Short Length Non-binary Rate-Adaptive LDPC Codes for Slepian-Wolf Source Coding

Zeina Mheich[1,2], Elsa Dupraz[2]
[1]University of Surrey, United Kingdom; [2]IMT Atlantique, Lab-STICC, UBL, 29238 Brest, France

***Abstract*—In this paper, we consider the construction of a Slepian-Wolf source coding scheme in a context where only a small amount of data has to be transmitted to the decoder. In this context, we propose a novel rate-adaptive Slepian-Wolf code construction that is based on non-binary LDPC codes. The construction we propose replaces the regular accumulator of the standard LDPCA method by a local graph that is optimized at every rate of interest. In our method, the local graph is specifically designed in order to give good decoding performance at short length, while existing LDPCA constructions are usually optimized under an infinite codeword length assumption. Our simulation results on short codes obtained from our design method show a FER improvement of up to an order to magnitude compared to the standard LDPCA construction.**

I. Introduction

The problem of lossless source coding with side information available at the decoder only (see Figure 1) was initially introduced by Slepian and Wolf in [1]. This problem has recently regained increased attention due to its use in many modern multimedia applications such as sensor networks [2], Free-Viewpoint Television [3], or Massive Random Access (MRA) to data [4]. In these applications, the source X often represents the current data to be transmitted (e.g. the current view of a video), while the side information Y corresponds to the previously transmitted data (the previously transmitted views). It is therefore crucial to design highly efficient and reliable Slepian Wolf coding schemes for the above applications.

It is well known that practical Slepian-Wolf codes can be constructed from error-correction codes such as Low Density Parity Check (LDPC) codes [5]. Standard LDPC code constructions assume that the level of correlation between the source X and the side information Y is fixed. However, in the aforementioned applications, the correlation level can vary from one transmission to another. Several constructions such as Rateless codes [6] or LDPC Accumulator (LDPCA) codes [7]–[9] have then been proposed in order to adapt the coding rate depending on the correlation level.

LDPC codes and their rate-adaptive counterparts are known to achieve their best performance for very long codewords (more than 10000 bits). However, in a context such as sensor networks, only a small amount of data (less than 1000 bits) needs to be transmitted. The data collected by the sensors indeed corresponds to measurements of e.g. temperature or pressure. In order to be able to use very long codes in this context, one could consider grouping several successive frames of data. Unfortunately, this strategy may induce an important latency as well as increased memory requirements for the sensors. As an alternative, one may consider short codes that match the measurement vector length, at the price of a reasonable loss in performance [10]. For short frames of data, non-binary LDPC codes are known to be more efficient than binary LDPC codes [11].

The objective of this paper is thus to construct short-length rate-adaptive non-binary LDPC codes for Slepian-Wolf coding. Most of the constructions proposed for this problem consider binary codes optimized from design tools that assume asymptotic codeword length [6]–[9]. The solutions that were proposed for the design of non-binary codes [12], [13] are mainly direct applications of the methods proposed for binary codes. However, at short-length, the code performance may be degraded by short cycles, or by a poor choice of the non-binary labels associated to the edges of the Tanner graph of the code. The existing constructions optimized from asymptotic design tools do not permit to address these issues.

In this paper, we propose a novel method for the design of short-length rate-adaptive non-binary LDPC codes for Slepian-Wolf coding. More in details, we propose a generalized LDPCA construction that replaces the regular accumulator of the original constructions of [7], [13] by a local graph designed so as to optimize the code performance at short length for every rate of interest. In the paper, we describe this new method and we provide the design rules for the construction of the local graph. The design rules we propose permit to address: 1) the fine control of the connections between the variable nodes and the check nodes in the Tanner graphs used for the decoding at different rates, 2) the issue of cycles in these Tanner graphs, 3) the careful non-binary edge labeling. Monte Carlo simulations realized on short codes over GF(256) and GF(16) show a clear performance improvement of up to an order of magnitude compared to the non-binary rate-adaptive construction proposed in [13].

The outline of the paper is as follows. Section II describes the existing solutions for the design of rate-adaptive Slepian-Wolf codes. Section III introduces our novel method for the construction of non-binary rate-adaptive Slepian-Wolf codes. Section IV shows the simulation results.

