Universal Wyner-Ziv Coding for Gaussian Sources

Elsa Dupraz, Aline Roumy, Michel Kieffer

UNIVERSAL WYNER-ZIV CODING FOR GAUSSIAN SOURCES

Elsa DUPRAZ† *Aline* ROUMY+, *Michel* KIEFFER†,◇

† L2S - CNRS - SUPELEC - Univ Paris-Sud, 91192 Gif-sur-Yvette, France
+ INRIA, Campus de Beaulieu, 35042 Rennes, France
◇ Institut Universitaire de France

ABSTRACT

This paper considers the problem of lossy source coding with side information at the decoder only, for Gaussian sources, when the joint statistics of the sources are partly unknown. We propose a practical universal coding scheme based on scalar quantization and non-binary LDPC codes, which avoids the binarization of the quantized coefficients. We first explain how to choose the rate and to construct the LDPC coding matrix. Then, a decoding algorithm that jointly estimates the source sequence and the joint statistics of the sources is proposed. The proposed coding scheme suffers no loss compared to the practical coding scheme with same rate but known variance.

***Index Terms*—** Source coding with side information, Rate-distortion, Low Density Parity Check codes, Expectation Maximization algorithm

1. INTRODUCTION

The problem of lossy source coding with side information at the decoder, also called the Wyner-Ziv problem [18], has been well investigated when the probability distribution between the source X and the side information Y is perfectly known. Practical coding schemes were proposed for this problem when the correlation channel between X and Y is assumed Gaussian [11], Gauss-Markov [4], or Bernoulli-Gaussian [3] with known parameters. The latter case accounts for the possible time variations of the variance of the correlation channel between the source and the side information.

Nevertheless, in many practical situations, such as distributed source coding in a network of sensors, the parameters of the probability distribution are in general difficult to obtain. In [12], it is shown that the performance of source coding with side information remains the same if the probability distribution $P(X)$ is unknown. Instead, we consider an uncertain correlation channel $P(Y|X)$, which is usually more difficult to characterize than the marginal $P(X)$. This has been investigated in [17], under the assumption that the correlation channel is given to the decoder but not perfectly known at the encoder. We further generalize this problem, and assume that the correlation channel is partly undisclosed at both the encoder and the decoder.

Usually this problem is addressed using a feedback channel [2]. This requires communication between the decoder and the encoder, leading to undesired delays. Alternatively, solutions based on learning sequences [7] induce a rate increase for finite-length source sequences. To avoid increasing either the delay or the rate, we propose a universal coding scheme that can handle correlated sources with unknown correlation channel parameters. As an example, we consider a Gaussian correlation channel with variance only known to belong to some interval. The encoder has to choose the rate and to design the coding process. The decoder has to reconstruct the source despite the lack of knowledge on the correlation.

The proposed coding scheme consists of a scalar quantizer that maps the real-valued source symbols into discrete symbols in GF(q), the Galois Field of size q. A Slepian-Wolf (SW) chain then transmits losslessly the quantized version of the source and a Minimum Mean Square Error (MMSE) estimator reconstructs the source from the side information and the quantization indexes. Usually, SW chain is based on Low Density Parity Check (LDPC) coding [14] of the quantized symbols mapped to bit planes. To prevent a performance loss, it is necessary for the LDPC decoder to account for the dependence between bit planes [1]. In order to avoid the binarization step and the loss in performance it incurs, LDPC codes in GF(q) [6] are considered here.

For the decoding, the classical min-sum LDPC algorithm does not perform well without the knowledge of the variance. Thus, we propose to jointly estimate the variance of the correlation channel and the source sequence, with the help of an Expectation Maximization (EM) algorithm based on the non-binary LDPC decoder. As one of the usual problems of the EM algorithm is its initialization, we also propose a method to produce a first raw estimate of the variance.

The paper is organized as follows. Section 2 introduces the notations and the considered signal model. Section 3 describes the proposed coding scheme. Section 4 shows simulation results.

