A new approach to optimise Non-Binary LDPC codes for Coded Modulations

Ahmed Abdmouleh[#], Emmanuel Boutillon[#], Laura Conde-Canencia[#],
Charbel Abdel Nour[*] and Catherine Douillard[*]

[#]Université de Bretagne Sud, Lab-STICC (UMR 6582), 56321 Lorient, France
[*]Institut Mines-Telecom; Telecom Bretagne; Lab-STICC (UMR 6582), Technopole Brest-Iroise, 29238 Brest, France
[#] ahmed.abdmouleh,emmanuel.boutillon,laura.conde-canencia@univ-ubs.fr
[*] charbel.abdelnour,catherine.douillard@telecom-bretagne.eu

***Abstract*—This paper is dedicated to the optimisation of Non-Binary LDPC codes when associated to high-order modulations. To be specific, we propose to specify the values of the non-zero NB-LDPC parity matrix coefficients depending on the corresponding check node equation and the Euclidean distance of the coded modulation. In other words, we explore the joint optimisation of the modulation mapping and the non-binary matrix. The performance gains announced by a theoretical analysis based on the Union Bound are confirmed by simulations results. We obtain an 0.2-dB gain in the high SNR regime compared to other state-of-the-art matrices.**

I. Introduction

Since their rediscovery in 1996, Low-Density Parity-Check (LDPC) codes designed over GF(2) have shown performance close to the Shannon limit for long code lengths [1] [2]. For moderate or small lengths, error performance can be improved by extending LDPC codes to high-order Galois Fields GF(q), $q > 2$ [3]. These so-called Non-Binary (NB) LDPC codes retain the benefits of steep waterfall region (typical of convolutional turbo-codes) and low error floor (typical of binary LDPC). Compared to their binary counterparts, NB-LDPC codes generally present higher girths, which leads to better decoding performance. Different works have also revealed the interest of NB-LDPC in MIMO systems ([4] [5] [6]).

Another advantage of NB-LDPC codes concerns their association with high-order q-ary modulations: by encoding directly over the q-ary constellation alphabet, binary-to-NB mapping and demapping operations are not needed, unlike for binary codes. In other words, NB symbol likelihoods are calculated directly and input to the NB decoder, without any marginalization [7]. Note that the demapping operation is costly in terms of complexity and introduces performance loss that would have to be partially countered by a proper choice of mapping or fully recovered by costly iterations over the demapper and decoder. For these reasons, NB-LDPC codes constitute a promising solution for high spectral efficiency coding, even if they present the drawback of high decoding complexity [8].

For finite code lengths, the construction of NB-LDPC matrices is generally solved in two different steps [9]. First, the positions of the non-zero entries of the parity check matrix $\mathcal{H}$ are optimised in order to maximize the girth of the code and minimize the impact of cycles when using the Belief Propagation (BP) algorithm on the associated Tanner graph. In [10] [11], it is widely accepted that good graphical codes have large girth and a small number of short cycles. This optimisation can be efficiently achieved with the Progressive Edge Growth (PEG) algorithm [12] or one of its variants. The second step in the matrix construction consists in choosing the values of the non-zero entries. This can be done either randomly from a uniform distribution (among the non-zero elements of GF(q)) [12] or carefully to meet some design criteria [13].

In [9] the problem of the selection and matching of the parity-check matrix non-zero entries with the code was considered. The authors proposed to optimise the position of the non-zero entries based on the binary image representation of matrix $\mathcal{H}$ and to maximise the minimum Hamming distance of the binary image of the code. Also, in [9], the authors showed the interest of NB-LDPC codes with minimum connectivity on the symbol nodes $d_v = 2$, where d_v represents the variable node degree.

However, when considering the association of NB codes and q-ary modulations, mapping symbols of a code optimised for Hamming distance into NB modulation signals does not guarantee that a good Euclidean distance structure is obtained. Squared Euclidean and Hamming distances are equivalent only in the case of binary modulation or four-phase modulation. Binary modulation systems with codes optimised for Hamming distance and soft decision decoding have been well established since the late 1960s for power-efficient transmission at spectral efficiencies of less than 2 bit/sec/Hz. For higher spectral efficiencies, the association of powerful error-correcting codes and high-order modulation has been largely considered in the literature, see for example [14] [15] [16] [17], among many others.