II. LDPC codes for Slepian-Wolf source coding

In this paper, we consider the lossless coding of a source X when a side information Y is available at the decoder only, see Figure 1. The sources X and Y generate symbols in the Galois Field GF(q) with q elements, where q is a power of two. The non-binary source symbols may come either from non-binary measurements with q possible values (e.g. the pixels of

Fig. 1. Slepian-Wolf source coding

an image), or by grouping $\log_2(q)$ binary digits into a non-binary symbol in GF(q). We assume that the sources X and Y are correlated and that each of them generates sequences of independent and identically distributed (i.i.d.) symbols. The probability mass function of the source X is denoted as $P(X = a)$, $\forall a \in$ GF(q). The conditional probability distribution of Y knowing X is denoted as $P(Y = b|X = a)$, $\forall a, b \in$ GF(q).

The minimimum achievable rate for lossless Slepian-Wolf coding of the source X is given by the conditional entropy $H(X|Y)$ [1]. It was shown in [5] that LDPC codes permit to build a practical Slepian-Wolf coding scheme that achieves a rate close to $H(X|Y)$.

A. Slepian-Wolf coding from LDPC codes

Assume that the source X generates a vector of n source symbols $\mathbf{x}^n = (x_1, \cdots, x_n)$. Denote by H the parity check matrix of size $m \times n$ of an LDPC code [14]. The matrix H is sparse and the non-zero components of H take values in GF(q) $\setminus \{0\}$. The matrix H is equivalently represented by a bipartite Tanner graph $\mathcal{T} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ defines the set of vertices, and $\mathcal{E}$ defines the set of edges in the graph. The set $\mathcal{V} = (\mathcal{X} \cup \mathcal{S})$ is composed by n variable nodes $\mathcal{X} = \{X_1, \cdots, X_n\}$ and m check nodes $\mathcal{S} = \{S_1, \cdots, S_m\}$. The edges in the set $\mathcal{E}$ are given by the non-zero positions in the parity check matrix H and they are labeled by the values at the corresponding positions in H.

In order to compress the source sequence $\mathbf{x}^n$, the encoder produces a codeword

$$\mathbf{s}^m = H(\mathbf{x}^n)^T \tag{1}$$

where T is the transpose operator and all the arithmetic operations are performed in the Galois Field GF(q). The codeword $\mathbf{s}^m$ is composed by m symbols in GF(q) and it is assumed to be transmitted without error to the decoder. The coding rate is hence given by $R = m/n$. The decoder has to estimate the source vector $\mathbf{x}^n$ from the received codeword $\mathbf{s}^m$ and from the side information vector $\mathbf{y}^n = (y_1, \cdots, y_n)$. For the estimation of $\mathbf{x}^n$, we rely on a standard non-binary Belief Propagation (BP) decoder [15] that is initialized with the side information vector $\mathbf{y}^n$.

In the above scheme, both the rate R and the parity check matrix H are designed so as to ensure good decoding performance for a fixed statistical model between X and Y. Consider for example a q-ary symmetric correlation channel of parameter p between X and Y. Since R and H are designed for a given p, a change in this value will result either in a rate loss if p is decreased, or even worse in a decoding failure if p is increased. This is why several rate-adaptive LDPC code constructions have been proposed in order to deal with a change in the correlation model between X and Y.

B. Rate-adaptive LDPC codes for Slepian-Wolf coding

Two main solutions were proposed to construct binary rate-adaptive Slepian-Wolf codes. Rateless codes proposed in [6] start from a matrix H with a low rate R. If a rate higher than R is required, the encoder transmits both the syndrom $\mathbf{s}^m$ and additional linear combinations of the source symbols $(x_1, \cdots, x_n)$. Unfortunately, this approach has difficulties to achieve good performance at very low rates [16].

As a reverse solution, LDPCA codes proposed in [7], [8] start with a high rate matrix H. In the LDPCA construction, the check nodes values $(s_1, \cdots, s_m)$ are first accumulated by calculating new parity equations $(a_1, \cdots, a_m)$ as

$$\begin{aligned} &a_1 = s_1, \\ &a_i = a_{i-1} + s_i, \quad \forall i \in \{2, \cdots, m\}. \end{aligned} \tag{2}$$

If a rate lower than R is required, only a part of the symbols $(a_1, \cdots, a_m)$ is transmitted to the decoder.

In order to improve the binary LDPCA construction, it was proposed in [9] to consider a non-regular accumulator optimized for any rate of interest. The optimization method proposed in [9] is based on a so-called density evolution analysis that assumes asymptotic codeword length. Unfortunately, in our context, the density evolution analysis may not predict accurately the performance of short-length codes which can be degraded for example by short cycles in the Tanner graph.

