Multilevel Coding for Non-Orthogonal Broadcast

Stephan Pfletschinger*, Monica Navarro*, Christian Ibars†
*Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Av. Carl Friedrich Gauss 7, 08860 Castelldefels, Spain
†Intel Corporation, 2200 Mission College Blvd, Santa Clara, CA, USA
{stephan.pfletschinger, monica.navarro}@cttc.es, christian.ibars.casas@intel.com

***Abstract*—This paper defines an information-theoretical framework for non-orthogonal broadcast systems with multilevel coding and gives design guidelines for the rate selection of multiple broadcast streams. This description includes hierarchical modulation and superposition coding with codes defined in a finite field as a special case. We show how multilevel coding can be applied to multiple antennas where, in contrast to most space-time coding and hierarchical modulation schemes, no capacity loss occurs.**

I. INTRODUCTION

This papers addresses broadcasting of multiple data streams to users with possibly, but not necessarily, different SNRs or receiver capabilities. A typical application includes broadcasting of data with different priorities or coverage areas. For this setting, it is usually required that the high-priority stream can be decoded at low SNR while for the low-priority stream better channel conditions are required. Non-orthogonal broadcast has recently received considerable attention as a possible enhancement for LTE and 5G mobile communication systems, as e.g. in [1], [2] which propose it for the downlink of LTE. This approach, referred to as *non-orthogonal multiple access (NOMA)*, shows significant gains when the scheduler is optimized to pair, in the same resource blocks, users with significant difference in received signal strength.

For single-antenna transmission, the optimum solution for broadcasting possibly different data streams to multiple users is given by superposition coding with random Gaussian codebooks [3], while for all practical purposes channel coding defined in a finite field, followed by a suitable modulation, is applied. This scheme is depicted in Fig. 1 along with the corresponding receiver structure which employs successive interference cancellation (SIC). This model can be seen as a direct realization of the optimum solution from information theory, and it corresponds directly to *hierarchical modulation (HM)*, which assembles a QAM constellation by two or more subconstellations corresponding to different data streams [4], [5].

The aspect of multi-stream codes for MIMO systems has been addressed in [6], [7]. In these papers, an optimization of space-time block codes for multi-stream transmission is proposed, which consists of a linear transformation of the encoder matrices. It is also shown that a simpler approach consisting of direct sum of the space time codewords is suboptimal. In this paper, we first provide an information-theoretic framework for MIMO multi-level coding (MLC), based on a binary component channel decomposition. This concept, developed in [8], is also applicable to MIMO systems. We then introduce the concept of multi-stream, space-time codes with non-binary constituent codes. The advantage of using non-binary codes is that multiple binary substreams can be grouped by grouping bits into higher order symbols, reducing the number of different code rates required, and reducing the possibility of error propagation.

This paper can be seen as an extension of [5]–[7] and as a building block for [1] since it brings together HM, multi-stream space-time codes and NOMA in the framework of MLC and provides guidelines for the selection of the substream rates. We also show two practically feasible implementations of MLC which make use of non-binary (de)coding. The paper is organized as follows. In Section II we provide the system model for MIMO multi-level coding. In Section III we provide a model for the capacity of MLC and compare it to BICM. Section IV introduces our design for MLC with non-binary codes, which are evaluated in Section V. Finally, we present the paper conclusions in Section VI.

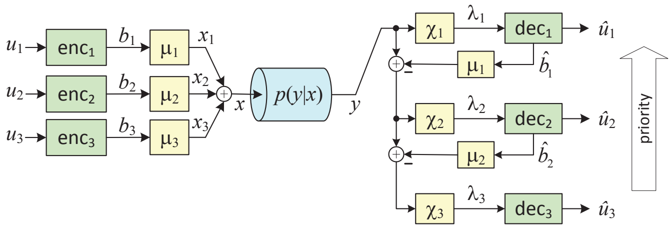

Figure 1. Superposition coding by adding modulated signals. The mapping functions μ_d map an integer number of coded bits b_d to a QAM symbol x_d while the soft demappers χ_d compute the posterior probabilities $\lambda_d \triangleq P[b_d = 1 \mid y]$ for each coded bit.

II. SYSTEM MODEL

A straightforward but decisive generalization of the transmitter structure of Fig. 1 is given in Fig. 2: instead of adding QAM symbols we combine the coded symbols of all data streams and map them to a common QAM symbol.

