Getting Closer to MIMO Capacity with Non-Binary Codes and Spatial Multiplexing

Stephan Pfletschinger*, D. Declercq†

* Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Av. del Carl Friedrich Gauss 7, 08860 Castelldefels (Barcelona), Spain
† ETIS, ENSEA, University of Cergy-Pontoise, CNRS
F-95000, Cergy-Pontoise, France

Abstract—In this paper, we discuss the combination of non-binary channel coding with higher-order modulation and MIMO transmission in the form of spatial multiplexing. In addition to the benefits of non-binary LDPC codes on the AWGN channel, we identify an inherent advantage for non-binary coding in multiple antenna schemes. By comparing binary and non-binary information processing at the receiver, we highlight the intrinsic advantages of non-binary coding. A performance comparison based on simulation results shows that, without increasing system complexity, non-binary information processing can actually outperform the Shannon limit of its binary counterpart.

Index Terms—Non-binary LDPC codes, MIMO, spatial multiplexing, soft demapping

I. INTRODUCTION

Virtually all modern wireless systems include channel coding and multiple antennas, often combined with multi-carrier modulation and adaptive coding and modulation. It is therefore of special interest to combine these approaches in an efficient way, both with respect to performance and system complexity. In this paper, we focus on the combination of non-binary channel coding and transmission over multiple antennas in the form of spatial multiplexing.

It has been observed that non-binary LDPC codes offer benefits over their binary counterparts, even for channels with binary inputs, as e.g. the binary symmetric channel or the AWGN channel with BPSK modulation [1], [2]. For higher-order modulations, the application of $q$-ary LDPC codes with $q = 2^n > 2$ has been proved very efficient, both with analytical approaches and in simulations [3], [4], [5]. For multiple-antenna schemes, non-binary coding has further advantages (see also [6]), which include the inherently higher capacity of the equivalent channel seen by the coding scheme. In this paper, we evaluate these benefits by considering the combination of non-binary channel coding with spatial multiplexing and higher-order modulation.

II. BINARY VS. NON-BINARY CODES FOR THE AWGN CHANNEL

A. Non-binary LDPC Codes

Binary LDPC codes can be generalized to non-binary LDPC codes (NB-LDPC). The parity-check equations are written using symbols in a Galois field of order $q$, denoted $\mathbb{F}_q$, where $q = 2$ is the particular binary case. The parity check matrix $A$ defining the code has only a few nonzero values which belong to $\mathbb{F}_q$, and a single parity equation involving $d_v$ codeword symbols has then the form: $\sum_{i=1}^{d_v} a_{ji} c_i = 0$, where $a_{ji}$ are the nonzero values of the $j$-th row of $A$.

Non-binary LDPC codes offer some advantages compared to their binary counterparts. As pointed out by several authors [1], [7], [8], [2], the Tanner graph of the NB-LDPC code is usually much sparser than the one of a binary code with the same parameters. As a consequence, the higher girth of NB-LDPC graphs helps to avoid the short cycles and also mitigates the effect of stopping or trapping sets, making the non-binary message passing decoders closer to maximum likelihood decoding. Actually, when $q \geq 2^6$, best error rate results on additive memoryless channels are obtained with the lowest possible variable node degree, that is $d_v = 2$. The codes with $d_v = 2$ have been named cycle-codes in literature, or ultra-sparse LDPC codes [1], [7]. For example, the girth of a binary irregular LDPC code with length $N = 848$ bits and rate $R_c = 0.5$ is at most $g_{th} = 6$ for the good degree distributions, while the girth of a NB-LDPC code with same parameters is $g_{th} = 14$ when a good graph construction is used [8], [9].

This makes the $q$-ary LDPC codes appealing especially for short to medium block-lengths, as well as for higher-order coded modulation, as explained in the next two sections.

B. Mapping and Demapping

For transmission over the physical channel, the code symbols defined in $\mathbb{F}_q$ have to be mapped to $M$-QAM symbols, where $M$ is a power of two, including $M \in \{2, 4\}$ for BPSK and QPSK. The coded modulation system is depicted in Figure 1: the message $\mathbf{u} \in \mathbb{F}_q^N$ is encoded into a codeword $\mathbf{c} \in \mathbb{F}_q^N$, which is passed via a symbol-wise interleaver to the modulator and then transmitted over a continuous-valued channel. At the receiver, the soft demapper computes APP (a posteriori probability) log-likelihood values (L-values), which constitute a sufficient statistic of the receive signal $\mathbf{y}$ and form the input of the channel decoder.