2. SIGNAL MODEL

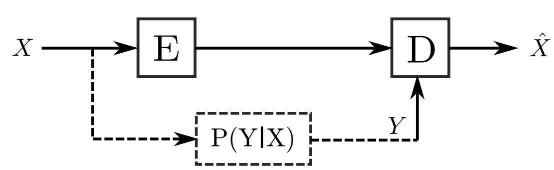

Fig. 1. The Wyner-Ziv coding scheme

The source X to encode and the SI Y, available at the decoder only, see Figure 1, produce sequences of symbols $\{X_n\}_{n=1}^{+\infty}$ and $\{Y_n\}_{n=1}^{+\infty}$. Bold upper-case letters, e.g., $\mathbf{X}_1^N = \{X_n\}_{n=1}^N$, denote random vectors, whereas bold lower-case letters, $\mathbf{x}_1^N = \{x_n\}_{n=1}^N$, represent their realizations. Moreover, when it is clear from the context that the distribution of a random variable X_n does not depend on n, the index n is omitted. Similarly, $\mathbf{X}_1^N$ is in general denoted $\mathbf{X}$.

In our setup, X is zero-mean Gaussian with variance σ_x^2, *i.e.*, $X \sim \mathcal{N}(0, \sigma_x^2)$. Assume that there exists a random variable Z independent of X such that $Y = X + Z$, $Z \sim \mathcal{N}(0, \sigma^2)$ and σ^2 is unknown but fixed for the sequence $\{(X_n, Y_n)\}_{n=1}^{+\infty}$. Denote I_{σ^2}

the set of possible σ^2 and assume that I_{σ^2} is an interval. Note that no prior distribution on σ^2 is assumed, either because this prior distribution is not known or because σ^2 is not modeled as a random variable. Such a model is stationary but non-ergodic [10, Section 3.5]. It corresponds to the WP-Model defined in [7]. Consider the quadratic distortion measure $d(X, \hat{X}) = E\left[|X - \hat{X}|^2\right]$.

3. UNIVERSAL WYNER-ZIV CODING SCHEME

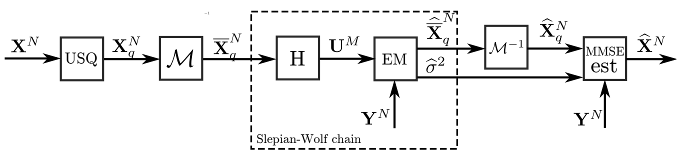

Fig. 2. Proposed universal Wyner-Ziv coding scheme

This section describes the proposed coding scheme represented in Figure 2. To encode a sequence $\mathbf{x}$ of length N, each symbol x_n is first quantized with a Uniform Scalar Quantizer (USQ). When the statistics of the source are known, [15] shows that USQ followed by SW encoding suffers only a 1.53 dB loss in the high rate regime. The USQ has $q = 2^l$ quantization levels. Denote Δ, q_k, and b_k, b_{k+1} respectively the size of the quantization cells, the level, and the boundaries of the k-th quantization level. The quantization operation gives a sequence $\mathbf{x}_\text{q}$ of real symbols with values given by the quantization levels. We assume that there exists a random variable $B_\text{q} \sim \mathcal{N}(0, \Delta^2/12)$ such that $X = X_\text{q} + B_\text{q}$. This defines a set of possible distributions $\mathcal{N}(0, \Delta^2/12 + \sigma^2)$ for the correlation channel between X_q and Y. The quantized symbols are then mapped one to one in GF(q), thus giving the sequence $\overline{\mathbf{x}}_q$.

Then, $\overline{\mathbf{x}}_\text{q}$ is transmitted losslessly to the decoder, with the help of a SW chain realized with an LDPC code in GF(q). In universal SW coding, the infimum of achievable rates for the correlation model between X_q and Y is given by [5]

$$R^{SW}_{X_q|Y} = \sup_{\sigma^2 \in I_{\sigma^2}} H(X_\text{q}|Y, \sigma^2)\,. \quad (1)$$

It follows that the rate of the code and the LDPC coding matrix H, have to be chosen for the worst possible σ^2 in I_{σ^2}. This gives a codeword $\mathbf{u}$ of length $M < N$ with symbols in GF(q). At the decoder, one cannot efficiently use a classical min-sum LDPC decoder [16] to reconstruct $\overline{\mathbf{x}}_\text{q}$ from $\mathbf{u}$ and $\mathbf{y}$, since such a decoder requires the knowledge of the true parameter σ^2. Therefore, we propose to jointly estimate the source sequence $\overline{\mathbf{x}}_\text{q}$ and the parameter σ^2 from the codeword $\mathbf{u}$ and the side information sequence $\mathbf{y}$ with the use of an EM algorithm based on the classical LDPC decoder, giving $\widehat{\overline{\mathbf{x}}}_\text{q}$ and $\widehat{\sigma}^2$. The min-sum LDPC decoder in GF(q) is presented in Section 3.1 for known σ^2. The EM algorithm is presented in Section 3.2 as well as the proposed initialization method.