A simple and common example for this generalization is e.g. 8-PSK which cannot be composed as the sum of three subconstellations.

The model of Fig. 2 corresponds to MLC, which has been treated in detail by Wachsmann et al. in [8]. Within the framework of MLC, we can determine the capacities of the binary subchannels and derive design rules for the selection of the code rate of each stream. The key advantage of this approach is that the subchannel capacities always sum up to the capacity of the constellation-constrained capacity of the channel whereas with HM, as also observed in [4], the sum capacity suffers a loss.

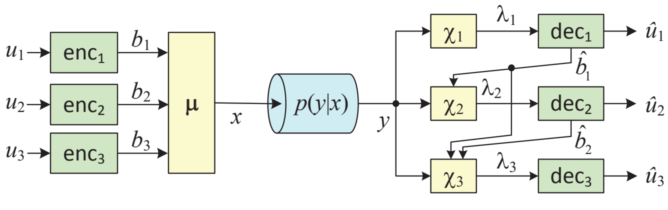

Figure 2. Multilevel coding with multistage decoding

III. Capacities of Binary Subchannels

A. Coded Modulation Capacity

We model the channel with input $\mathbf{x}$ and output $\mathbf{y}$ by its conditional probability density function (pdf) $p(\mathbf{y} \mid \mathbf{x})$, which includes single-antenna as well as MIMO channels. For all realizable transmission systems, the transmit signal is taken out of a discrete constellation $\mathbf{x} \in \mathcal{A} \triangleq \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_M\}$, and hence the capacity of this channel is given by the *coded modulation (CM)* or *constellation-constrained* capacity

$$C_{\mathsf{cm}} \triangleq I(\mathbf{x};\mathbf{y}) = \frac{1}{M}\sum_{m=1}^{M}\int p(\mathbf{y} \mid \mathbf{a}_m)\log_2\frac{p(\mathbf{y} \mid \mathbf{a}_m)}{p(\mathbf{y})}\mathrm{d}\mathbf{y} \tag{1}$$

for which we additionally assume that all symbols $\mathbf{x} \in \mathcal{A}$ are transmitted with the same probability.

B. Decomposition into Binary Subchannels

The transmit symbols $\mathbf{x} \in \mathcal{A}$ correspond to $D = \log_2 M$ bits by the bijective mapping $\mu : \mathbb{F}_2^D \to \mathcal{A}$, i.e. the transmitted symbols are given by $\mathbf{x} = \mu(b_1, \ldots, b_D)$ where $\mathbf{b} = [b_1, b_2, \ldots, b_D]$ denotes a D-dimensional bit vector. We denote by $b_1, \ldots, b_D$ the bits at the input of the mapping device which are drawn from the coded bits at the output of the decoder. With this bijective mapping between $\mathbf{b}$ and $\mathbf{x}$ and the chain rule of mutual information, we can decompose the CM capacity as follows

$$C_{\mathsf{cm}} = I(\mathbf{b};\mathbf{y}) = \sum_{d=1}^{D} C_{\mathsf{mlc}}(d) \tag{2}$$

where

$$C_{\mathsf{mlc}}(d) \triangleq I(b_d;\mathbf{y} \mid b_1, \ldots, b_{d-1}) \tag{3}$$

The MLC subchannel capacities $C_{\mathsf{mlc}}(d)$ correspond to the channels from b_d to λ_d in Fig. 2, assuming that subchannels $1, \ldots, d-1$ have been correctly decoded [8]. This decomposition hence establishes a decoding order which determines the possible choices of code rates per subchannel.

On the other hand, if we do not use information from other subchannels and decode every binary subchannel independently, we are limited by the BICM (bit-interleaved coded modulation) [9] subchannel capacities

$$C_{\mathsf{bicm}}(d) \triangleq I(b_d;\mathbf{y}) \le C_{\mathsf{mlc}}(d). \tag{4}$$

C. Basic Example: 2×2 MIMO with 8-PSK per Antenna

In the following, unless otherwise noted, we consider as main example the 2×2 MIMO channel with fast i.i.d. Rayleigh fading and 8-PSK modulation per antenna. This setting is also chosen for comparability with recent related work [6], [7], and the received signal is hence given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \ \mathbf{w} \sim \mathcal{CN}(0, N_0\mathbf{I}_2) \tag{5}$$

with $H_{i,j} \sim \mathcal{CN}(0,1)$. With this model, we have $D = 6$ binary subchannels, whose capacities according to (2) and (4) for 8-PSK per antenna are plotted in Figure 3, where we can observe that, except for $d = 1$, the subchannel capacities with MLC are indeed significantly higher than with BICM.