![Block diagram for a coded transmission system](image-url)
To obtain a bijective mapping, we have to map $m_1$ code symbols to $m_2$ QAM symbols such that $q^{m_1} = M^{m_2}$. We denote the QAM alphabet by $\mathcal{X}_M$, and the mapping function by $\mu$, i.e.

$$\mu : \mathbb{F}_q^{m_1} \rightarrow \mathcal{X}_M^{m_2} \tag{1}$$

The code symbols $\mathbf{b} = (b_1, b_2, \ldots, b_{m_1})$ belong to the same codeword $\mathbf{c} = (c_1, c_2, \ldots, c_N)$ and are mapped to a vector of QAM symbols,

$$\mathbf{x} = (x_1, x_2, \ldots, x_{m_2})^T = \mu(\mathbf{b}) \in \mathcal{X}_M^{m_2} \tag{2}$$

For binary codes (i.e. for $q = 2$), we always have $m_2 = 1$, while for codes in higher order Galois fields, for many modulations $m_1 = 1$, which is quite beneficial for the demapping, as we will see below.

The soft demapper computes the APP L-vector $L_i \triangleq (L_{i,0}, L_{i,1}, \ldots, L_{i,q-1})$, which corresponds to the code symbol $b_i$, and whose components are given by

$$L_{i,k} \triangleq \ln \frac{P[b_i = k | y]}{P[b_i = 0 | y]} \quad \text{for} \quad i = 1, \ldots, m_1 \quad k \in \mathbb{F}_q$$

For a memoryless scalar channel given by $y_j = h_j x_j + w_j$ with $x_j = \mu_j(\mathbf{b})$, $j = 1, \ldots, m_2$, $w_j \sim \mathcal{N}(0, N_0)$ and assuming that all code symbols are equiprobable, we obtain

$$L_{i,k} = \ln \left( \sum_{b \in B_i^k} P(y | b) \right) / \left( \sum_{b \in \mathbb{F}_q^k} P(y | b) \right) = \text{jaclog} \left\{ P(y | b) \right\} - \text{jaclog} \left\{ P(y | b) \right\}$$

where $p(y | b) = \frac{1}{N_0} \sum_{j=1}^{m_2} | y_j - h_j \mu_j(b) |^2$ and $B_i^k \triangleq \{ b \in \mathbb{F}_q^{m_2} : b_i = k \}$ is the set of all code symbol vectors whose $i$-th component is fixed to $k$, and $\text{jaclog} \in \mathcal{L}(x) \triangleq \ln \sum_{x \in \mathcal{L}} \exp(x)$ is the Jacobian algorithm [10], [11]. For the max-log approximation, the jaclog $\cdot$ is simply replaced by $\max \cdot$.

**C. Binary Demapping**

For binary codes, the L-vector is replaced by a scalar L-value, which is defined by (3) for $k = 1$.

In the following, we will compare the possibilities and performance of employing binary and $q$-ary channel codes in transmission systems with higher-order modulation and multiple antennas. In order to evaluate the binary and $q$-ary information processing, rather than comparing some particular channel codes of each kind, we actually will use the same $q$-ary LDPC code for both cases. This is possible since we only consider GF orders which are a power of two, i.e. $q = 2^p$ with $p \in \mathbb{N}$. In this way, all obtained performance differences are entirely due to the binary or $q$-ary demapping since for memoryless binary-input channels, the two approaches are identical.

For the $q$-ary decoder, we have to reconstruct the L-vectors from the binary L-values. For $k \in \mathbb{F}_q$ we denote the corresponding bit vector by $(k_1, k_2, \ldots, k_p) \in \mathbb{F}_m^p$. Under the assumption that the binary L-values are independent, the $q$-ary L-vector is given by

$$L_{i,k} = \ln \prod_{j=1}^{ldq} \frac{P[b_{i,j} = k_j | y]}{P[b_{i,j} = 0 | y]} = \sum_{j=1}^{ldq} k_j \lambda_i$$

For modulations higher than QPSK, the binary L-values are actually not independent and therefore we cannot perfectly reconstruct the $q$-ary L-vector with (5).