A rate-adaptive LDPC code construction was further proposed in [13] in order to deal with non-binary source symbols. The design method proposed in [13] starts with an intermediate rate, usually $R = 1/2$. The rates higher than $1/2$ are obtained from the solution of [6] by revealing a part of the non-binary source symbols to the decoder. The rates lower than $1/2$ are obtained from a non-binary LDPCA construction. The method of [13] however straightforwardly applies the LDPCA construction of [7] without optimizing the accumulator, which lowers its performance. In this paper, we thus propose a novel non-binary LDPCA construction that greatly improves the performance of the coding scheme of [13]. Compared to the approach of [9], our design method is based on a finite length performance evaluation rather than on density evolution, and it is thus well adapted to the construction of short length LDPC codes.

III. Proposed Method

This section describes our design method for the construction of short-length rate-adaptive non-binary Slepian-Wolf codes. As in [13], we start with an LDPC code of rate $R = 1/2$ called the mother code. For rates higher than $1/2$, we rely on the source revealing method proposed in [13] which is already shown to be efficient for high rates (out of the scope of the paper). For rates lower than $1/2$, we propose a new construction which we present in this section. We first describe our construction for only one rate lower than $R = 1/2$, we then generalize this construction to several rates.

A. Mother code

In our method, the mother code consists of a matrix H_1 of size $m_1 \times n$ with coding rate $R_1 = m_1/n = 1/2$. The

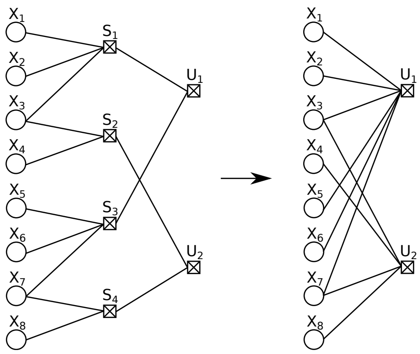

Fig. 2. The left part of the figure shows the combination of $\mathcal{T}_1$ with $\mathcal{T}_{1\to 2}$. The right part of the figure shows the resulting $\mathcal{T}_2$. Here, the matrix $H_{1\to 2}$ is full rank, and one may choose $\mathcal{S}' = \{S_1, S_2\}$, $\mathcal{S}' = \{S_3, S_4\}$, $\mathcal{S}' = \{S_1, S_4\}$, or $\mathcal{S}' = \{S_2, S_3\}$.

Tanner graph $\mathcal{T}_1 = (\mathcal{X} \cup \mathcal{S}, \mathcal{E}_1)$ associated to H_1 connects the n variable nodes $\mathcal{X} = \{X_1, \cdots, X_n\}$ to m_1 check nodes $\mathcal{S} = \{S_1, \cdots, S_{m_1}\}$. For the construction of H_1, we apply a standard non-binary LDPC code design method that consists of three steps. We first select a code degree distribution or a protograph that maximizes the code ensemble threshold [17]. We then construct a binary parity check matrix of the selected code ensemble from a progressive edge growth (PEG) algorithm that permits to avoid as much as possible short cycles in the Tanner graph of the code [18]. To finish, we assign non-binary labels to all the non-zero components of the parity check matrix of the code. The non-binary labels are chosen so as to maximize the minimum distance of the binary image of each parity check equation [11]. In this paper, we do not give more details on the construction of H_1 from this procedure. The code construction we propose indeed works for any choice of H_1.

B. Daughter code

From the mother matrix H_1, we want to construct a daughter matrix H_2 of size $m_2 \times n$, with $m_2 < m_1$, and rate $R_2 = m_2/n < 1/2$. The Tanner graph $\mathcal{T}_2 = (\mathcal{X} \cup \mathcal{U}, \mathcal{E}_2)$ associated to H_2 will connect the n variable nodes $\mathcal{X} = \{X_1, \cdots, X_n\}$ to m_2 check nodes $\mathcal{U} = \{U_1, \cdots, U_{m_2}\}$. In order to obtain the daughter matrix H_2 from the mother matrix H_1, we construct an intermediate matrix $H_{1\to 2}$ of size $m_2 \times m_1$. The Tanner graph $\mathcal{T}_{1\to 2} = (\mathcal{S} \cup \mathcal{U}, \mathcal{E}_{1\to 2})$ of $H_{1\to 2}$ connects the check nodes $\mathcal{S}$ of $\mathcal{T}_1$ to the check nodes $\mathcal{U}$ of $\mathcal{T}_2$. With this construction, the matrix H_2 of rate R_2 is then equal to

$$H_2 = H_{1\to 2} H_1. \tag{3}$$

Figure 2 shows an example of the construction of $\mathcal{T}_2$ from $\mathcal{T}_1$ and $\mathcal{T}_{1\to 2}$. Note that LDPCA codes can be seen as a particular case of this construction.

