GTTI 2009: sessione Trasmissione
Spectrally efficient LDPC coded modulations

Andrea Marinoni
Università degli Studi di Pavia
Dipartimento di Elettronica
Via Ferrata 1, 27100, Pavia, Italy
Email: andrea.marinoni@unipv.it

Thomas A. Courtade, Richard D. Wesel
University of California, Los Angeles
Electrical Engineering Department
Los Angeles, CA 90095-1594, USA
Email: {tacourta, wesel}@ee.ucla.edu

***Abstract*— In recent times the need of spectral efficiency has become a relevant topic for many communication systems, especially for wireless services. In order to achieve the best trade-off between bandwidth occupancy and error-rate performance, several structures that involve large constellations have been proposed in literature. This paper focuses on LDPC-coded systems using 16-QAM constellations on a channel affected by Additive White Gaussian Noise (AWGN). The LDPC codes that have been used include both binary and non-binary systems. In order to be compared, they have been designed such that they are equivalent in terms of blocklength, rate and average column weight. Simulation results show how the structure that involves a q-ary LDPC code outperforms the other schemes: new possible scenarios to be analyzed and ongoing works are then introduced.**

I. Introduction

With an ever-increasing demand for wireless services, the need for spectral efficiency in data communications has become an important topic. To alleviate the crowding of the radio-frequency spectrum, it is desirable to make more efficient use of currently allocated frequency bands. Historically, the most popular scheme to improve bandwidth-efficiency has been to utilize higher-order modulation. This approach allows more bits per transmitted symbol, but the higher symbol density requires increased power to achieve acceptable bit-error-rate (BER) performance. In order to achieve the best possible performance, capacity approaching codes as Turbo-Codes (TC) and Low-Density Parity-Check (LDPC) codes have been adopted by a multitude of systems - from storage devices to optical communications. LDPC codes [1] are algebraic codes characterized by a sparse parity-check (PC) matrix, H, having M rows and N columns. LDPC codes can be classified as either regular or irregular depending on their row and column degree-distributions. Regular LDPC codes have a parity check matrix in which all rows (and columns) have equal weight, while the irregular LDPC codes do not exhibit this property. Non-binary (or q-ary) LDPC codes have codewords (and also a PC matrix) whose symbols are elements of the finite field $GF(q)$, with $q > 2$. These non-binary LDPC codes typically have steeper bit-error-rate curves, however the decoding complexity is $O(Ntq^2)$, where N is the blocklength, t is the average column weight, and q is the alphabet width [4], [5]. Using their bipartite graph representation, [6] and [23] showed that LDPC codes may perform very close to capacity on AWGN channels and achieve capacity on binary erasure channels. Therefore, it is natural to ask if LDPC codes can improve the bit-error-rate performance of a code in a communication system that requires high bandwidth efficiency [20]. In this paper, we compare three different coding architectures, paying particular attention to the properties of the LDPC code selected for each one. The paper is organized as follows. In Section II the three different architectures are introduced and we highlight the features related to the application of LDPC codes to these architectures. Further, we comment on the bandwidth efficiency of each of the architectures. In Section III the simulation results are given, and we also discuss the more practical aspects of the code construction and decoding.

II. System model

In this section, we analyze the performance of a higher-order coded modulation system over an AWGN channel. In each system that we consider, the input to the modulator is encoded by an LDPC code whose properties depend on the particular system under consideration. At the receiver, the received signal is sent to the LDPC decoder. Depending on the transmitter model that was used, the receiver decodes in a manner consistent with how the transmitter encoded the message. In the three different architectures that will be introduced, the first two are based on binary LDPC codes, while the last is based on a q-ary LDPC code.

Here we make two notes. First, in the Multi-level coding architecture that we will introduce, the error correction coding is performed by means of p properly synchronized binary LDPC codes, where $p = log_2(q)$ and q represents the order of the modulation. In our error-rate performance analysis, we do not consider the influence of the inherent decoding delays associated such a structure. Second, all the LDPC codes used in this paper have been constructed using Quasi-Regular PC matrices [20], [22] generated by the Progressive Edge-Growth (PEG) algorithm [19].