**D. Performance of Selected Non-binary LDPC Codes**

For all simulations in this paper, we selected two regular non-binary LDPC codes with the following parameters

<table>
<thead>
<tr>
<th>Code</th>
<th>$q$</th>
<th>$N$</th>
<th>$N_{\text{bin}}$</th>
<th>$N_{\text{ldq}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code 1</td>
<td>64</td>
<td>2688</td>
<td>16 128 bits</td>
<td>16 128 bits</td>
</tr>
<tr>
<td>Code 2</td>
<td>256</td>
<td>360</td>
<td>2880 bits</td>
<td>2880 bits</td>
</tr>
</tbody>
</table>

Both codes have rate $R_c = 1/2$ and have variable and check node degrees $d_v = 2$ and $d_c = 4$, respectively. The performance of these codes in terms of word error ratio (WER) over the AWGN channel with BPSK modulation is depicted in Figure 2, and is determined by their length and not by their GF order. For this channel, there is no difference between binary and $q$-ary demapping.

![Figure 2](image)

From information theory, an advantage for $q$-ary information processing is to be expected [12], [13], [14]: the capacity of coded modulation (CM) is equal or greater to the capacity of bit-interleaved coded modulation (BICM), which neglects the dependencies between the received bit layers, and the difference is increasing for higher-order modulations and multiple antennas. For all binary schemes considered in this paper, we assume that the interleaver in Figure 1 is bitwise and the binary L-values are processed independently. In particular, we do not consider multilevel coding, which would require information passing between binary decoders, known as multistage decoding [15].
III. TRANSMISSION WITH MULTIPLE ANTENNAS: SPATIAL MULTIPLEXING

A. System Model

We consider spatial multiplexing with \( n_T \) transmit and \( n_R \) receive antennas. The received signal is then

\[
y = Hx + w, \quad w \sim CN(0, N_0 I_n)
\]  

(6)

The channel matrix \( H \in \mathbb{C}^{n_R \times n_T} \) is assumed to be i.i.d. Rayleigh fading, i.e. the components are complex Gaussian distributed with zero mean and unit variance, \( H_{i,j} \sim CN(0,1) \). The transmit power per antenna is \( E_S = \mathbb{E}[|x|^2] \) and hence the SNR is given by \( SNR = \frac{E_S}{N_0} \).

Although many other MIMO schemes have been proposed, spatial multiplexing is for many situations the preferable approach since it does not require channel state information (CSI) at the transmitter and, especially when combined with strong channel coding, there is little need for gathering additional diversity. For broadband systems with OFDM and channel coding, diversity is already obtained in the form of frequency and time diversity and hence for the MIMO scheme, the multiplexing gain is more important than possible diversity gains.

B. Mappings for Multiple Antennas

For spatial multiplexing with \( n_T \) transmit antennas, there are several possibilities to map the \( m_2 \) QAM symbols to the antennas. For the case that these symbols are allocated to only one transmit vector, e.g. for \( m_2 > n_T \), we have to adopt a matrix notation, \( Y = HX + W \), where \( X \in \mathbb{C}^{m_2 \times T} \) contains the \( m_2 \) QAM symbols which correspond to the code symbol vector \( b \). The L-values for the general case are given by

\[
L_{i,k} = \ln \frac{\sum_{(X: b_i = k)} p(Y|X)}{\sum_{(X: b_i = 0)} p(Y|X)}
\]  

(7)

In the following, we consider some relevant special cases for the allocation of the QAM symbols to the transmit antennas. For higher-order Galois fields, e.g. GF(64) or GF(256), the complexity of the soft demapper becomes impractical if more than two code symbols are involved in the mapping. For this reason, in the following, we focus on cases with \( m_1 = 1 \), i.e. the QAM symbols \( (x_1, \ldots, x_{m_2}) \) depend on only one code symbol.