At the end, the matrix $H_{1\to 2}$ should be chosen not only to give a good decoding performance for H_2, but also to allow H_1 and H_2 to be rate-adaptive in a sense we now define.

C. Rate-adaptive condition

In our construction, we set the following transmission rule in order to allow H_1 and H_2 to be rate-adaptive. In order to get a rate R_2, we simply transmit all the parity check values defined by the set $\mathcal{U}$, which corresponds to m_2 equations. The decoding is then realized with the matrix H_2. In order to get a rate R_1, we transmit all the parity check values in $\mathcal{U}$ but also a subset $\mathcal{S}' \subseteq \mathcal{S}$ of size $m_1 - m_2$ of the values in $\mathcal{S}$. In order to use the matrix H_1 for decoding, the receiver must be able to recover the values in $\mathcal{S}$ from $\mathcal{U}$ and $\mathcal{S}'$. The code that results from the choice of (H_1, $H_{1\to 2}$, $\mathcal{S}'$) is then said to be rate-adaptive if is satisfies the following condition.

Definition 1. *$\mathcal{U}$ and $\mathcal{S}'$ define a system of m_1 equations with m_1 unknown variables $\mathcal{S}$. If this system has a unique solution in GF(q), then the triplet (H_1, $H_{1\to 2}$, $\mathcal{S}'$) is said to be a rate-adaptive code.*

This condition is equivalent to the 1-SR condition formalized in [19] and considered in [20] in the context of channel coding. Note that the standard LDPCA construction satisfies the above rate-adaptive condition. The following proposition gives a simple condition that permits to verify whether a given choice of $H_{1\to 2}$ can lead to a rate-adaptive code.

Proposition 1. *If the matrix $H_{1\to 2}$ is full rank, then there exists a set $\mathcal{S}' \subseteq \mathcal{S}$ of size $m_1 - m_2$ such that (H_1, $H_{1\to 2}$, $\mathcal{S}'$) is a rate-adaptive code.*

Proof: The variables $U_1, \cdots, U_{m_2}$ define m_2 equations of m_1 variables $S_1, \cdots, S_{m_1}$. Consider $m_1 - m_2$ additional virtual equations $U_{m_2+1}, \cdots, U_{m_1}$ such that for all $i \in \{m_2 + 1, \cdots, m_1\}$, U_i equals one of the variables $S_1, \cdots, S_{m_1}$. Also define an extended matrix $\tilde{H}_{1\to 2}$ of size $m_1 \times m_1$. The first m_2 rows of $\tilde{H}_{1\to 2}$ are given by $H_{1\to 2}$. The last $m_1 - m_2$ rows will define to which variable of $\mathcal{S}$ each U_i ($i > m_2$) is equal. Since $H_{1\to 2}$ is full rank, we can always identify m_2 linearly independent columns in $H_{1\to 2}$. For each of the $m_1 - m_2$ remaining columns, it is always possible to put a value 1 in one of the last $m_1 - m_2$ rows of $\tilde{H}_{1\to 2}$ such that each of these rows contains only one non-zero value. This guarantees that $\tilde{H}_{1\to 2}$ is full rank. The positions of the $m_1 - m_2$ remaining columns in fact defines a possible set $\mathcal{S}'$, which concludes the proof. ∎

The above proposition shows that if $H_{1\to 2}$ is full rank, it is always possible to find a set $\mathcal{S}'$ that ensures that H_1 and H_2 are rate-adaptive. As a consequence, as long as $H_{1\to 2}$ is full-rank, the decoding performance does not depend on the choice of the set $\mathcal{S}'$, since at rate R_1, the decoder uses H_1 and at rate R_2, the decoder uses H_2. On the contrary, according to equation (3), the decoding performance of the matrix H_2 still heavily depends on the choice of the matrix $H_{1\to 2}$.