In this paper we propose to optimise the NB-LDPC coded modulation through a matrix optimisation that is aware of the modulation and mapping. In others words, the optimisation criterion is not the Hamming distance of the binary image

978-1-5090-3401-7/16/.00 ©2016 IEEE

of the code as in [12] [13] [9], but the Euclidean distance of the modulated codewords. To the best of our knowledge, this problem has never been considered in the literature for NB-LDPC codes. Our approach assumes that the positions of non-zero entries in the parity check matrix $\mathcal{H}$ are already defined and we focus on the determination of the values of non-zero entries when the NB-LDPC code over GF(q) is directly associated to a q-ary Quadrature Amplitude Modulation (QAM).

The paper is organised as follows: Section II presents notations and some key definitions to describe our work. Section III introduces a theoretical analysis of the decoding performance based on the Union Bound. The evaluation of the distance spectrum of a code is considered in section IV where we also introduce a method to simplify the calculation of the Union Bound. We then consider in section V the properties of the Euclidean distance for Gray-mapped QAM modulations. Our contribution is then described in Section VI where we optimise the distance spectrum of non-binary coded modulations leading to better results than those in the state-of-the-art. Section VII presents simulation results to show the interest of our approach. Finally, section VIII concludes the paper.

II. Notations and definitions

Let us define a (N, K) NB-LDPC code over GF(q) with code length N and information length K. Its parity check matrix $\mathcal{H}$ has N columns and $N-K$ rows. The code is assumed to be regular where each row has a number of non-zero entries equal to d_c and each column has $d_v = 2$ non-zero entries. Assuming a full rank matrix, the code rate is given by $r = 1 - d_v/d_c$. The non-zero entries of $\mathcal{H}$ are denoted by $h_{m,n}$, where m is the row index and n the column index.

We assume that each symbol in GF(q) is associated with an element of the q-ary constellation $\mathcal{M}$ through a mapping function π: GF(q) $\rightarrow \mathbb{R}^2$ such that for each $x \in$ GF(q), $\pi(x) = (\pi_I(x), \pi_Q(x)) \in \mathbb{R}^2$. Note that $\pi_I(x)$ and $\pi_Q(x)$ represent the in-phase and the quadrature amplitudes of the modulated signal, respectively. For example, if $\mathcal{M}$ is a 64-QAM, then both $\pi_I(x)$ and $\pi_Q(x)$ belong to the set $\{-7, -5, -3, -1, 1, 3, 5, 7\}$. Figure 1 provides three different mappings that are considered in our study for the 64-QAM.

The propose approach consists in optimising the coefficients of one row of the matrix [13], i.e. a single parity check equation of the code expressed as:

$$\sum_{k=1}^{d_c} h_k x_k = 0, \tag{1}$$

where $h_k \in$ GF(q), $k = 1 \ldots d_c$, are the non-zero entries and $x_k \in$ GF(q), $k = 1 \ldots d_c$ are the d_c variables of the parity check equation. Note that this equation defines a code over GF$(q)^{d_c}$ and that we denote by $\mathcal{C}$ the set of d_c-uple $\mathbf{x} = (x_k)_{k=1\ldots d_c}$ in GF$(q)^{d_c}$ that verify (1).

Let the Euclidean distance $D(x, y)$ between two elements x and y in GF(q) be the Euclidean distance between $\pi(x)$ and $\pi(y)$ in $\mathbb{R}^2$. Note that for a given modulation, $D(x, y)$ depends on the mapping function π. The Squared Euclidean distance $D^2(x, y)$ is then expressed as:

$$D^2(x, y) = |\pi_I(x) - \pi_I(y)|^2 + |\pi_Q(x) - \pi_Q(y)|^2. \tag{2}$$

Let us also define the Euclidean distance between two codewords $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}^2$ as:

$$D^2_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{d_c} D^2(x_k, y_k). \tag{3}$$

Finally, let the Distance Spectrum (DS) of a code $\mathcal{C}$ be the enumeration of all the possible distances between two different codewords and the number of distinct ordered couples at each of those distances. The DS function $S_{\mathcal{C}}(d)$ can then be defined as:

$$S_{\mathcal{C}}(d) = |\{(\mathbf{x}, \mathbf{y}) \in \mathcal{C}^2, D^2_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) = d^2\}|. \tag{4}$$

where $|.|$ represents the cardinality of a set.