Then, each symbol in GF(q) of $\widehat{\overline{\mathbf{x}}}_\text{q}$ is mapped one to one into its real version, giving $\widehat{\mathbf{x}}_\text{q}$. Scalar dequantization is performed via MMSE estimation of $\mathbf{x}$ and uses $\widehat{\mathbf{x}}_\text{q}$, $\widehat{\sigma}^2$ and $\mathbf{y}$. For a known $\mathbf{x}_\text{q}$ and σ^2, as the sources are Gaussian, the MMSE estimator is linear and $\widehat{x}_n = \alpha y_n + \beta x_{\text{q},n}$, where $\alpha = \left(\frac{\sigma^2}{\Delta^2/12} + \frac{\sigma_x^2+\sigma^2}{\sigma_x^2}\right)^{-1}$ and $\beta = \frac{\sigma^2}{\Delta^2/12}\left(\frac{\sigma^2}{\Delta^2/12} + \frac{\sigma_x^2+\sigma^2}{\sigma_x^2}\right)^{-1}$.

3.1. Non-binary LDPC codes

LDPC codes are binary [9] or non-binary [6] linear error-correcting codes. In [14], they have been adapted to SW coding for binary sources with known correlation channel. This section presents an extension to the SW non-binary case when σ^2 is known. Most of the material presented in Section 3.1 was already given in [8] but we recall it here for the sake of clarity.

The SW coding of a vector $\mathbf{x}$ of length N is performed by producing a vector $\mathbf{u}$ of length $M < N$ as $\mathbf{u} = H^T\mathbf{x}$. The matrix H is sparse, with non-zero coefficients uniformly distributed in GF(q)$\backslash\{0\}$. In the following, $\oplus$, $\ominus$, $\otimes$, $\oslash$ are the usual operators in GF(q). In the bipartite graph representing the dependencies between the random variables of $\mathbf{X}$ and $\mathbf{U}$, the entries of $\mathbf{X}$ are represented by Variable Nodes (VN) and the entries of $\mathbf{U}$ are represented by Check Nodes (CN). The set of CN connected to a VN n is denoted $\mathcal{N}(n)$ and the set of VN connected to a CN m is denoted $\mathcal{N}(m)$. The sparsity of H is determined by the VN degree distribution $\lambda(x) = \sum_{i\geq 2} \lambda_i x^{i-1}$ and the CN degree distribution $\rho(x) = \sum_{i\geq 2} \rho_i x^{i-1}$ with $\sum_{i\geq 2}\lambda_i = 1$ and $\sum_{i\geq 2}\rho_i = 1$. In SW coding, the rate $r(\lambda, \rho)$ of a code is given by $r(\lambda, \rho) = \frac{M}{N} = \frac{\sum_{i\geq 2}\rho_i/i}{\sum_{i\geq 2}\lambda_i/i}$.

The decoder performs a Maximum *A Posteriori* (MAP) estimation of $\mathbf{x}$ from the received codeword $\mathbf{u}$ and the observed side information $\mathbf{y}$ via a Message Passing (MP) algorithm. The messages exchanged in the dependency graph are vectors of length q. The initial messages for each VN n are denoted $\mathbf{m}^{(0)}(n, y_n)$, with components

$$m_k^{(0)}(n, y_n) = \log \frac{P(X_n = 0|Y_n = y_n)}{P(X_n = k|Y_n = y_n)}\,. \quad (2)$$

With our model when σ^2 is known, this gives

$$m_k^{(0)}(n, y_n) = \log \frac{P(X_q = 0)}{P(X_q = k)} + \frac{(q_k - q_0)(q_k - q_0 - 2y_n)}{2\sigma^2} \quad (3)$$

where $P(X_q = k)$ can be calculated as

$$P(X_q = k) = \frac{1}{2}\left(\text{erf}\left(\frac{b_{k+1}}{2\sigma_x^2}\right) - \text{erf}\left(\frac{b_k}{2\sigma_x^2}\right)\right)\,. \quad (4)$$

The messages from CN to VN are computed with the help of a particular Fourier Transform (FT), denoted $\mathcal{F}(\mathbf{m})$. Denoting r the unit-root associated to GF(q), the i-th component of the FT is $\mathcal{F}_i(\mathbf{m}) = \sum_{j=0}^{q-1} r^{i\otimes j} e^{-m_j} / \sum_{j=0}^{q-1} e^{-m_j}$, see [13].