Given a rate R and the average column weight (i.e. the average variable-node degree in the Tanner graph), $\overline{d}_v$, it is possible to compute the average row weight (i.e. the average check-node degree), $\overline{d}_c$ as follows:

$$\overline{d}_c = \frac{\overline{d}_v}{1-R}. \tag{1}$$

Furthermore, the column (variable-node) profile is provided by this rule:

$$d_{v_j} = \begin{cases} \lfloor \overline{d}_v \rfloor - \overline{d}_v + 1 & \text{if } j = \lfloor \overline{d}_v \rfloor \\ \lfloor \overline{d}_v \rfloor - \overline{d}_v & \text{if } j = \lfloor \overline{d}_v \rfloor + 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Where d_{v_j} represents the fraction of columns with weight j in the given PC matrix, and $\lfloor z \rfloor$ is defined as the largest integer less than or equal to z. Analogously, the row (check-node) profile can be computed as follows:

$$d_{c_j} = \begin{cases} \lfloor \overline{d}_c \rfloor - \overline{d}_c + 1 & \text{if } j = \lfloor \overline{d}_c \rfloor \\ \lfloor \overline{d}_c \rfloor - \overline{d}_c & \text{if } j = \lfloor \overline{d}_c \rfloor + 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Here, d_{c_j} represents the fraction of rows with weight j in the given PC matrix.

A. Turbo-like receiver

In the architecture given in Figure 1, the transmitted signal is a binary LDPC codeword that has been properly mapped to the constellation associated with the given higher-order modulation scheme.

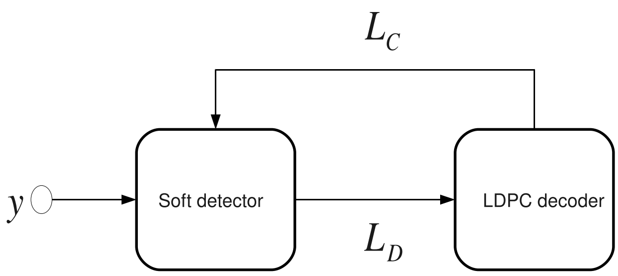

Fig. 1. Turbo iterative detection-and-decoding receiver for a LDPC coded system

At the receiver, the soft detector incorporates extrinsic information provided by the binary LDPC decoder, and the LDPC decoder incorporates soft information provided by the detector. Extrinsic information between the detector and decoder is exchanged in an iterative way until an LDPC codeword is found or a maximum number of iteration is performed [5], [12]. With LDPC codes, convergence to a codeword is easily detected by the receiver when the parity check equations are satisfied. The decoding Message Passing Algorithm (MPA) is described in detail in [4], [5]. In this architecture, the received vector **y** is demapped by a log-likelihood ratio (LLR) calculation for each of the coded bits included in the transmitted vector **x**. The extrinsic information provided by the detector is the difference of the soft-input and soft-output LLR values for the coded bits. For the κ-th code bit of **x**, x_κ, the extrinsic LLR value of the estimated bit is computed as follows:

$$\begin{aligned} L_D(x_\kappa) &= log \frac{P(x_\kappa = +1|\mathbf{y})}{P(x_\kappa = -1|\mathbf{y})} - log \frac{P(x_\kappa = +1)}{P(x_\kappa = -1)} \\ &= log \frac{P(x_\kappa = +1|\mathbf{y})}{P(x_\kappa = -1|\mathbf{y})} - L_C(x_\kappa), \end{aligned} \tag{4}$$

where $L_C(x_\kappa)$ is the extrinsic information of x_κ computed by the LDPC decoder in the previous turbo iteration. Note that $L_C(x_\kappa) = 0$ at the first iteration. Assuming the bits associated with **x** are statistically independent of one another, the $a\ priori$ probability $P(\mathbf{x})$ can be expressed in the following way:

$$P(\mathbf{x}) = \prod_{i=1}^{N} P(x_i) = \prod_{i=1}^{N} \left[1 + \exp(-\mathbf{x}^{x_i} \cdot L_C(x_i))\right], \tag{5}$$

where $\mathbf{x}^{x_i}$ corresponds to the value (either +1 or -1) of the i-th bit in the vector **x**.