1) Horizontal mapping: The QAM symbols \( (x_1, \ldots, x_{m_2}) = \mu(b) \) are placed in the rows of \( X \), hence \( T = m_2 \) and the cardinality of the set \( \{X: b_i = k\} \) over which the sum must be taken is

\[
N_x = |\{X: b = k\}| = M^{m_2(n_T-1)} = q^{m_1(n_T-1)}
\]  

(8)

since each code symbol fixes one row of \( X \). The soft demapper for this case is given by (7), and due to the huge set \( \{X: b_i = k\} \), this is only practical for few antennas and small QAM constellations.

2) Vertical mapping: A great complexity reduction can be achieved if the QAM symbols \( (x_1, \ldots, x_{m_2}) \) directly form the transmit vector, i.e. \( n_T = m_2 \) and \( T = 1 \), in addition to \( m_1 = 1 \). In this case, each transmit vector corresponds to exactly one code symbol \( b \in \mathbb{F}_q \), i.e. \( x = (x_1, \ldots, x_{m_2}) = (\mu_1(b), \ldots, \mu_{m_2}(b)) \) and the set \( \{X: b_i = k\} \) reduces to one element. Therefore, (7) simplifies to

\[
L_k = -\frac{1}{N_0} \left( ||y - H\mu(k)||^2 - ||y - H\mu(0)||^2 \right)
\]

\[= \frac{2}{N_0} \left( \sum_{y_H \mu(k)} y_{HG} H \mu(k) - \mu(k) \mu_{HG} H \mu(k) \right) + c_0
\]  

(9)

where \( c_0 \) is a constant that does not depend on \( k \) and can be omitted if the decoder is invariant to additive constants.

3) Binary demapping: As in the SISO case, we can view a code symbol \( b \in \mathbb{F}_q \) as a vector of \( ldq \) bits and derive their binary L-values. The \( q \)-ary L-vectors are then obtained by (5). While for a single antenna, this is equivalent to q-ary demapping for BPSK and QPSK (with Gray labeling), for multiple antennas, this is sub-optimum for all modulations.

C. Channel Capacities

The channel capacity for the MIMO system defined above in Subsection III-A is given by [16]

\[
C_G = \mathbb{E}_H \left[ \log_2 \left( 1 + \frac{SNR}{n_T} HH^H \right) \right]
\]  

(10)

This expression holds for Gaussian distributed transmit signals. However, in our system the transmit symbols are taken out of a QAM constellation \( \mathcal{X}_{qT} \). We denote by \( R_q = \log_2 M \) the number of bits per QAM symbol (the “modulation rate”), by \( R_0 = n_T R_q \) the number of transmitted bits per channel use and define the set \( \mathcal{X} \), of cardinality \( |\mathcal{X}| = 2^{R_0} \), of all possible transmit vectors.

The coded modulation (CM) capacity [13], [14] is then given by

\[
C_{CM} = R_0 - \mathbb{E}_{x,y,H} \left[ \log_2 \left( \sum_{z \in \mathcal{X}} \frac{p(y|z,H)}{p(y|x,H)} \right) \right]
\]  

(11)

For binary processing, the BICM capacity applies [13], [14]: We first define the subsets \( \mathcal{X}'_j \), which contain the transmit vectors whose associated bit vector has the value \( j \in \{0,1\} \) in its \( r \)-th position. The BICM capacity is given by

\[
C_{BICM} = R_0 - \mathbb{E}_{x,y,H} \left[ \log_2 \left( \sum_{z \in \mathcal{X}'} \frac{p(y|z,H)}{\sum_{z \in \mathcal{X}'} p(y|z,H)} \right) \right]
\]  

(12)

Figures 3 and 4 show these capacities for a \( 2 \times 2 \) and \( 4 \times 4 \) MIMO channel. The considered modulation orders are such that in both cases \( R_0 = n_T R_q = \log_2 q \in \{6,8\} \), which allows to apply the LDPC codes specified in Section II-D, defined over GF(64) and GF(256).

Figure 3 depicts several capacity curves for CM and BICM and different modulations. While for 16-QAM, the usual Gray
labeling has been applied (note that only the BICM capacity but not the CM capacity depends on the labeling), for 8-ary modulation there are several possibilities. For 8-QAM, we selected the constellation with minimum energy for a given minimum distance [17, p. 279]. It is interesting to note that this constellation is not optimum in terms of the BICM capacity, which is the reason why for the binary approach we selected the 8-PSK constellation. On the other hand, in terms of the CM capacity, the differences among different constellations are minor (not shown in the figure). For the following simulations, we apply 8-PSK for the binary demapping and 8-QAM for the $q$-ary case.