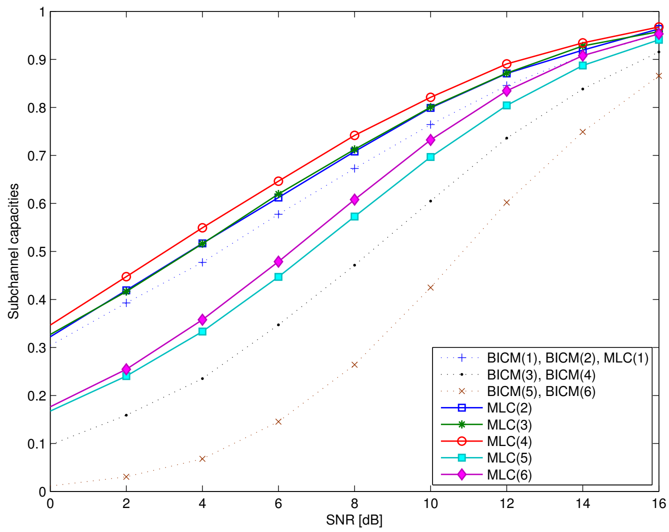

Figure 3. Capacities of the binary subchannels with BICM and with MLC for a 2×2 fast Rayleigh fading MIMO channel with 8-PSK per antenna.

From Fig. 3, we can observe that the binary subchannels have different capacities, although the 8-PSK constellation is highly symmetric. The main advantage with MLC compared to HM is that the subchannels capacities of the former always sum up to the CM capacity. On the other hand, with the usual approaches for HM, the CM capacity is typically reduced.

IV. Subchannel Grouping

A. Transmitter and Receiver Structures

Assume that for the given setting we wish to transmit two data streams and the low-priority (high SNR) stream should carry twice as many information bits. The simplest approach is to group subchannels 1 and 2 into the high-priority stream while subchannels $3, \ldots, 6$ form the low-priority stream. Before discussing in some detail the different transmitter and receiver structures which are possible with this subchannel grouping, we note that there are more possible choices, e.g.

1) For the same constellation $\mathcal{A}$, different mappings (bit labelings) $\mu(\cdot)$ can be chosen (e.g. Gray labeling, natural labeling, set partitioning labeling, etc.). While this choice does not affect the CM capacity, it leads to different MLC and BICM subchannel capacities.
2) For the same constellation $\mathcal{A}$ and mapping μ, another decoding order for MLC leads to different MLC subchannel capacities $C_{\mathsf{mlc}}(d)$.
3) For each binary subchannel or for each subchannel group, another code rate may be chosen. For MLC, restrictions which are discussed below, apply.
4) Any other subchannel grouping is possible for both MLC and BICM.

B. BICM and BICM-SIC

The simplest approach is BICM, which treats the bits and their related a posteriori probabilities (APP) or L-values at the receiver independently. With this approach, the sub-stream capacities with the selected subchannel grouping are given by

$$\text{stream A: } C_{\mathsf{bicm},\mathrm{A}} = C_{\mathsf{bicm}}(1) + C_{\mathsf{bicm}}(2)$$

$$\text{stream B: } C_{\mathsf{bicm},\mathrm{B}} = \sum_{d=3}^{6} C_{\mathsf{bicm}}(d)$$

and are plotted in Fig. 4. From these substream capacities, we obtain the required minimum SNR for a desired code rate R_c by setting $C_{\mathsf{high-prio}} = 2R_c$ and $C_{\mathsf{low-prio}} = 4R_c$. For our example and $R_c = 1/2$, we obtain from Fig. 4 the SNRs 4.2 dB and 9.7 dB, respectively. With the same modulation, we can choose different code rates to obtain other SNR thresholds.