B. Multilevel Coding

Imai's idea of multilevel coding (MLC) is to protect each address bit x_i of the constellation points by an individual binary code ξ_i at level i [3]. At the receiver, each code ξ_i is decoded individually starting from the lowest level and taking into account decisions of prior decoding stages. This procedure is called multistage decoding (MSD). In contrast to Ungerboeck's trellis coded modulation (TCM) [7]-[9], the MLC approach provides flexible transmission rates because it decouples the dimensionality of the signal constellation from the code rate. Furthermore, any kind of code may be used as component code. Although MLC offers excellent asymptotic coding gains, it achieved only theoretical interest in the past. In practice, system performance was severely degraded due to high error rates at low levels. A straightforward generalization of Imai's approach is to use q-ary component codes based on non-binary partitioning of the signal set; however, using binary codes in conjunction with multilevel codes turns out to be asymptotically optimal. For practical coded modulation schemes where boundary effects have to be taken into account, Huber and Kofman [13], [14] proved that the capacity of the adopted modulation scheme can be achieved by multilevel codes together with MSD if and only if the individual rates of the component codes are properly chosen. Here it is assumed that the signal points are equiprobable and the partitioning is regular. Further yet, in [11], the authors generalized these results to arbitrary signaling and labeling of signal points by means of the chain rule for mutual information. In this way we can create a model with virtually independent parallel channels for each address bit at the different partitioning levels, these levels are called equivalent channels. In order to better understand the idea beneath this concept, consider the previously described modulation scheme with $L = 2^\lambda$ signal points. Since each of the signal points exists in a D-dimensional signal space, every signal point is taken from the signal set $T = \{\tau_0, \tau_1, \ldots, \tau_{L-1}\}$ where $T \subset \mathbf{R}^D$ ($\mathbf{R}$ being the field of real numbers).

When considering the AWGN channel, the channel output signal points come from the alphabet $Y = \mathbf{R}^D$. In order to create effective error-correcting codes for such an L-ary signal alphabet, labels have to be assigned to each signal point, using a bijective mapping between the set of all possible $\mathbf{x}$ and T. Since the mapping is bijective independently of the partitioning strategy, the mutual information, $I(Y;T)$, between the transmitted signal point $\tau \in T$ and the received signal point $y \in Y$ equals the mutual information, $I(Y; X_0^{\lambda-1})$, between the mapper binary input $\mathbf{x} \in \{0,1\}^\lambda$ and the received signal point $y \in Y$. Here we use the notation $X_a^b = [X_a, X_{a+1}, \ldots, X_b]$. Note that the physical channel is characterized by the set $\{f_Y(y|\tau)|\tau \in T\}$ of conditional probability density functions of the received point y given the transmitted signal point τ. Applying the chain rule of mutual information, we obtain the following:

$$\begin{aligned} I(Y;T) &= I(Y;X_0^{\lambda-1}) \\ &= I(Y;X_0) + I(Y;X_1|X_0) + \ldots \\ &\quad + I(Y;X_{\lambda-1}|X_0^{\lambda-2}). \end{aligned} \tag{6}$$

Essentially, this shows that the transmission of binary vectors over the physical channel can be separated into the parallel transmission of each single bit x_i over λ equivalent channels with $x_0, \ldots, x_{i-1}$ known. In other words, the mutual information $I(Y; X_\kappa | X_0^{\kappa-1})$ of the κ-th equivalent channel can be easily calculated as the following:

$$I(Y;X_\kappa|X_0^{\kappa-1}) = \quad I(Y;X_\kappa^{\lambda-1}|X_0^{\kappa-1}) - I(Y;X_{\kappa+1}^{\lambda-1}|X_0^{\kappa}). \tag{7}$$