III. Decoding performance of the elementary check node

The d_c symbols of an element $\mathbf{x} \in \mathcal{C}$ are transmitted through an Additive White Gaussian Noise (AWGN) channel. The received message is thus $\mathbf{r} = \pi(\mathbf{x}) + \mathbf{w}$ where $\mathbf{w}$ is a complex vector of size d_c, with each coordinate being the realization of a complex Gaussian noise of variance $\sigma^2 = N_0/2$, where N_0 is the power spectral density of the AWGN. For a given Signal-to-Noise Ratio (SNR), the probability $P(\mathbf{x} \rightarrow \mathbf{y})$ of transmitting $\mathbf{x}$ and decoding $\mathbf{y} \neq \mathbf{x}$ when using a Maximum Likelihood (ML) decoder is given by:

$$P(\mathbf{x} \rightarrow \mathbf{y}) = \text{Prob}(||\mathbf{r} - \pi(\mathbf{x})||^2 > ||\mathbf{r} - \pi(\mathbf{y})||^2). \tag{5}$$

Since we consider the AWGN channel, this probability can be expressed as:

$$P(\mathbf{x} \rightarrow \mathbf{y}) = Q\left(\frac{D_{\mathcal{C}}(\mathbf{x}, \mathbf{y})}{2\sigma}\right), \tag{6}$$

where $Q(u)$ is the Q-function defined as:

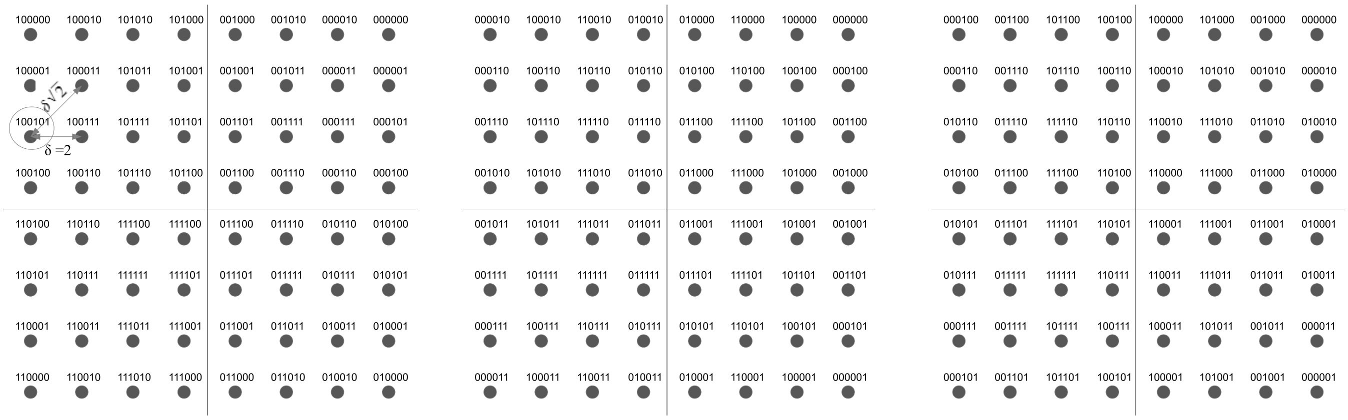
$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^{+\infty} e^{\frac{-t^2}{2}} dt. \tag{7}$$

The probability of error on a received codeword $P_e(\sigma)$ is thus upper bounded (Union Bound inequality) by $U_b(\sigma)$: $P_e(\sigma) \leq U_b(\sigma)$, with

$$U_b(\sigma) = \frac{1}{|\mathcal{C}|} \sum_{\mathbf{x} \in \mathcal{C}} \sum_{\mathbf{y} \in \mathcal{C}/\mathbf{x}} P(\mathbf{x} \rightarrow \mathbf{y}) = \frac{1}{|\mathcal{C}|} \sum_d S_{\mathcal{C}}(d) Q(\frac{d}{2\sigma}), \tag{8}$$

where $|\mathcal{C}|$ represents the cardinality of set $\mathcal{C}$.