At iteration ℓ, the message $\mathbf{m}^{(\ell)}(m, n, u_m)$ from a CN m to a VN n is

$$\mathbf{m}^{(\ell)}(m, n, u_m) = \mathcal{A}[\overline{u}_m]\mathcal{F}^{-1}\left(\mathbf{M}^{(\ell)}(m, n)\right) \quad (5)$$

with

$$\mathbf{M}^{(\ell)}(m, n) = \prod_{n' \in \mathcal{N}(m)\backslash n} \mathcal{F}\left(W\left[\overline{H}_{n'm}\right]\mathbf{m}^{(\ell-1)}(n', m, y_{n'})\right) \quad (6)$$

where $\bar{u}_m = \ominus u_m \oslash H_{n,m}$, $\overline{H}_{n'm} = \ominus H_{n',m} \oslash H_{n,m}$ and $W[a]$ is a $q \times q$ matrix such that $W[a]_{k,n} = \delta(a \otimes n \ominus k), \forall 0 \leq k, n \leq q-1$. $\mathcal{A}[k]$ is a $q \times q$ matrix that maps a vector message $\mathbf{m}$ into a vector message $\boldsymbol{\ell} = \mathcal{A}[k]\mathbf{m}$ with $\ell_j = m_{j\oplus k} - m_k$. Note that the matrix $\mathcal{A}$ does not appear in the channel coding version of the algorithm and is specific to SW coding. At a VN n, a message $\mathbf{m}^{(\ell)}(n, m, y_i)$ is sent to the CN m and an *a posteriori* message $\tilde{\mathbf{m}}^{(\ell)}(n, y_n)$ is computed.

They both satisfy

$$\mathbf{m}^{(\ell)}(n,m,y_n) = \sum_{m'\in\mathcal{N}(n)\backslash m} \mathbf{m}^{(\ell)}(m',n,u_{m'}) + \mathbf{m}^{(0)}(n,y_n)\ ,$$
$$\tilde{\mathbf{m}}^{(\ell)}(n,y_n) = \sum_{m'\in\mathcal{N}(n)} \mathbf{m}^{(\ell)}(m',n,u_{m'}) + \mathbf{m}^{(0)}(n,y_n)\ . \tag{7}$$

From (7), each VN n produces an estimate of x_n as $\widehat{x}_n^{(\ell)} = \arg\max_k \tilde{m}_k^{(\ell)}(n,y_n)$. The algorithm ends if $\mathbf{u} = H^T\widehat{\mathbf{x}}^{(\ell)}$ or if $l = L_{\max}$, the maximum number of iterations.

3.2. EM algorithm

In our problem, the described LDPC decoder cannot be applied directly, because the initial messages (2) depend on the unknown σ^2. Consequently, we propose to jointly estimate $\mathbf{X}_q$ and σ^2 with an EM algorithm based on the described LDPC decoder. However, the main problem of the EM algorithm is its sensitivity to initialization. We thus provide a method to produce a first raw estimate of σ^2 in order to initialize the EM algorithm.

3.2.1. Initialization

The main idea of the initialization method is to produce a Maximum Likelihood (ML) estimate of σ^2 from $\mathbf{u}$ and $\mathbf{y}$ before using the LDPC decoder. Because of the encoding operation, each U_m can be seen as a sum of random variables $U_m = \sum_{j=1}^{\deg(m)} h_j^{(m)} X_j^{(m)}$, where $\deg(m)$ is the degree of the Check Node m and $h_j^{(m)}$ are the coefficients contained in H. Two assumptions are made in order to get an approximate version of the likelihood function $L(\sigma^2) = \log P(\mathbf{u},\mathbf{y}|\sigma^2)$. First, we assume that the random variables $X_j^{(m)}$ are independent $\forall j,m$. Second, we assume that $P(U_m|\mathbf{Y}) = P(U_m|\{Y_j^{(m)}\}_{j=1}^{\deg(m)})$, *i.e.*, each U_m depends only on the $Y_j^{(m)}$ associated to the $X_j^{(m)}$ composing the sum. Although these two assumptions are false in general, they may lead to a reasonable first raw estimate for σ^2. Indeed, with these assumptions, one just neglects part of the available information, *i.e.*, the dependencies between the U_m's and the dependencies between U_m and the remaining symbols of $\mathbf{Y}$. Using these assumptions, $L(\sigma^2)$ can be expressed as