Since the subsets at one partitioning level may not be congruent, the mutual information $I(Y; X_\kappa, \ldots, X_{\lambda-1})$ is calculated by averaging over all possible combinations of $x_0^{\kappa-1} = x_0, \ldots, x_{\kappa-1}$. Specifically:

$$I(Y;X_\kappa^{\lambda-1}|X_0^{\kappa-1}) = E_{x_0^{\kappa-1}\in\{0,1\}^\kappa}\left[I(Y;X_\kappa^{\lambda-1}|x_0^{\kappa-1})\right]. \tag{8}$$

Assuming the bits in the lower levels, $x_0^{\kappa-1}$, are fixed, we see that the κ-th equivalent channel is characterized by the pdf $f_y(y|x_\kappa, x_0^{\kappa-1})$. The underlying signal subset for the equivalent κ-th modulator is given by $T(x_0^{\kappa-1})$, which denotes the partition of the signal set with the set of bits $x_0^{\kappa-1}$ in common. Since the binary symbol x_κ is potentially represented several times in this subset, the signal point τ is in effect chosen uniformly from the subset $T(x_0^\kappa)$. Therefore, $f_Y(y|x_\kappa, x_0^{\kappa-1})$ is given by the expected value of the pdf $f_Y(y|\tau)$ over all signal points τ out of the subset $T(x_0^\kappa)$, as follows:

$$\begin{aligned} f_Y(y|x_\kappa, x_0^{\kappa-1}) &= E_{\tau\in T(x_0^\kappa)}\left[f_Y(y|\tau)\right] \\ &= \frac{1}{P(T(x_0^\kappa))}\sum_{\tau\in T(x_0^\kappa)} P(\tau)\cdot f_Y(y|\tau). \end{aligned} \tag{9}$$

The κ-th equivalent channel is completely characterized by a set of probability density functions $\mathbf{f}_Y(y|x_\kappa)$ of the received point y if the binary symbol x_κ is transmitted. Moreover, since the subset for transmission of symbol x_κ depends on the symbols at levels 0 through $\kappa - 1$, the set of pdf's, $\mathbf{f}_Y(y|x_\kappa)$, is the set of $f_Y(y|x_\kappa, x_0^{\kappa-1})$ for each possible combination of $x_0^{\kappa-1}$. Specifically:

$$\mathbf{f}_Y(y|x_\kappa) = \left\{f_Y(y|x_\kappa, x_0^{\kappa-1})|x_0^{\kappa-1}\in\{0,1\}^\kappa\right\}. \tag{10}$$

The multilevel coding approach together with its multistage decoding procedure is a consequence of the chain rule described in (6). The binary symbols x_i, $i = 0, \ldots, \lambda-1$, come from independently encoding different data symbols. Each binary encoder generates words $\mathbf{x}_i = [x_{i_1}, \ldots, x_{i_N}]$ of the component code ξ_i, where $x_{i_j} \in \{0,1\} \forall j \in \{1, \ldots N\}$. Even if the choice of the component codes is arbitrary, we assume that the blocklength, N, of each code, ξ_i, equal for all levels. Nevertheless, we can still define different rates for every ξ_i, resulting in different lengths of the encoder inputs, denoted K_i.

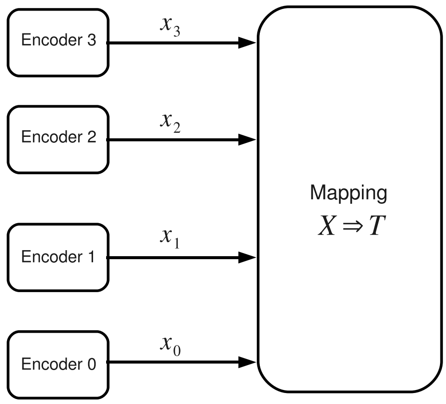

Fig. 2. Multilevel encoder for 16-ary modulation

Using this notation, we define the rate of the i-th encoder to be $R_i = K_i/N$. The codeword symbols, $x_{i_j} \in \mathbf{x}_i$, form the the binary address $\mathbf{x}_j = \left[x_{0_j}, \ldots, x_{\lambda-1_j}\right]$, which is mapped to the signal point τ_j (Figure 2). The code rate, R, of this scheme is equal to the sum of the individual code rates, R_i, as follows:

$$R = \sum_{i=0}^{\lambda-1} R_i = \sum_{i=0}^{\lambda-1} \frac{K_i}{N}. \quad (11)$$