D. Construction of the matrix $H_{1\to 2}$

In order to get a good decoding performance for H_2, we can play on two degrees of freedom for $H_{1\to 2}$. These two degrees of freedom are the construction of the set of edges $\mathcal{E}_{1\to 2}$ in the Tanner graph $\mathcal{T}_{1\to 2}$, and their non-binary labeling. We now discuss the constraints we set for the optimization of these two degrees of freedoms.

1) Set of edges: We impose two constraints for the construction of the set of edges $\mathcal{E}_{1\to 2}$. The first constraint is that every node in $\mathcal{U}$ must be connected to at least one node in $\mathcal{S}$, and conversely. If this constraint is not satisfied, the matrix $H_{1\to 2}$ cannot be full rank. The second constraint is that $\mathcal{T}_{1\to 2}$ does not contain any cycle (see Figure 2 for an example). The rationale of this constraint is to avoid as much as possible adding cycles in the resulting H_2.

2) Non-binary labeling: Once the set of edges $\mathcal{E}_{1\to 2}$ is defined, we choose the non-binary labels of the edges as follows. For each row of $H_{1\to 2}$, we choose a combination of labels that maximizes the minimum distance of the binary image of the corresponding parity check equation [11]. There are several combinations of labels that satisfy the previous condition. We verify that the selected combination: 1) does not make a column of H_2 empty (each variable node in H_2 is connected to at least one check node), 2) does not give twice the same coefficient in the same row of H_2 (this could degrade the performance of H_2), 3) produces a full rank matrix $H_{1\to 2}$.

From [11] and from Proposition 1, we know that once it satisfies the above constraints, the labeling will not change much the decoding performance. This is why, for each possible set of edges $\mathcal{E}_{1\to 2}$, we choose at random one labeling that satisfies the above conditions. If, for a given set $\mathcal{E}_{1\to 2}$, no labeling satisfying these conditions exists, the considered set $\mathcal{E}_{1\to 2}$ is discarded. We then select the set of edges $\mathcal{E}_{1\to 2}$ that, combined with the labeling chosen for this set of edges, gives the best matrix H_2 (verified from Monte Carlo simulations).

The number of possible sets of edges $\mathcal{E}_{1\to 2}$ is rather small due to the constraints we impose on these sets. In particular, discarding the sets of edges that contain cycles does not allow high node degrees in the resulting Tanner graph. The number of possible sets $\mathcal{E}_{1\to 2}$ is also reduced by the fact that we consider short codes. This is why, in our code construction process, we fixed a maximum node degree of three and we selected the set of edges by a random search over the admissible sets $\mathcal{E}_{1\to 2}$. The performance of each considered set $\mathcal{E}_{1\to 2}$ was evaluated from Monte Carlo simulations over the resulting H_2. More efficient methods will be developed in future works.

At the end, since the retained $H_{1\to 2}$ is guaranteed to be full-rank, we know that there exists a set $\mathcal{S}'$ that permits the rate adaptation between H_1 and H_2. Since the choice of this set does not affect the decoding performance, we choose any set $\mathcal{S}'$ that gives a (H_1, $H_{1\to 2}$, $\mathcal{S}'$) rate-adaptive code.

E. Generalization to several rates

The above construction permits to obtain the matrix H_2 of rate $R_2 < R_1$ from the matrix H_1. In order to obtain lower rates $R_I < R_{I-1} < \cdots < R_2 < R_1$, we need to construct the corresponding matrices H_i, $i = 2, \cdots, I$. The successive matrices R_i are constructed recursively from the above method for two rates. More precisely, the matrix H_i is obtained from H_{i-1} by constructing an intermediate matrix $H_{i-1\to i}$ such that $H_i = H_{i-1\to i}H_{i-1}$. The intermediate matrix $H_{i-1\to i}$ is constructed following the design rules of Section III-D.

We now consider two constructions of rate-adaptive codes in GF(256) and in GF(16), and we evaluate their performance for different rates.