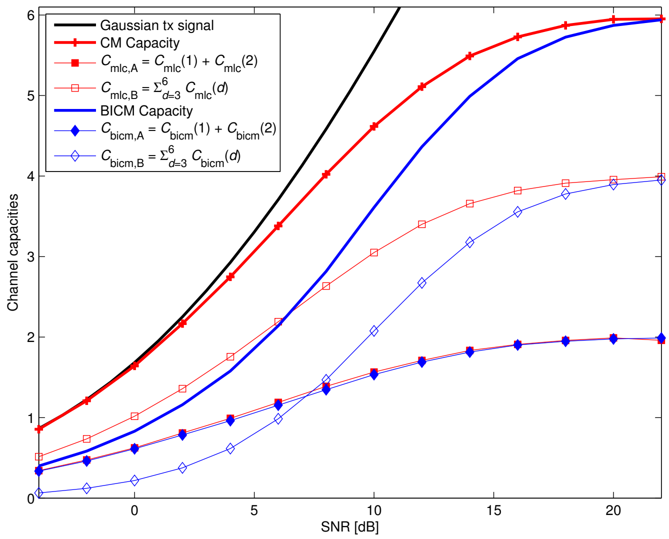

Figure 4. Capacities of the two data streams: stream A is composed of the binary subchannels 1 and 2, while subchannels $3, \ldots, 6$ form stream B.

This simple procedure, which mimicks the HM approach, is clearly suboptimum as it does not take advantage of the knowledge of the high-priority stream for decoding the low-priority stream. A straightforward improvement is to apply the decoded bits of the high-priority stream to cancel their impact on the other stream, as illustrated in Fig. 5. In addition to improving the performance of stream B, this step *reduces* the complexity of the demapper. Note that for this kind of interference cancellation, it is not necessary that the bit sequence c_A of the high-priority stream can be mapped to a symbol sequence.

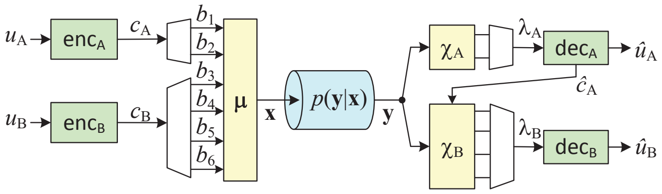

Figure 5. BICM-SIC: The binary subchannels in each substream are demapped independently.

C. MLC with Non-Binary Decoding

The direct application of MLC to the 6 binary subchannels has the disadvantage that for every subchannel a different code rate according to the capacities $C_{\mathsf{mlc}}(d)$ has to be selected. In addition, if the SNR thresholds are close and the codewords are not long enough, error propagation will limit the performance. This problem can be overcome by a non-binary decomposition as follows,

$$\begin{aligned} C_{\mathsf{cm}} &= I(b_1, \ldots, b_6; \mathbf{y}) \\ &= I(b_1 b_2; \mathbf{y}) + I(b_3, \ldots, b_6; \mathbf{y} \mid b_1 b_2) \end{aligned} \tag{6}$$

which defines the substream capacities

$$C_{\mathrm{A}} \triangleq I(b_1 b_2; \mathbf{y}) = C_{\mathsf{mlc}}(1) + C_{\mathsf{mlc}}(2) \tag{7}$$

$$C_{\mathrm{B}} \triangleq I(b_3, \ldots, b_6; \mathbf{y} \mid b_1 b_2) = \sum_{d=3}^{6} C_{\mathsf{mlc}}(d) \tag{8}$$

This decomposition corresponds to grouping the bits b_1, b_2 into a quaternary and the bits $b_3, \ldots, b_6$ into a 16-ary symbol and has the advantage that it allows to choose *one* code rate per substream (i.e. per subchannel group) without loss of capacity. A possible transmitter and receiver structure for this approach is shown in Fig. 6. The transmitter applies parallel encoding with two and four binary encoders per substream while the receiver applies two joint decoders which decode all codewords of the substream simultaneously. These decoders work on the joint APPs of each substream, which are defined by

$$\mathbf{p}_A = \begin{bmatrix} P[b_1 = 0, b_2 = 0 \mid \mathbf{y}] \\ P[b_1 = 0, b_2 = 1 \mid \mathbf{y}] \\ P[b_1 = 1, b_2 = 0 \mid \mathbf{y}] \\ P[b_1 = 1, b_2 = 1 \mid \mathbf{y}] \end{bmatrix},$$