As determined by the MSD procedure, the component codes ξ_i are succesively decoded by the corresponding decoders, D_i (Figure 3). At the i-th stage, D_i processes the block, $\mathbf{y} = [y_1, \ldots, y_N]$ ($y_j \in Y$), of received signal points using the decisions, $\widehat{\mathbf{x}}_l$, from the i previous decoding stages (i.e. $l = 0, \ldots, i-1$).

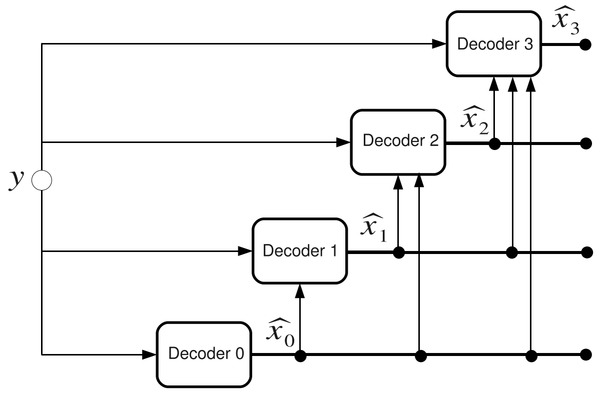

Fig. 3. Multistage decoding for 16-ary modulation

As noted earlier, this procedure necessarily introduces delays in the decoding process. In order to satisfy the chain rule (6) and preserve the mutual information, we require that the estimated symbol, $\widehat{\mathbf{x}}_l$, is equal to the transmitted symbol, $\mathbf{x}_l$. Therefore, if assume that error free decisions are generated by the decoders D_i, MSD can be interpreted as an implementation of the chain rule (6), and hence is mutual information preserving. In order to approach channel capacity, we need to maximize the mutual information over all controllable parameters. Usually, these are the *a priori* probabilities of the signal points. Therefore, we require a specific channel-input probability distribution, $P(\tau)$, in order to achieve the channel capacity, C. These probabilities can not be optimized independently for each individual level, and hence we must consider the entire signal set. The capacity of the i-th equivalent channel, C_i, is given by the respective mutual informations, $I(Y; X_i | X_0^{i-1})$, corresponding to the channel input probabilities. C_i is then given as follows:

$$\begin{aligned} C_i &= I(Y; X_i | X_0^{i-1}) \\ &= E_{x_0^{i-1}}\left[C(T(x_0^{i-1}))\right] - E_{x_0^i}\left[C(T(x_0^i))\right], \end{aligned} \quad (12)$$

where $C(T(x_0^i))$ denotes the capacity when using the subset $T(x_0^i)$ with *a priori* probabilities $P(\tau)/P(T(x_0^i))$. At this point, it is possible to determine the capacity $C = C(T)$ for a 2^λ-ary digital modulation scheme given the *a priori* probability distribution, $P(\tau)$, of the signal points $\tau \in T$. In particular, C is equal to the sum of the capacities of the equivalent channels, C_i, in the MLC scheme:

$$C = \sum_{i=0}^{\lambda-1} C_i. \quad (13)$$

The capacity, C, can be approached via MLC-MSD if the individual rates, R_i, are chosen to be arbitrarily close to (but not greater than) the capacities of the equivalent channels C_i. In order to lower the latency of the MLC system, a different decoding scheme has been studied in [10] and [11]. In the MLC with Parallel Independent Decoding (PID) structure each decoder D_i does not use the decisions of the other levels $j \neq i$. In [11], the authors showed how the mutual information of the modulation scheme can be approached with MLC-PID if and only if the rate R_i of each code is set in order to fulfill $R_i = I(Y; X_i)$. Moreover, they showed that the MLC-PID approach represents a suboptimal solution of an optimum coded modulation scheme and that the capactiy of such a scheme strongly depends on the particular labeling of signal points. However, they also showed how the gap to an optimum scheme can be very small using a Gray labeling of the signal points.