For large SNRs, equation (8) can be approximated by only using the first (dominating) terms in the Union Bound which concern the codewords at minimum distances.



(a) Mapping 1: DVB-T2 (b) Mapping 2 (c) Mapping 3

Fig. 1: Gray mappings for coded modulations C_1, C_2 and C_3

IV. DISTANCE SPECTRUM EVALUATION

The exact evaluation of the DS of a code is a computationally intensive task. For a check node of degree d_c, the first $d_c - 1$ inputs can be set arbitrarily to any value of GF(q), while the last one is determined by (1), thus $|\mathcal{C}| = q^{d_c-1}$. The evaluation of all couples has a complexity of $|\mathcal{C}|(|\mathcal{C}|-1)/2$. For instance, for values of $q = 64$ and $d_c = 4$, more than 3.4×10^{10} distances have to be evaluated. Fortunately, the Q function decreases very rapidly and only the first terms in the DS are useful to accurately estimate the Union Bound for high SNRs.

Let us consider the following approach: we define d_u as the maximum value of the Euclidean distance for which the DS is exactly evaluated (in others words, if $d \leq d_u$, then $S_{\mathcal{C}}(d)$ should be exactly evaluated) and δ the minimum Euclidean distance between two points of the constellation $\mathcal{M}$. For each point $\pi(x)$ in $\mathcal{M}$, we define its near neighborhood as the set $V(x)$ expressed by:

$$V(x) = \{y \in \text{GF}(q)/D(x,y)^2 \leq d_u^2 - \delta^2\}. \qquad (9)$$

Since two distinct codewords $\mathbf{x}$ and $\mathbf{y}$ of $\mathcal{C}$ satisfy (1), then $\mathbf{x}$ and $\mathbf{y}$ differ at least by two distinct symbols among the possible d_c symbols. From this property, we deduce that having $D_{\mathcal{C}}(\mathbf{x},\mathbf{y})^2 \leq d_u^2$ implies that $y_k \in V(x_k)$ for $k = 1\dots d_c$ (the proof can be done by contraposition). Thus, for a given codeword $\mathbf{x}$, estimating the codewords $\mathbf{y} \in \mathcal{C}$ such that $D_{\mathcal{C}}(\mathbf{x},\mathbf{y}) \leq d_u$ requires a maximum of v^{d_c-1} distance evaluations, where v is the maximum cardinality of $V(x)$, i.e., $v = \max\{|V(x)|, x \in \text{GF}(q)\}$.

For the example previously introduced with $q = 64$, $d_c = 4$, and a 64-QAM constellation, let us set $d_u = 4$. The minimum distance in the 64-QAM constellation is $\delta = 2$. Therefore, the neighborhood $V(x)$ of x should contain all points of the constellation at a distance smaller or equal to $\sqrt{d_u^2 - \delta^2} = \sqrt{12}$ from x. Since $2\sqrt{2} < \sqrt{12} < 4$, $V(x) = \{y \in \text{GF}(64)\}$ such that $|\pi_I(x) - \pi_I(y)| \leq 2$ and $|\pi_Q(x) - \pi_Q(y)| \leq 2$. In that case, $|V(x)| \leq 9, x \in \text{GF}(64)$ and hence $v = 9$. Enumerating the set of points $\mathbf{y}$ of $\mathcal{C}$ at a distance smaller than or equal to $d_u = 4$ from a given point $\mathbf{x}$ of $\mathcal{C}$ requires a maximum number of distance computations equal to $v^{d_c-1} = 9^3 = 729$. Then, the exact evaluation of the first terms of the DS is bounded by $64^3\times9^3 \cong 1.91\times10^8$, which is computationally more tractable.

V. EUCLIDEAN DISTANCE IN CODED MODULATIONS

The Galois Field GF(q), with $q = 2^r$, can be defined by the set of polynomials over $\mathbb{Z}/2\mathbb{Z}[\alpha] \mod P[\alpha]$, where $P[\alpha]$ is an irreducible polynomial of degree r. In that case, each element of GF(q) can be represented by a binary vector of size r as $x = (x_0, x_1, \dots x_{r-1})_2$, with $x = x_0\alpha^0 + x_1\alpha^1 \dots + x_{r-1}\alpha^{r-1}$. The non-null elements of GF(q) can also be represented as $x = \alpha^\mu$, $\mu = 0\dots q-2$.