$$\mathbf{p}_B = \begin{bmatrix} P[b_3 = 0, b_4 = 0, b_5 = 0, b_6 = 0 \mid \mathbf{y}] \\ P[b_3 = 0, b_4 = 0, b_5 = 0, b_6 = 1 \mid \mathbf{y}] \\ \vdots \\ P[b_3 = 1, b_4 = 1, b_5 = 1, b_6 = 1 \mid \mathbf{y}] \end{bmatrix}$$

and provide as output the corresponding binary codewords. For LDPC codes, this joint decoding is very similar to decoding in the extension fields $\mathbb{F}_4$ and $\mathbb{F}_{16}$ [10]. The parallel encoding with the same LDPC code per substream allows to apply joint decoding of multiple codewords without altering the Tanner graph and thus preserving the code properties. This allows to take advantage from the benefits of non-binary LDPC decoding while still using binary codes[1] . On the other hand, it is also possible to apply the usual binary decoders for the transmitter structure of Fig. 6. This will not exploit the full potential of the scheme but it will result in a performance very similar to BICM-SIC at a reduced complexity.

Instead of the parallel encoding of binary codes, non-binary codes in $\mathbb{F}_4$ and $\mathbb{F}_{16}$ and the corresponding decoders can also be applied to this structure. The performance and complexity with non-binary LDPC codes is in the same range as for parallel encoding.

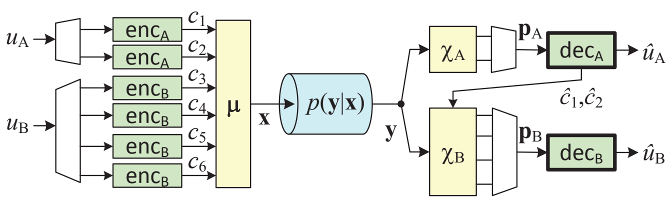

Figure 6. MLC with binary encoding and non-binary decoding

V. Simulation Results

A. *Performance with DVB-T2 Binary LDPC Codes*

This section evaluates the performance of MLC and BICM-SIC in terms of BER as a function of the SNR for a variety of code rates. For channel coding, we have chosen a family of binary LDPC from the DVB-T2 standard [11] with parameters listed in Table I. These codes, which have a common word length of $N = 16200$ bits, have been chosen for their rather wide range of rates and, being widely implemented, to facilitate reproducibility of our results. We assumed a 2×2 MIMO channel with 8-PSK modulation per antenna as described in Section III-C.

Table I
Code rates and message lengths of DVB-T2 code familiy, for "short FEC frame" with codeword length $N = 16200$.

R_c	1/5	4/9	3/5	2/3	11/15	7/9	37/45
K	3240	7200	9720	10800	11880	12600	13320

Fig. 7 shows the BER curves with the code rates $R_c \in \left\{\frac{1}{5}, \frac{2}{3}, \frac{37}{45}\right\}$ for both streams with MLC and joint, non-binary decoding. For comparison, we also included the performance of the UT-A (unitary-transform-Alamouti) scheme of Stauffer and Hochwald [6], [7], for which we applied the highest-rate code of Table I, which is similar in rate to the turbo code with $R_c = 5/6$ in their work. While on the transmitter side we reproduced the UT-A space-time code with QPSK-in-64-QAM modulation, on the receiver side, instead of the proposed QR decomposition, we implemented an optimum APP soft demapper, which has an unrealistically high computational complexity but provides an upper-bound for the performance of the UT-A scheme. For this reason, rather than a using this result for a rigorous comparison, we can see it as a reference.

[1]Although the terminology might be confusing, we note here that LDPC codes defined in an extension field of $\mathbb{F}_2$, which are typically denoted as *non-binary* LDPC codes, are actually a subset of binary LDPC codes.