C. Combination of q-ary LDPCC and q-ary modulation

In this final method that we analyze, we combine LDPC codes over $GF(q)$ ($q = 2^p$, p a positive integer) with q-ary modulation to achieve bandwidth-efficient transmission (Figure 4). For a chosen code rate, R, and a blocklength, N, it is necessary to find a parity-check (PC) matrix, $H = \{h_{ij}\}_{i=1,\ldots,M, j=1,\ldots,N}$, where $h_{ij} \in GF(q)$ and $R = 1 - \dfrac{M}{N}$. In this manner, the $K = NR$ information symbols and the M parity symbols are encoded into a q-ary vector $\mathbf{x} \in GF(q)^N$. After q-ary LDPC encoding, the N elements of $\mathbf{x}$ are mapped into the modulated sequence $\mathbf{s} = \{s_j\}_{j=1,\ldots,N}$. This sequence depends on the address given by $\mathbf{x}_b = \left\{\overline{x}_b^j\right\}_{j=1,\ldots,N}$, where $\overline{x}_b^j = \left\{x_{b_k}^j\right\}_{k=0,\ldots,p-1}$ is the binary representation of the non-binary codeword symbol x_j. Therefore, the bandwidth efficiency of this structure is equal to $R \cdot p$.

At the receiver, the output of the AWGN channel may be expressed as:

$$y_\kappa = s_\kappa + n_\kappa = (s_{\kappa_I} + j s_{\kappa_Q}) + (n_{\kappa_I} + j n_{\kappa_Q}) = y_{\kappa_I} + j y_{\kappa_Q}, \quad (14)$$

where $\kappa = 1, \ldots, N$ and $n_{\kappa_I}, n_{\kappa_Q}$ are two independent noises with the same variance, σ^2, related to the in-phase and quadrature component of the modulated signal. Starting with $P(y_\kappa | s_\kappa)$, and using the Bayes' theorem [20], the *a posteriori* probability distribution can be written as:

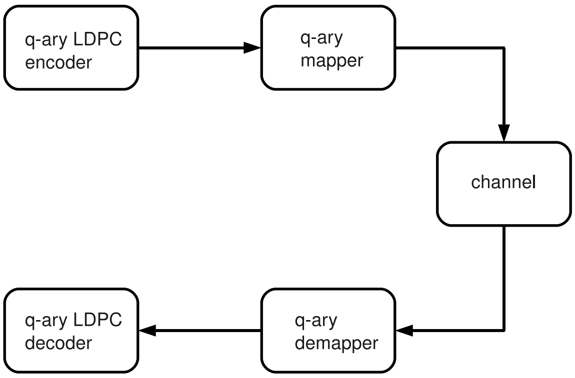

Fig. 4. Block diagram of the structure that combines q-ary LDPC code and q-ary modulation

$$P(s_\kappa|y_\kappa) = exp\left(-\frac{(y_{\kappa_I} - s_{\kappa_I})^2 + (y_{\kappa_Q} - s_{\kappa_Q})^2}{2\sigma^2}\right). \quad (15)$$

The probabilities in (15) are used to initialize the Message Passing algorithm in the decoder [4]. We remark here that the computational complexity of the algorithm provided by [4] may be reduced by employing the Fast Fourier Transform (FFT) or the Fast Hadamard Transform (FHT) approach [20].

III. Simulation results

In this section, we discuss simulation results obtained by implementing the three structures introduced in the previous section. In each of these implementions, we use Gray-mapped 16-QAM modulation, a global bandwidth efficiency of 2 bits/symbol (i.e. a coding rate equal to 0.5), and an input blocklength of 5000 bits per codeword. For the system described in subsection II-A, the binary LDPC code has blocklength $N = 10000$ and rate 0.5. The variable-node degree distribution, following the notation introduced in [21] and [22] and according to (2) and [20] , is λ_8 = 0.2 and λ_9 = 0.8, where $\lambda(x) = \sum_{i=2}^{d_v} \lambda_i x^{i-1}$, and d_v is the maximum symbol-node degree. In what follows, the maximum number of iterations between the soft-detector and the LDPC decoder is set to 30 [12].