In the case of a Binary Phase-Shift Keying (BPSK) modulation, the binary representation $(x_0, x_1 \dots x_{r-1})$ of $x \in \text{GF}(q)$ is used to modulate r BPSK symbol as $s_i = (1 - 2x_i)$, $i = 0\dots r-1$. Then, the Euclidean distance $D(x,y)$ between two symbols of GF$(q)^2$ is exactly twice the Hamming distance $d_H(x,y)$ between the binary representation of x and y. By extension, $D_{\mathcal{C}}(\mathbf{x},\mathbf{y})$ is also twice the Hamming distance $d_H(\mathbf{x},\mathbf{y})$ between the binary representation of the two codewords. This means that the coefficients h_k, $k = 1\dots d_c$ in (1) should be chosen so as to optimise the Hamming distance of the code. This approach was proposed in [9]. For GF(64), $d_c = 4$, and $P[\alpha] = \alpha^6+\alpha+1$, the best coefficients found are $\{h_k\}_{k=1\dots4} = \{\alpha^0, \alpha^9, \alpha^{22}, \alpha^{37}\}$ (the order has no importance). With these optimal coefficients, the minimum Hamming distance between two codewords is 3 and there are exactly 20 codewords at distance 3 of the all-zero codeword (see [9]).

When using a q-ary modulation, there is no longer any direct connection between Hamming distance and Euclidean distance, except if a Gray mapping (for q-QAM modulation) or Gray-like mapping (for q-APSK modulation) is used. Considering, for example, the mapping in Fig. 1a, $x = (x_0, x_1, \dots, x_5)_2$ is associated with $\pi^0(x) = (\pi_I^0(x), \pi_Q^0(x))$, where $\pi_I^0(x) = G(x_0 + 2x_2 + 4x_4)$ and $\pi_Q^0(x) = G(x_1 + 2x_3 + 4x_5)$, with $\{G(i)\}_{i=0\dots7} = \{+7, -7, +1, -1, +5, -5, +3, -3\}$. For example, $x = (100101)_2$ is assigned to $\pi^0(x) = (G(1 + 2 \times$

$0+4\times 0), G(0+2\times 1+4\times 1)) = (G(1), G(6)) = (-7, 3)$. For a 64-QAM Gray-mapped constellation, the following properties are satisfied for all $(x, y) \in \mathrm{GF}(64)^2$:

- Property 1: $D(x, y) = 2 \Rightarrow d_H(x, y) = 1$, for example $x = (100101)_2$ and $y = (100111)_2$ in Mapping 1 (Fig. 1a)
- Property 2: $D(x, y) = 2\sqrt{2} \Rightarrow d_H(x, y) = 2$ (corresponding to two points in opposite positions in a square of side $\delta = 2$). For example, $x = (100101)_2$ and $y = (100011)_2$ in Fig. 1a.

Thus, we can deduce:

- Property 3: $d_H(x, y) = 2 \Rightarrow D(x, y) \geq 2\sqrt{2}$
- Property 4: $d_H(x, y) \geq 3 \Rightarrow D(x, y) \geq 4$.

From these properties, we can infer that, if the Hamming distance between two codewords $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{C}^2$ is greater than or equal to 3, then $D_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) \geq 2\sqrt{3}$.

To summarize, using a 64-QAM Gray-mapped constellation and a parity check equation that guarantee a minimum Hamming distance of three yields a code with a minimum Euclidean distance of $2\sqrt{3}$. In order to reduce the number of couples with an Euclidean distance of $2\sqrt{3}$, we first investigate the impact of Gray mapping. In fact, shuffling the binary representation of x before applying the mapping π^0 leads also to a Gray mapping. More formally, we define the mapping π^0_σ, where σ is a permutation in the set $\{0, 1, \ldots, 5\}$ as $\pi^0_\sigma(x) = \pi^0(\sigma(x))$, with $\sigma : \mathrm{GF}(64) \rightarrow \mathrm{GF}(64)$, $x = (x_0, x_1, \ldots x_5) \rightarrow \sigma(x) = (x_{\sigma(0)}, x_{\sigma(1)}, \ldots x_{\sigma(5)})$. Note that a permutation of the binary representation does not affect the Hamming distance, i.e., $\forall(x, y) \in \mathrm{GF}(64), d_H(x, y) = d_H(\sigma(x), \sigma(y))$ but does affect the Euclidean distance after mapping, and can therefore eventually improve the spectrum of the coded modulation. Mappings in Fig. 1b and 1c can then be obtained from Mapping 1 (Fig. 1a) through this kind of permutation.