$$L(\sigma^2) = \log P(\mathbf{u}|\mathbf{y},\sigma^2) + \log P(\mathbf{y}|\sigma^2) \tag{8}$$
$$= \sum_{m=1}^{M} \log P(u_m|\{y_j^{(m)}\}_{j=1}^{\deg(m)}) + \sum_{n=1}^{N} \log P(y_n|\sigma^2)\ .$$

Moreover, one can show that

$$P(u_m|\{y_j^{(m)}\}_{j=1}^{\deg(m)}) = \mathcal{F}_{u_m}^{-1}\Big(\prod_{j=1}^{\deg(m)} \mathcal{F}(W[h_j^{(m)}]\mathbf{p}_j)\Big) \tag{9}$$

with $\mathbf{p}_j = \big(P(X_j^{(m)}=0|y_j^{(m)})\dots P(X_j^{(m)}=q-1|y_j^{(m)})\big)$ and $\mathcal{F}_{u_m}^{-1}(.)$ is the u_m-th component of the inverse FT. $W[h_j^{(m)}]$ enables to evaluate the $P(h_j^{(m)}X_j^{(m)} = k|y_j^{(m)})$, $k = 0\dots q-1$. (9) follows from the fact that U_m is the sum of independent and identically distributed random variables of conditional distributions $P(h_j^{(m)}X_j^{(m)} = k|y_j^{(m)})$. Its probability is a convolution product, and can be efficiently evaluated in the Fourier domain from $\mathbf{p}_j$. Finally, the initial estimate $\widehat{\sigma}^{2(0)}$ is given by

$$\widehat{\sigma}^{2(0)} = \arg\max_{\sigma^2\in I_{\sigma^2}} L(\sigma^2)\ . \tag{10}$$

3.2.2. Iterations of the EM algorithm

Then, knowing some estimate $\widehat{\sigma}^{2(\ell)}$ obtained at iteration ℓ, the EM algorithm maximizes

$$\begin{aligned} Q(\sigma^2,\widehat{\sigma}^{2(\ell)}) =& E_{\mathbf{X}|\mathbf{y},\mathbf{u},\widehat{\sigma}^{2(\ell)}}\left[\log P(\mathbf{X}|\mathbf{y},\mathbf{u},\sigma^2)\right] \\ =& \sum_{\mathbf{x}\in\mathrm{GF}(q)^n} P(\mathbf{x}|\mathbf{y},\mathbf{u},\widehat{\sigma}^{2(\ell)}) \log P(\mathbf{y}|\mathbf{x},\mathbf{u},\sigma^2) \\ =& \sum_{n=1}^{N}\sum_{k=0}^{q-1} P(X_n=k|y_n,\mathbf{u},\widehat{\sigma}^{2(\ell)}) \\ & \log P(y_n|X_n=k,\sigma^2) \end{aligned} \tag{11}$$

with respect to σ^2. Denote $P_{k,n}^{(\ell)} = P(X_{q,n}=k|y_n,\mathbf{u},\widehat{\sigma}^{2(\ell)})$. These probabilities are the output of the LDPC decoder assuming that the true parameter is $\widehat{\sigma}^{2(\ell)}$. From (11) we show that the update equation is given by

$$\widehat{\sigma}^{2,(\ell+1)} = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=0}^{q-1} P_{k,n}^{(\ell)}(y_n-q_k)^2 \tag{12}$$

if $\widehat{\sigma}^{2,(\ell+1)} \in I_{\sigma^2}$, or by one of the bounds of I_{σ^2} if $\widehat{\sigma}^{2,(\ell+1)} \notin I_{\sigma^2}$. Moreover, at each iteration, an estimate of $\mathbf{x}_q^{(\ell)}$ can be obtained from

$$\widehat{x}_n^{(\ell)} = \arg\max_{k\in\mathrm{GF}(q)} P_{k,n}^{(\ell)}\ . \tag{13}$$

4. NUMERICAL RESULTS

For the simulations, $\sigma_x^2 = 1$, $N = 1000$, $\sigma^2 = 0.5$. The encoder and the decoder only know that $I_{\sigma^2} = [\underline{\sigma}^2,\overline{\sigma}^2] = [0,0.55]$. The number of experiments is choosen as $K = 100$.