In order to make a fair comparison between architectures, the PC matrix of the 16-ary LDPC code used in the architecture introduced in II-C also has a rate equal to 0.5, while the blocklength N is set to 2500 symbols, and the variable-node degree distribution is λ_2 = 0.8 and λ_3 = 0.2 [20]. For this decoding architecture and the MSD architecture, we set the maximum number iterations performed by the LDPC decoder to 25.

The MLC structure is defined by $4 = log_2(16)$ binary LDPC codes corresponding to each address bit. . They each have blocklength $N = 2500$ and variable-node degree distribution λ_2=0.8 and λ_3=0.2. Each rate is defined to be $[R_0, R_1, R_2, R_3] = [0.337, 0.663, 0.337, 0.663]$ in the MSD case and $[R_0, R_1, R_2, R_3] = [0.349, 0.651, 0.349, 0.651]$ in the PID case. These values agree with the ones in [10], since 16-QAM can be interpreted as product of two indepedent 4-PAM costellations. The simulation results in Figure 5 show how the q-ary LDPC code architecture from subsection II-C outperforms the binary LDPC Turbo-like architecture of subsection II-A. In particular, the gain is about 2.5 dB in terms of Signal-to-Noise Ratio (SNR). Moreover, we also observed that the MLC architectures outperform the Turbo-like architecture, however they do not perform as close to capacity as the q-ary LDPC coded architecture introduced in subsection II-C.

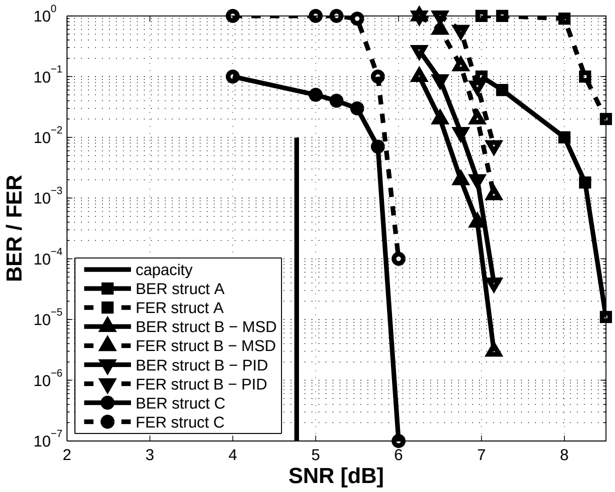

Fig. 5. Performance of the analyzed architectures on the AWGN channel.

IV. Conclusions

Three higher-order coded modulations employing LDPC codes were introduced and analyzed in order to study their corresponding trade-offs between bandwidth-effciency and bit-error-rate performance.

Simulation results for 16-QAM modulation schemes showed that the best performance can be achieved by using a code whose alphabet size matches the modulation order. Consequently, using such an architecture, associating each non-binary coded symbol to a modulated symbol appears to be the best solution in an environment (such the wireless one) where high bandwidth-efficiency and good error-correction capability is desirable.

Ongoing research that promises high spectral-efficiency includes the analysis of different structures. Future directions for research could investigate the behavior of the proposed architectures over different channels and with different modulation schemes , as well as different typologies of LDPC codes, having different codeword length or degree-distribution profile as in [24]. Further, since a complete analysis of decoding architectures in terms of latency and complexity is lacking in the literature, future works could potentially highlight such features.

REFERENCES

[1] R.G. Gallager, *Low-Density Parity-Check Codes,* Cambridge, MA, MIT Press, 1963.

[2] R.G. Gallager, *Information Theory and Reliable Communication,* New York, Wiley, 1968.

[3] H. Imai, S. Hirakawa, "A new multilevel coding method using error correctin codes," *IEEE Trans. Inform. Theory* vol IT-23, pp.371-377, May 1977.