VI. Joint optimisation of mapping and NB-LDPC matrix coefficients

In this section, we propose to jointly optimise both mapping and check node coefficients. To this end, we start by performing an exhaustive search among possible mappings for the d_c-uple of coefficients that optimise the DS, or in practice, that minimise the first two terms in $S_{\mathcal{C}}(d)$.

If we consider again the example for the 64-QAM with $q = 64$ and $d_c = 4$, the exhaustive search for each mapping should minimise $S_{\mathcal{C}}(2\sqrt{3})$ and $S_{\mathcal{C}}(2\sqrt{4})$ in the DS, since these two terms are considered as the dominating terms that determine the high SNR regime performance of the coded modulation.

Table I presents three different coded modulations $\mathcal{C}_i, i = 1, 2, 3$, each one defined by a mapping $\pi_i, i = 1, 2, 3$ as described in Figure 1 and a set of $d_c = 4$ coefficients $h_k \in \mathrm{GF}(q)$, $k = 1 \ldots d_c$. These $\mathcal{C}$'s have been chosen as follows: $\mathcal{C}_1$ uses the DVB-T2 Gray mapping [18] and coefficients $(\alpha^0, \alpha^9, \alpha^{22}, \alpha^{37})$ as proposed in [9]; $\mathcal{C}_2$ uses the same coefficients as $\mathcal{C}_1$ and Mapping 2 (Gray mapping that maximizes $S_{\mathcal{C}}(2\sqrt{3})$ in DS). Note that this corresponds to the worst case, or equivalently the mapping that should show the worst performance for the coded modulation, and it is considered for comparison purposes. Finally, $\mathcal{C}_3$ is our proposed combination of mapping and coefficients, i.e. Mapping 3 in Fig. 1 with coefficients $(\alpha^0, \alpha^8, \alpha^{16}, \alpha^{42})$, obtained after an exhaustive search which consists in calculating the two first terms of DS for a large number of possible mapping/coefficients combinations. Note that $\mathcal{C}_3$ significantly reduces $S_{\mathcal{C}}(2\sqrt{3})$. A reduction around 25% compared to the optimised NB-LDPC code in [9] ($\mathcal{C}_1$) and around 58% compared to $\mathcal{C}_2$. This reduction should have a positive impact on the NB-LDPC coded modulation performance as we shall see next.

TABLE I: First terms of DS for coded modulations $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$

Code	Mapping σ	Coeff. of (1)	$S_{\mathcal{C}}(2\sqrt{3})$	$S_{\mathcal{C}}(2\sqrt{4})$
$\mathcal{C}_1$	$\{5,4,3,2,1,0\}$	$\alpha^{\{0,9,22,37\}}$	516,096	3,868,672
$\mathcal{C}_2$	$\{3,0,2,1,5,4\}$	$\alpha^{\{0,9,22,37\}}$	909,312	2,910,208
$\mathcal{C}_3$	$\{4,2,1,0,5,3\}$	$\alpha^{\{0,8,16,42\}}$	385,024	3,499,008

VII. Simulation results

We first consider the performance of a single parity check code of size $d_c = 4$ in GF(64) associated to a 64-QAM modulation for coded modulations $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$. Fig. 2 presents curves that correspond to the Union Bound calculation with $S_{\mathcal{C}}(2\sqrt{3})$ and $S_{\mathcal{C}}(2\sqrt{4})$, i.e. the first two terms in DS, as well as the Maximum Likelihood (ML) decoding performance curves expressed in Frame Error Rate (FER), i.e. Monte-Carlo simulations with a stopping criterion of 100 errors. From this figure we can first observe that the Union Bound is an accurate approximation for SNR values greater than 16 dB and becomes an exact bound in the high SNR regime region starting from about 20 dB.

