.1121x^7 + \\ & 0.12938x^8 + 0.1054x^9 + 0.0807x^{10} + \\ & 0.1109x^{11} + 0.2470x^{100}. \end{aligned} \quad (12)$$

The assumed SNR of the virtual channel during this optimization was also set to 3 dB.

One can note the poor performance of applying the non-systematic Raptor code with a classical Soliton-like distribution to the proposed problem of DJSCC and considerably better performance of the systematic Raptor code scheme. However, non-systematic Raptor code with optimized Ω does approach the performance of the systematic Raptor code scheme, but is also sensitive to the assumption on the value of the SNR of the virtual channel, since this assumption is a crucial part in the distribution design.

IV. CONCLUSION

The study of the design of the DF codes based on the assumption that side information is available at the decoder is presented. We assume that side information is modelled as the “virtual” channel output when original information sequence is its input. We investigate two cases, the simple “erasures-only” case when both the virtual and the actual transmission channel are binary erasure channels and the more realistic case when both channels are BIAWGN channels. While in both cases a solution based on systematic Raptor codes seems the most advantageous, its complexity and intricate design motivate us to study the limitations of the non-systematic Raptor codes when applied to these DJSCC problems. We show how to improve their performance by optimizing the output degree distribution based on the assumption that the decoder side

information and the estimates of the correlation between the source and the decoder side information are available.

ACKNOWLEDGMENT

The authors would like to acknowledge the generous support of Toshiba Research Europe Ltd Telecommunications Research Laboratory in Bristol.

REFERENCES

- [1] M. Luby, “LT Codes,” *Proc. of the 43rd Annual IEEE Symp. Foundations of Computer Science (FOCS)*, Vancouver, Canada, Nov. 2002.
- [2] A. Shokrollahi, “Raptor Codes,” *IEEE Transactions on Information Theory*, vol. 52, No. 6, pp. 2551–2567, June 2006.
- [3] O. Etesami, A. Shokrollahi, “Raptor Codes on Binary Memoryless Symmetric Channels,” *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2033–2051, May 2006.
- [4] H. Jenkač, T. Mayer, T. Stockhammer, W. Xu. “Soft Decoding of LT Codes for Wireless Broadcast,” *Proc. of IST Summit 2005*, Dresden, Germany, June 2005.
- [5] D. Slepian, J.K. Wolf. “Noiseless Coding of correlated information sources,” *IEEE Transactions on Information Theory*, vol. IT-19, pp. 471–480, July 1973.
- [6] Q. Xu, V. Stanković, Z. Xiong, “Distributed source-channel coding of video using Raptor codes,” *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 851–861, May 2007.
- [7] N. Dütsch, H. Jenkač, T. Mayer, J. Hagenauer, “Joint Source-Channel-Fountain Coding for Asynchronous Broadcast,” *Proc. of IST Summit 2005*, Dresden, Germany, June 2005.
- [8] D. Sejdinović, V. Ponnampalam, R. Piechocki, A. Doufexi, “The Throughput Analysis of Different IR-HARQ Schemes Based on Fountain Codes,” in *Proc. of the IEEE Wireless Communications and Networking Conference*, Las Vegas, USA, April 2008.
- [9] 3GPP, “3GPP TS 26.346 V7.0.0, Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Protocols and Codecs”, Sept. 2007.
- [10] M. Luby, M. Mitzenmacher, A. Shokrollahi, “Analysis of random processes via and-or tree evaluation,” *Proc. of the 9th Annual SIAM Symp. on Discrete Algorithms*, pp. 364–373, 1998.
- [11] R. Karp, M. Luby, A. Shokrollahi, “Finite Length Analysis of LT Codes”, *Proc. of IEEE International Symposium on Information Theory 2004*, Chicago, USA, June 2004.
- [12] H. Jin, A. Khandekar, R. McEliece, “Irregular Repeat Accumulate Codes”, *Proc. of the 2nd International Symp. on Turbo Codes*, Brest, France, Sept. 2000.
- [13] M. Ardakani, F.R. Kschischang, “A More Accurate One-Dimensional Analysis and Design of Irregular LDPC Codes”, *IEEE Transactions on Communications*, vol. 52, no. 12, pp. 2106–2114, Dec. 2004.