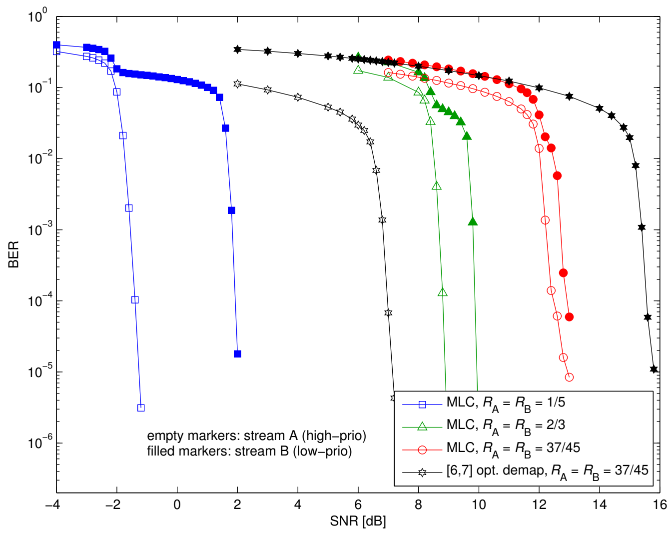

Figure 7. BER curves with MLC and joint decoding for a 2×2 MIMO channel and 8-PSK modulation. As comparison, the performance of the UT-A scheme from [6], [7] with optimum soft demapping is included.

Fig. 8 illustrates these results as rate-SNR pairs for a target BER of 10^{-4} for all codes of Table I and also for BICM-SIC. The diagram also includes the rate-SNR pairs for UT-A with the optimum demapper and the DVB-T2 $R_c = 37/45$ code, as well as the rate-SNR pairs directly taken from [6, Fig. 2]. Note that different code rates may be selected for the two streams. The only condition is that the SNR threshold for stream A has to be lower than for stream B, since the decoder of the low-priority stream B relies on the decoded high-priority stream A. This condition is satisfied for the chosen modulation and subchannel grouping if the code rate of stream A is lower or equal than the one of stream B.

The substream capacities C_A and C_B according to (7), (8) constitute an upper bound for the MLC scheme with joint decoding and in consequence also for BICM-SIC, although it is tight only for the former scheme. For all code rates, MLC is clearly superior to BICM-SIC. For this particular choice of parameters, we observe that the UT-A scheme shows better performance for stream A but is inferior for stream B.

B. *MLC with Non-Binary LDPC Codes*

Instead of the parallel encoding of binary codes in Fig. 6, we can employ non-binary LDPC codes in a suitable extension field of the binary field $\mathbb{F}_2 = \{0, 1\}$. We have applied two non-binary LDPC codes from the DAVINCI project [12] with parameters listed in Table II and 64-QAM per antenna. The

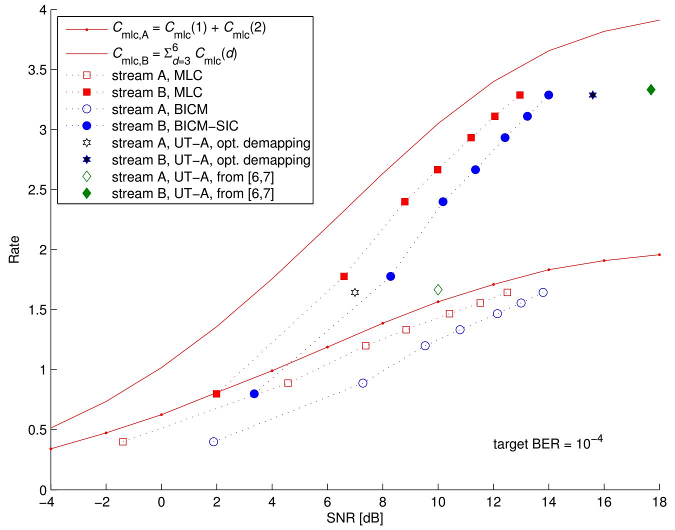

Figure 8. Rate-SNR pairs with MLC and BICM-SIC for all codes of Table I for a target BER of 10^{-4}. As comparison, the rate-SNR pairs of the UT-A scheme with optimum demapping and the DVB-T2 code of rate $R_c = 37/45$ and the result of [6], Fig. 2 are included.

word length in bits is for both codes $N_{\text{bin}} = N_A \log_2 q_A = N_B \log_2 q_B = 1440$ bits, while both codes share the same code rate of $R_c = 1/2$. With 64-QAM per antenna, we have 12 coded bits per channel use, of which 4 are allocated to stream A to form $2^4 = 16$-ary code symbols and the other 8 bits are gathered into $\mathbb{F}_{256}$-symbols for stream B. With this partition, we obtain a rather small difference in terms of required SNR for the two streams, as we can observe from Fig. 9. Like for parallel encoding, all possibilities for defining the substreams as described in Section IV-A apply also to non-binary codes.