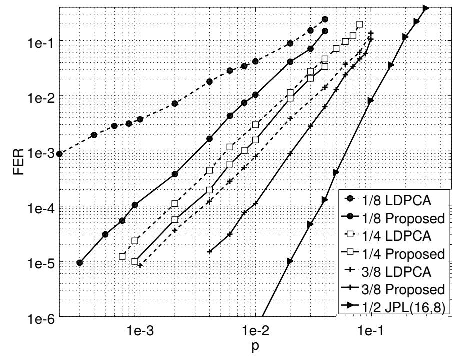

Fig. 3. FER performance in GF(256) of the proposed construction compared with LDPCA

IV. SIMULATION RESULTS

This section evaluates the performance of our design method compared to the standard LDPCA construction. In our simulations, we assume that the source X is distributed uniformly. The correlation between X and Y is modeled by the following q-ary symmetric channel

$$\forall a \in \text{GF}(q),\ P(Y = a|X = a) = 1 - p$$
$$\forall a, b \in \text{GF}(q), a \neq b,\ P(Y = b|X = a) = \frac{p}{q-1}. \quad (4)$$

Note that this model is considered here as a test case for the simulations, although the proposed construction applies whatever the statistical model between X and Y. In this section, we consider short codes of length $N = 128$ bits and source symbols in GF(256) or in GF(16).

A. Code construction in GF(256)

We first consider source symbols X in GF(256). We consider a code of length $N = 128$ bits, which corresponds to 16 source symbols in GF(256). For the mother code of rate $R = 1/2$, we use the already well designed $(16, 8)$ JPL matrix given in [21]. We then apply the method described in Section III in order to construct all the lower rates $R_i = i/16$, $i = 2 \cdots, 7$.

We want to evaluate the performance of our method compared to the standard LDPCA construction proposed in [13] for non-binary LDPC codes. In [13] the non-binary coefficients in the accumulator were chosen uniformly at random, despite the fact that this choice may negatively impact the code performance. This is why, here, in order to allow a fair comparison with our method, the LDPCA construction was realized by choosing the non-binary labels in the accumulator so as to maximize the minimum distance of the binary images of the parity check equations. This labelling comes for the method of [11] for the construction of non-binary LDPC channel codes. It is also employed for the labeling in our rate-adaptive method, see Section III-D.

Figure 3 represents the obtained Frame Error Rates (FER) for the LDPCA construction and for our design method for four rates $1/8$, $1/4$, $3/8$, and $1/2$. Note that we represented only three rates lower than $1/2$ in order not to overload the Figure. At all the considered rates, even those that are not shown in the figure, we see a clear gain of sometimes an order

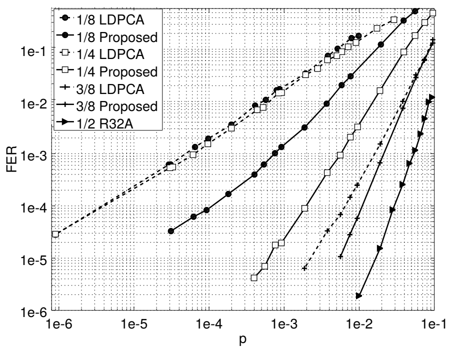

Fig. 4. FER performance in GF(16) of the proposed construction compared with LDPCA

of magnitude compared to LDPCA (except for the rate $3/16$, which shows a slight loss in performance). This gain does not come from the non-binary labelling since we have employed the same labeling both for our method and for the LDPCA construction. This gain most surely comes from the fact that our construction is optimized from design rules that lead to a good performance at finite length for each considered rate.

B. Code construction in GF(16)

We now consider source symbols in GF(16) and a code of length $N = 128$ bits, which gives 32 source symbols in GF(16). In order to obtain the mother code of rate $R = 1/2$, we followed the code construction procedure of [11]. More in details we constructed a Quasi-Cyclic code from the R32A protograph

$$B = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{5}$$

by applying a two-steps lifting described in [22], with a first lifting factor of 4 and a second lifting factor of 2. We then applied the method described in Section III in order to construct the three lower rates $1/8$, $1/4$, and $3/8$. We compared the performance of our method to the performance of the LDPCA construction with a carefully chosen labeling as we did for the code in GF(256).

Figure 4 represents the obtained Frame Error Rates (FER) for the LDPCA construction and for our design method for the three considered rates. In this case again, our method gives better performance than the LDPCA construction, whatever the considered rate. Rates $1/8$ and $1/4$ exhibit a gain of more than an order of magnitude compared to LDPCA, which shows that our method is particularly efficient for very low rates.

These two sets of simulations permit to conclude that our construction method outperforms the LDPCA solution by specifically optimizing the codes performance at short length.