We can observe that at a capacity of 3 bits per channel use, corresponding to the application of a half-rate channel code, the capacity gap between 8-QAM CM and 8-PSK BICM is about 1.3 dB, and around 2 dB between BICM and CM for 8-QAM. This implies a significant drawback for binary information processing from the theoretical point of view and will also be apparent with practical channel coding.

For 4 × 4 MIMO with QPSK, the capacity gap is around 1 dB at $R_c = 1/2$, which is less but nevertheless noteworthy since for the single antenna case, the two capacities are identical with QPSK.

The next section will evaluate if this gain in channel capacity translates to a similar advantage for a practical non-binary LDPC code.

### IV. Simulation Results

In this section, we evaluate the potential of non-binary channel coding for multiple antenna transmission for some selected cases. Since each non-binary code over GF($q$) with $q = 2^p$ can be viewed as a binary code, we can apply the same code for both cases. In particular, for all simulations, we used the two ultra-sparse LDPC codes specified in Section II-D with a BP decoder and a maximum of 100 iterations. Note that at the transmitter, apart from the bit or symbol-wise interleaver, there is no distinctive difference between the two approaches.

The simulated BER for different modulations in conjunction with spatial multiplexing the $q$-ary LDPC codes is shown in figures 5, 6 and 7. In all cases, the number of transmit antennas, the modulation order and the GF order are such that one code symbol corresponds to one channel use, i.e. $n_T R_{q} = \log_2 q$ and $m_1 = 1$. The Shannon limits for non-zero error probabilities are obtained by $R_{0} = C_{\text{BICM}}(\text{SNR})$, where $H_2(\cdot)$ is the binary entropy function.

In all cases, the expected performance differences indicated by the respective channel capacities translate to similar gaps in BER. Especially for the longer code, the performance gain is extraordinary: the BER curves for the vertical mapping with $q$-ary processing beat the Shannon limit for binary processing! In other words, no binary scheme can approach the performance of the non-binary vertical mapping scheme, as can be seen in Figs. 5, 6.

For the shorter GF(256) code, the gains also correspond
to the differences between CM and BICM capacity and the performance comes close to the Shannon limit of their binary counterparts. It is expected that a longer code in GF(256) will also yield results beyond the binary Shannon limit.

The complexity of the demapper is determined mainly by the number of elements of the sum in (7). For the horizontal mapping, the cardinality of this set is given by (8) and takes on impractical values for all cases except for \( n_T = 2 \) and \( M = 8 \). In contrary, for the vertical mapping, this set has only one element. For the binary demapper, the complexity is proportional to \( \rho 2^p = gldq \), which is the same factor by which the decoding complexity of the \( q \)-ary channel decoder increases with respect to a binary decoder [18]. In other words, with \( q \)-ary vertical mapping, complexity is shifted from the demapper to the decoder, while the performance increases significantly. Note that if we applied a binary channel code for the binary information processing, the gap would be even more pronounced.

\[
3 \times 3 \text{ MIMO, QPSK, } q = 64
\]

Figure 6. BER for 3 × 3 spatial multiplexing and QPSK for binary and vertical mapping with the GF(64) code

\[
4 \times 4 \text{ MIMO, QPSK, } q = 256
\]

Figure 7. BER for 4 × 4 spatial multiplexing with QPSK, binary and vertical mapping and the GF(256) code

V. CONCLUSIONS

We have analyzed the combination of non-binary channel coding with spatial multiplexing and have pointed out the superior performance of \( q \)-ary receiver processing. The advantage of \( q \)-ary information processing compared to its binary equivalent stems from the greater channel capacity of non-binary schemes, while the complexity of the soft demapper for non-binary codes is reduced for adequately selected parameters. We have shown some examples where an LDPC code defined in GF(64) achieves a performance beyond the Shannon limit for binary codes.

ACKNOWLEDGMENT

This work is supported by INFSCO-ICT-216203 DAVINCI “Design And Versatile Implementation of Non-binary wireless Communications based on Innovative LDPC Codes” (www.ict-davinci-code.eu) funded by the European Commission under the Seventh Framework Programme (FP7).

REFERENCES