Table II
CODE PARAMETERS FOR TWO NON-BINARY LDPC CODES

	Field size q	Word length in symbols	Code rate
stream A	$q_A = 16$	$N_A = 360$	$R_c = 1/2$
stream B	$q_B = 256$	$N_B = 180$	$R_c = 1/2$

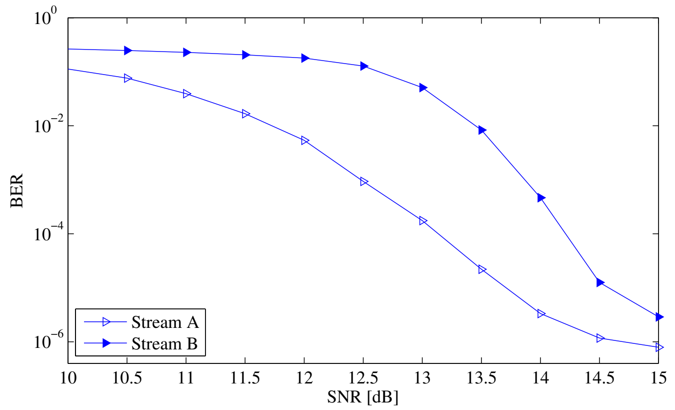

Figure 9. BER with non-binary LDPC codes and 64-QAM per antenna

VI. Conclusion

In this paper, we have presented several solutions for the simultaneous broadcast of different data streams that can operate with different service requirements. We have applied the approach of multilevel coding for multiple antennas, which is based on a solid information-theoretic framework and is optimum in the sense of sum capacity. For MLC, we have shown how substreams can be defined by grouping binary subchannels and we presented two solutions which allow to select one code rate per substream, i.e. for all involved binary subchannels, without a loss of capacity. Furthermore, we like to point out that MLC solutions allow for a higher flexibility in the antenna mapping and modulation than other approaches based on hierarchical modulation and space-time codes.

Acknowledgment

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306) and by the Catalan and Spanish Governments under SGR (2009SGR1046) and CICYT (TEC2011-29006-C03-01), respectively.

References

[1] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *IEEE Vehicular Technology Conference (VTC)*, Dresden, Germany, June 2013.

[2] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Int. Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)*, Okinawa, Japan, Nov. 2013.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[4] H. Méric, J. Lacan, C. Amiot-Bazile, F. Arnal, and M.-L. Boucheret, "Generic approach for hierarchical modulation performance analysis: Application to DVB-SH," in *IEEE Wireless Telecomm. Symp. (WTS)*, New York, April 2011.

[5] H. Méric, J. Lacan, F. Arnal, G. Lesthievent, and M.-L. Boucheret, "Combining adaptive coding and modulation with hierarchical modulation in satcom systems," *IEEE Trans. Broadcasting*, vol. 59, no. 4, pp. 627–637, Dec. 2013.

[6] E. Stauffer and B. Hochwald, "Low-complexity multi-stream space-time codes—Part I: Direct-sum codes and design criteria," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 81–88, Jan. 2012.

[7] B. Hochwald and E. Stauffer, "Low-complexity multi-stream space-time codes—Part II: Unitary-transform codes," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 375–383, Feb. 2012.

[8] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1361–1391, July 1999.

[9] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[10] S. Pfletschinger, "Joint decoding of multiple non-binary LDPC codewords," in *IEEE ICC Workshop on Massive Uncoordinated Access Protocols (MASSAP)*, Sydney, Australia, June 2014.

[11] *Digital Video Broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)*, ETSI Std. EN 302 755, 2011.

[12] D. Declercq, S. Pfletschinger, I. Andriyanova, E. Biglieri, G. Boccolini, and C. Poulliat, "D4.8 performance comparison of NB-LDPC, NB-root-LDPC, and NB-DGLDPC," European FP7 project DAVINCI, Tech. Rep., July 2010. [Online]. Available: http://www.ict-davinci-codes.eu/