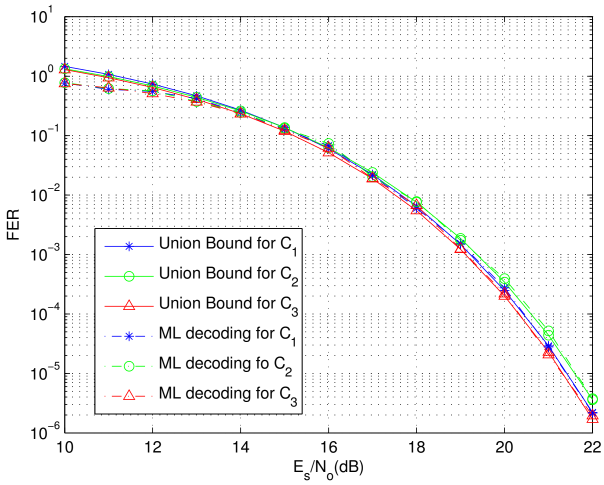


Fig. 2: Union Bound and FER performance for the Single Parity Check (SPC) coded modulations $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$.

We now consider a regular GF(64)-LDPC code of length $N = 48$ symbols, with $d_v = 2$, $d_c = 4$ (coding rate 1/2). The positions of the non-zero matrix coefficients are the one proposed in [9], the mapping and the coefficient values are those in Table I which are randomly assigned to the non-zero positions at each row in the matrix, i.e. for a single check node. We consider the L-Bubble EMS decoding algorithm [19] [20] with a number of significant values $n_m = 25$ and 20 decoding iterations [1]. The demapping step follows the principle described in [22] for simplified intrinsic Log-Likelihood Ratio generation. Fig. 2 shows that $\mathcal{C}_3$ outperforms $\mathcal{C}_1$ and $\mathcal{C}_2$, specially in the high SNR regime region. To be specific, a gain of 0.2 dB (0.15 dB) at a FER = 2×10^{-8} with respect to $\mathcal{C}_2$ ($\mathcal{C}_3$) is achieved with the proposed solution. Note that this performance gain does not entail any additional complexity at the transmitter nor at the receiver compared to existing schemes, as the enhancement comes from the matrix construction and the Gray mapping choice. Also note that even if the proposed approach is based on an exhaustive search to optimise the DS properties of the coded modulation, this step is performed only once during the code design.

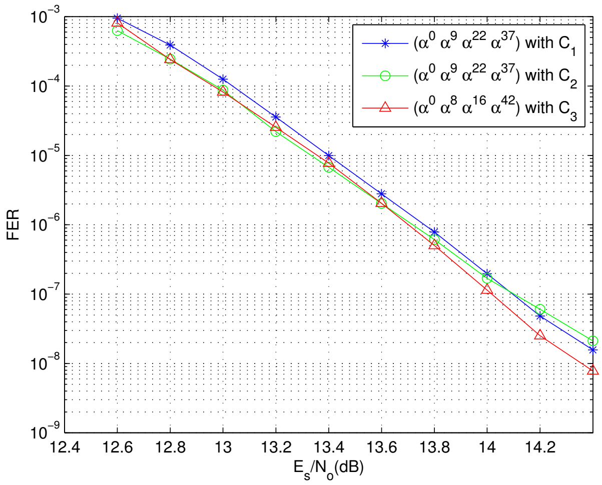


Fig. 3: Decoding performance of a $N = 48$ GF(64)-LDPC code with coded modulations $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$.

VIII. Conclusion

In this paper we have considered the design of advanced high-spectral efficiency communications with error-decoding performance. We focused on high-order NB-LDPC coded modulations where the order of the Galois Field and modulation order coincide. We based the NB-LDPC matrix optimisation on the analysis of a single check node to find the best GF(q) values for the d_c coefficients for a given modulation mapping. To show the good agreement between the theoretical analysis and the simulation results, we calculated and compared the Union Bound with the ML decoding curves. Finally, we presented simulation results to show how the NB-LDPC coded modulation designed with the proposed method outperforms the state of the art.

[1]The coding matrices and the simulation software are available in [21]

References

[1] M. G. Luby, M. Mitzenmacher, and M. A. Shokrollahi, "Improved low density parity check codes using irregular graphs," *Information Theory, IEEE Transactions on*, vol. 2, pp. 585–598, Feb 2001.