V. Conclusion

In this paper, we proposed a novel rate-adaptive construction for short-length non-binary Slepian-Wolf codes. The construction we proposed permits to optimize the code performance at short length for all the rates of interest. Simulation results confirmed the efficiency of the proposed approach compared to the standard LDPCA construction. Future works will be dedicated to the development of more efficient methods in order to construct the successive codes in a more systematic way.

VI. Acknowledgement

This work has received a French government support granted to the Cominlabs excellence laboratory and managed by the National Research Agency in the Investing for the Future program under reference ANR-10-LABX-07-01.

References

[1] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *IEEE Transactions on information Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[2] J. Zheng, P. Wang, and C. Li, “Distributed data aggregation using Slepian-Wolf coding in cluster-based wireless sensor networks,” *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2564–2574, 2010.

[3] A. Roumy and T. Maugey, “Universal lossless coding with random user access: the cost of interactivity,” in *IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 1870–1874.

[4] E. Dupraz, T. Maugey, A. Roumy, and M. Kieffer, “Rate-storage regions for massive random access,” *arXiv preprint arXiv:1612.07163*, 2016.

[5] A. D. Liveris, Z. Xiong, and C. N. Georghiades, “Compression of binary sources with side information at the decoder using LDPC codes,” *IEEE Communications Letters*, vol. 6, no. 10, pp. 440–442, 2002.

[6] A. W. Eckford and W. Yu, “Rateless Slepian-Wolf codes,” in *Asilomar conference on signals, systems and computers*, 2005, pp. 1757–1761.

[7] D. Varodayan, A. Aaron, and B. Girod, “Rate-adaptive codes for distributed source coding,” *Elsevier Signal Processing*, vol. 86, no. 11, pp. 3123–3130, 2006.

[8] F. Cen, “Design of degree distributions for LDPCA codes,” *IEEE Communications Letters*, vol. 13, no. 7, 2009.

[9] C. Yu and G. Sharma, “Improved low-density parity check accumulate (LDPCA) codes,” *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3590–3599, 2013.

[10] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[11] C. Poulliat, M. Fossorier, and D. Declercq, “Design of regular (2, dc)-LDPC codes over GF (q) using their binary images,” *IEEE Transactions on Communications*, vol. 56, no. 10, 2008.

[12] Z.-l. Wang, X. Li, and J. Li, “Q-LDPCA based distributed coding for asymmetrical correlated non-binary sources,” in *9th International Conference on Signal Processing (ICSP)*, 2008, pp. 1842–1845.

[13] K. Kasai, T. Tsujimoto, R. Matsumoto, and K. Sakaniwa, “Rate-compatible Slepian-Wolf coding with short non-binary LDPC codes,” in *Data Compression Conference (DCC)*, 2010, pp. 288–296.

[14] T. J. Richardson and R. L. Urbanke, “The capacity of low-density parity-check codes under message-passing decoding,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 599–618, 2001.

[15] E. Dupraz, A. Roumy, and M. Kieffer, “Source coding with side information at the decoder and uncertain knowledge of the correlation,” *IEEE Transactions on Communications*, vol. 62, no. 1, pp. 269–279, 2014.

[16] S.-Y. Shin, M. Jang, S.-H. Kim, and B. Jeon, “Design of binary LDPCA codes with source revealing rate-adaptation,” *IEEE Communications Letters*, vol. 16, no. 11, pp. 1836–1839, 2012.

[17] G. Li, I. J. Fair, and W. A. Krzymien, “Density evolution for nonbinary LDPC codes under Gaussian approximation,” *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 997–1015, 2009.

[18] X.-Y. Hu, E. Eleftheriou, and D.-M. Arnold, “Regular and irregular progressive edge-growth tanner graphs,” *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 386–398, 2005.

[19] J. Ha, J. Kim, D. Klinc, and S. W. McLaughlin, “Rate-compatible punctured low-density parity-check codes with short block lengths,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 728–738, 2006.

[20] S. Pfletschinger, D. Declercq, and M. Navarro, “Adaptive HARQ with non-binary repetition coding,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4193–4204, 2014.

[21] Orange Book, “Short Blocklength LDPC codes for TC synchronization and channel coding,” 2012.

[22] D. G. Mitchell, R. Smarandache, and D. J. Costello, “Quasi-cyclic LDPC codes based on pre-lifted protographs,” *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 5856–5874, 2014.