[4] M.C. Davey, D. MacKay, "Low-Density Parity-Check Codes over GF(q)," *IEEE Communications Letters*, vol. 2, no. 6, June 1998.

[5] M.C. Davey, '"Error-Correction Using Low-Density Parity-Check Codes," Ph.D. Thesis, University of Cambridge, UK, December 1999.

[6] N. Wiberg, '"Code and decoding on general graphs," Dissertation no. 440, Dept. Elec. Eng., Linkoping Univ., Linkoping, Sweden, 1996.

[7] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory,* vol. IT-28, pp. 55-67, Jan 1982.

[8] G. Ungerboeck, "Trellis coded modulation with redundant signal sets, part I," *IEEE Commun. Mag.*, vol. 25, pp. 5-11, Feb 1987.

[9] G. Ungerboeck, "Trellis coded modulation with redundant signal sets, part II," *IEEE Commun. Mag.*, vol. 25, pp. 12-21, Feb 1987.

[10] J. Hou, P.H. Siegel, L.B. Milstein, H.D. Pfister, "Capacity-Approaching Bandwidth-Efficient Coded Modulation Schemes Based on Low-Density Parity-Check Codes," *IEEE Trans. Inform. Theory,* vol. 49, no. 9, Sept. 2003.

[11] U. Wachsmann, R.F.H. Fischer, J.B. Huber, "Multilevel Codes: Theoretical Concepts and Practical Design Rules," *IEEE Trans. Inform. Theory,* vol. 45, no. 5, July 1999.

[12] A. Matache, C. Jones, R.D. Wesel, "Reduced Complexity MIMO Detectors for LDPC Coded Systems," in proc. *2004 IEEE Military Communications Conference*, 31 Oct.-3 Nov. 2004, pp 1073-1079 .

[13] J. Huber, U. Wachsmann, "Capacities of equivalent channels in multilevel coding systems," *Electron. Lett.*, vol. 30, pp. 557-558, Mar. 1994.

[14] Y. Kofman, E. Zehavi, S. Shamai, "Analysis of a multilevel coded modualtion system," in proc. *1990 Bilkent Int. Cond. New Trends in Communications, Control and Signal Processing*, pp. 376-382, Ankara, Turkey, July 1990.

[15] G. Hosoya, H. Yagi, S. Hirasawa, "Modification Methods for Construction and Performance Analysis of Low-Density Parity-Check Codes over the Markov-Modulated Channel," in proc. *International Symposium on Information Theory and its Applications, ISITA 2004* , October 10-13 2004, Parma, Italy.

[16] E.A. Krouk, S.V. Semenov, "Low-Density Parity-Check Burst Error-Correcting Codes," in proc. *2nd International Workshop Algebraic and combinatorial coding theory*, pp. 121-4, 1990.

[17] E.A. Krouk, S.V. Semenov, *Error Correcting Coding and Security for Data Networks*, Mac-Graw Hill.

[18] E. Eleftheriou, S. Õlçer, "Low-Density Parity-Check Codes for Digital Subscriber Lines," in proc. *IEEE International Conference on Communications, ICC 2002*, April 28-May 2, 2002.

[19] X.Y. Hu, E. Eleftheriou, D.M. Arnold, "Regular and Irregular Progressive Edge-Growth Tanner Graphs," *IEEE Trans. on Information Theory*, vol. 51, no.1, Jan. 2005.

[20] B. Rong, T. Jiang, X. Li, M.R. Soleymani,"Combine LDPC Codes Over GF(q) With q-ary Modulations for Bandwidth Efficient Transmission," *IEEE Trans. on Broadcasting*, vol. 54, no. 1, March 2008.

[21] M. Luby, M. Mitzenmacher, A. Shokrollahi, D. Spielman, V. Stemann, "Practical Loss-Resilient Codes," in proc. 29^{th} *Annu. ACM Symp. Theory of Computing*, 1997, pp. 150-159.

[22] T. Richardson, A. Shokrollahi, R. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, Feb. 2001.

[23] T. J. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599-618, Feb. 2001.

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