[2] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity approching irregular low-density parity-check codes," *Information Theory, IEEE Transactions on*, vol. 47, pp. 657–670, Feb 2001.

[3] M. C. Davey and D. MacKay, "Low-density parity-check codes over GF(q)," *IEEE Communication Letters*, vol. 2, pp. 165–167, June 1998.

[4] F. Guo and L. Hanzo, "Low-complexity non-binary LDPC and modulation schemes communicatins over MIMO channels," in *IEEE Vehicular Technology Conference (VTC'2004)*. Los Angeles, USA, Sept. 2004.

[5] X. Jiand, Y. Yan, X. Xia, and M. Lee, "Application of non-binary LDPC codes based on euclidean geometries to MIMO systems," in *Int. Conference on wireless communications and signal processing, WCSP'09*. Nanjing, China, Nov. 2009, pp. 1–5.

[6] A. Haroun, C. Abdel Nour, M. Arzel, and C. Jego, "Low-complexity LDPC-coded iterative MIMO receiver based on belief propagation algorithm for detection," in *Turbo Codes and Iterative Information Processing (ISTC), 2014 8th Int. Symp. on*, Aug. 2014, pp. 213–217.

[7] D. Declercq, M. Colas, and G. Gelle, "Regular GF(2^q)-LDPC coded modulations for higher order QAM-AWGN channel," in *Proc. ISITA*. Parma, Italy, Oct. 2004.

[8] D. Declercq and M. Fossorier, "Decoding algorithms for non-binary LDPC codes over GF(q)," *IEEE Trans. on Commun.*, vol. 55, pp. 633–643, April 2007.

[9] C. Poulliat, M. Fossorier, and D. Declercq, "Design of regular $(2, d_c)$-LDPC codes over GF(q) using their binary images," *IEEE Trans. on Communications*, vol. 56, pp. 1626–1635, Oct 2008.

[10] R. J. M. Hui Jin, "General coding theorems for turbo-like codes," *Proceedings IEEE International Symposium on Information Theory*, vol. 10, p. 120, 2000.

[11] X. C. K. Chugg, A. Anastasopoulos, "Adaptivity, complexity reduction, and applications," *Norwell, MA: Kluwer*.

[12] X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *Information Theory, IEEE Transactions on*, vol. 51, pp. 386–398, Jan 2005.

[13] D. MacKay. (2003, Aug) Optimizing sparse graph codes over GF(q). [Online]. Available: {online}

[14] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets part I: Introduction," *IEEE Communications Magazine*, vol. 25, no. 2, pp. 5–11, February 1987.

[15] A. Martinez, A. G. i Fabregas, G. Caire, and F. M. J. Willems, "Bit-interleaved coded modulation revisited: A mismatched decoding perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2756–2765, June 2009.

[16] Y. Kofman, E. Zehavi, and S. Shamai, "Performance analysis of a multi-level coded modulation system," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 299–312, Feb 1994.

[17] S. Nowak and R. Kays, "Interleaver design for spectrally-efficient bit-interleaved LDPC-coded modulation," in *Turbo Codes and Iterative Information Processing (ISTC), 2012 7th International Symposium on*, Aug 2012, pp. 240–244.

[18] "Digital video broadcasting (DVB); frame structure, channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)," in *ETSI EN 302 755*, 2008.

[19] E. Boutillon, L. Conde-Canencia, and A. Al Ghouwayel, "Design of a GF(64)-LDPC decoder based on the EMS algorithm," *Circuits and Systems I: Regular Papers, IEEE Trans. on*, vol. 60, no. 10, pp. 2644–2656, 2013.

[20] E. Boutillon and L. Conde-Canencia, "Simplified check node processing in nonbinary ldpc decoders," in *2010 6th International Symposium on Turbo Codes Iterative Information Processing*, Sept 2010, pp. 201–205.

[21] C. Marchand and E. boutillon. (2015) Non-binary LDPC codes website. [Online]. Available: http://www-labsticc.univ-ubs.fr/nb_ldpc/

[22] L. Conde-Canencia and E. Boutillon, "Application of bubble-check algorithm to non-binary LLR computation in QAM coded schemes," *Electronics Letters*, vol. 50, no. 25, pp. 1937–1938, 2014.