We consider uniform scalar quantization over the interval $[-3\sigma_x, 3\sigma_x]$ and $q = 4, 8, 16$ and 32. Therefore, the quantization step is $\Delta = 6\sigma_x/q$. If a source symbol is outside $[-3\sigma_x, 3\sigma_x]$, its quantized version corresponds to the associated extreme quantization level. The quantizer induces some distortion D, which depends on q and on the true but unknown σ^2. The simulations have been carried out under a rate constraint achieved by the LDPC encoder. More precisely, for an LDPC over GF(q) with coding rate $r(\lambda,\rho)$, the overall coding rate (in bits/source symbol) is $R = r\log_2(q)$. In this section, we chose LDPC codes with $d_v = 2$ (this leads to good non binary codes) and tested the pairs $(q,r) = (4,2/3)$, $(8,2/3)$, $(16,3/4)$, $(32,4/5)$. These pairs have been determined to insure successful decoding in the worst case $\sigma^2 = \overline{\sigma}^2$. Each (q,r) is equivalent to a rate constraint, which induces a distortion D. The resulting set of points (R,D) correspond to the operational rate-distortion curve for uniform scalar quantization with LDPC codes.

Six setups are compared and the corresponding curves are plotted in Figure 3.

In the setups 1,2 and 3 we evaluate the rate-distortion performance of ideal Wyner-Ziv and SW schemes. Setup 1 (Ideal WZ) gives the rate-distortion bound $R(D) = \frac{1}{2}\log_2\frac{\sigma^2}{D}$ for the true parameter σ^2, corresponding to ideal Wyner-Ziv coding when σ^2 is

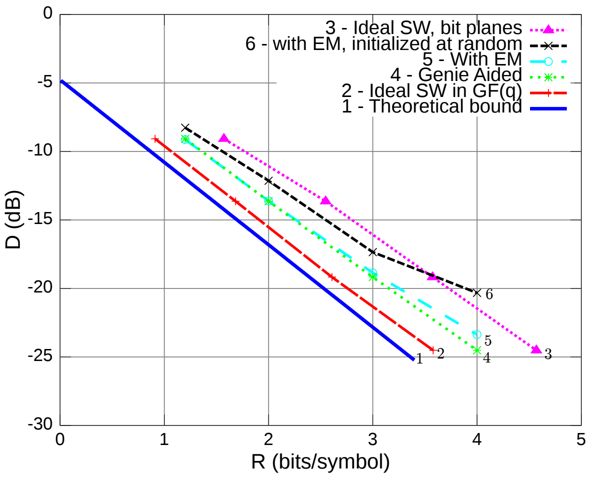

Fig. 3. Comparison of the rate-distortion performance of six schemes

known. Setup 2 (Ideal SW, GF(q)) assumes that σ^2 is known and that the coding scheme is composed by USQ + Ideal SW + MMSE estimator. The ideal SW rate corresponds to $H(X_q|Y)$, evaluated numerically, for the true σ^2, in order to evaluate the loss due to the USQ. The distortion is calculated as $D = \sum_{n=1}^{NK} \frac{(x_n^{(k)} - \widehat{x}_n^{(k)})^2}{KN}$. One can observe a loss of about 1.53 dB, as expected from [15] in the high rate regime. Setup 3 (Ideal SW, GF(q)) considers again USQ + ideal SW + MMSE estimator, except that the rate of the ideal SW is evaluated for binary codes when the quantized symbols are mapped in bit planes and the dependence between bit planes is not taken into account. An important loss is observed despite the ideal SW chain.

In the setups 4, 5 and 6, we evaluate the rate-distortion performance of coding schemes based on LDPC codes. In Setup 4 (genie-aided), the complete chain is implemented : USQ + non-binary LDPC coder and decoder + MMSE estimator, with 20 iterations of the LDPC decoder. The rate is dimensioned for $\overline{\sigma}^2$ but the value of σ^2 is assumed perfectly known at the decoder. A loss of about 2.5 dB is observed that is both due to the worst-case coding and to the fact that the non-binary LDPC codes we use are not optimized for this problem. Setup 5 (EM), is the proposed coding scheme with σ^2 unknown. The EM algorithm is initialized with the proposed method and has 4 iterations. If suffers only a little loss in the high-rate regime compared to Setup 4. In setup 6, the EM algorithm is initialized at random on I_{σ^2} and a loss of approximately 2 dB is observed.

5. CONCLUSION

In this paper, we propose a universal coding scheme for the Wyner-Ziv coding of Gaussian source with unknown variance of the correlation channel. Simulation results show that the proposed coding scheme suffers no loss compared to the case where the variance is known. The proposed scheme could be generalized to other signal models, such as Laplacian, or Bernoulli-Gaussian.

From the simulations, we observe that the loss is mainly due to the uniform scalar quantizer and to the choice of the LDPC codes. Consequently, future works will be dedicated to the construction of a nested quantizer and to the design of good non-binary LDPC codes for this problem.

6. REFERENCES

[1] A. Aaron, S. Rane, E. Setton, and B. Girod. Transform domain Wyner-Ziv codec for video. In *Proc of SPIE Vis. Com. and Im. Proc.*, pages 520–528, 2004.

[2] A. Aaron, R. Zhang, and B. Girod. Wyner-Ziv coding of motion video. In *Conference Record of the 36th Asilomar Conf. on Sig., Sys. and Comp.*, volume 1, pages 240–244, 2002.

[3] F. Bassi, M. Kieffer, and C. Weidmann. Wyner-Ziv Coding with uncertain side information quality. In *Proc. Eusipco 2010*, pages 2141–2145.

[4] T. Chu and Z. Xiong. Coding of Gauss-Markov sources with side information at the decoder. In *Proc. of WSSP 2009*, pages 26–29.

[5] I. Csiszar. Linear codes for sources and source networks: Error exponents, universal coding. *IEEE Trans. Inf. Th.*, 28(4):585–592, 1982.

[6] M. Davey and D. MacKay. Low Density Parity Check codes over GF (q). In *Proc. IEEE ITW*, pages 70–71, 1998.

[7] E. Dupraz, A. Roumy, and M. Kieffer. Source coding with side information at the decoder : models with uncertainty, performance bounds, and practical coding schemes. In *Proc of ISITA 2012*.

[8] E. Dupraz, A. Roumy, and M.Kieffer. Practical coding scheme for universal source coding with side information at the decoder. *Submitted to DCC*, 2013.

[9] R. Gallager. *Low-Density Parity Check Codes*. PhD thesis, Massachusetts Institute of Technology, 1963.

[10] R. Gallager. *Information theory and reliable communication*. Wiley, 1968.

[11] M. Gastpar, P. Dragotti, and M. Vetterli. The distributed Karhunen-Loeve transform. *IEEE Trans. Inf. Th.*, 52(12):1–10, 2006.

[12] S. Jalali, S. Verdú, and T. Weissman. A universal scheme for Wyner-Ziv coding of discrete sources. *IEEE Trans. Inf. Th.*, 56(4):1737–1750, 2010.

[13] G. Li, I. Fair, and W. Krzymien. Density evolution for nonbinary LDPC codes under Gaussian approximation. *IEEE Trans. Inf. Th.*, 55(3):997–1015, 2009.

[14] A. Liveris, Z. Xiong, and C. Georghiades. Compression of binary sources with side information at the decoder using LDPC codes. *IEEE Comm. Letters*, 6(10), 2002.

[15] D. Rebollo-Monedero, S. Rane, A. Aaron, and B. Girod. High-rate quantization and transform coding with side information at the decoder. In *EURASIP - Journal on Applied Sig. Proc. (Special Issue on Distributed Source Coding)*, volume 86, 2006.

[16] T. Richardson and R. Urbanke. The capacity of Low-Density Parity-Check codes under message-passing decoding. *IEEE Trans. Inf. Th.*, 47(2):599–618, 2001.

[17] A. Sgarro. Source coding with side information at several decoders. *IEEE Trans. Inf. Th.*, 23(2):179–182, 1977.

[18] A. Wyner and J. Ziv. The rate-distorsion function for source coding with side information at the decoder. *IEEE Trans. Inf. Th.*, 22(1):1–10, Jan 1976.