Granada *et al. EURASIP Journal on Wireless Communications and Networking* (2019) 2019:11
https://doi.org/10.1186/s13638-018-1330-z

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Asymptotic BER EXIT chart analysis for high rate codes based on the parallel concatenation of analog RCM and digital LDGM codes

Imanol Granada[1*], Pedro M. Crespo[1] and Javier Garcia-Frías[2]

Abstract

This paper proposes an extrinsic information transfer (EXIT) chart analysis and an asymptotic bit error rate (BER) prediction method to speed up the design of high rate RCM-LDGM hybrid codes over AWGN and fast Rayleigh channels. These codes are based on a parallel concatenation of a rate compatible modulation (RCM) code with a low-density generator matrix (LDGM) code. The decoder uses the iterative sum-product algorithm to exchange information between the variable nodes (VNs) and the two types of constituent check nodes: RCM-CN and LDGM-CN. The novelty of the proposed EXIT chart procedure lies on the fact that it mixes together the analog RCM check nodes with the digital LDGM check nodes, something not possible in previous multi-edge EXIT charts proposed in the literature.

Keywords: Rate compatible modulation (RCM), Low-density generation matrix (LDGM), EXIT charts, BER prediction, Joint source-channel coding (JSCC)

1 Introduction

We propose an extrinsic information transfer (EXIT) chart analysis and an asymptotic bit error rate (BER) prediction method to speed up the design of high rate (greater than 2 bits per complex channel symbol) RCM-LDGM hybrid channel codes for the transmission of memoryless binary sources over additive white Gaussian noise (AWGN) and fast Rayleigh channels. These hybrid codes consist of the parallel concatenation of a rate compatible modulation (RCM) code (see, e.g., [1, 2]) and a low-density generator matrix (LDGM) code (see, e.g., [3, 4]). In what follows, we will refer to these schemes as parallel RCM-LDGM codes. Both uniform and non-uniform sources are considered. The reason for considering non-uniform sources is that many data sources (e.g., image or speech signals) are non-uniformly distributed, containing substantial amount of natural redundancy [5–8]. Even when these sources are compressed, they still exhibit a residual redundancy due to the sub-optimality of the compression scheme [9].

RCM codes generate random projections (RP) from weighted linear combinations and are able to achieve smooth rate adaptation in a broad dynamic range. However, they present error floors at high signal to noise ratios (SNRs). In order to solve this drawback, [10, 11] suggested the use of an LDGM code in parallel with the RCM code, aiming at reducing the error floor. Simulation results in [10, 11] show that the parallel RCM-LDGM code outperforms RCM schemes significantly, achieving a performance close to the Shannon limit if suitable design parameters are chosen.

One of the main advantages of this class of high rate RCM-LDGM codes over other high rate codes, such as the widely adopted bit-interleaved coded modulation (BICM) [12], is the easiness of performing adaptive coded modulation (ACM). Conventional ACM is done by selecting the best combination of channel coding and modulation based on the estimated channel condition. Due to the limited number of rate combinations, a stair-shaped rate curve is often obtained. Moreover, ACM requires instant and accurate channel estimation. Due to their intrinsic

*Correspondence: igranada@tecnun.es
[1]Department of Basic Science, University of Navarra, Mikeletegi Pasealekua, 48, 20018 San Sebastian, Spain
Full list of author information is available at the end of the article

design, RCM codes are well suited to overcome these adaptation challenges (refer to [2] for a comparison with other ACM schemes). The reason is that their coded symbols are generated by a set of weighted linear combinations of the source binary symbols. By varying the number of linear combinations on a per symbol basis, a smooth rate adaptation is possible.

Another advantage of RCM-LDGM codes is in the transmission of non-uniform memoryless sources. Existing low rate joint source-channel coding schemes [7, 8, 13, 14] present a gap to the Shannon theoretical limit of about 2 dB for sources with low non-uniformity. However, this gap increases when the source becomes more non-uniform, i.e., when its entropy decreases. Unlike these low rate joint source-channel codes, it is shown in [11] that RCM-LDGM codes are able to maintain the gap to the theoretical limit as the non-uniformity increases, while keeping very large throughputs. Their robustness against channel and source variations, together with the fact that smooth rate adaption is possible, makes RCM-LDGM codes excellent candidates in applications where channel and source variations are encountered. However, the proposed RCM-LDGM codes found in the literature [10, 11], have been designed by a trial-error procedure, something that requires a large amount of simulation time. Here, to circumvent this design drawback, we propose an EXIT chart analysis that facilitates the selection of suitable code parameters.

EXIT charts were first introduced in [15] to analyze and design an iterative coding scheme. Later, [16] proposed a curve fitting procedure based on EXIT charts to design a low-density parity check (LDPC) code valid for modulation and detection. Due to the iterative decoding nature of parallel RCM-LDGM codes, EXIT charts are a good method to visually explore the iterative exchange of information that occurs in the decoders of these schemes. The authors in [17] were the first to use EXIT charts as a design aid for pure RCM codes. However, the EXIT analysis for pure RCM codes is not valid in our case, since two different types of check nodes, RCM and LDGM, have to be considered (refer to Fig. 1). EXIT chart analysis considering multiple edge node types has been extensively investigated in the literature, e.g., in [18] multi-edge type EXIT charts are used to design the bit mapping for LDPC coded BICM schemes, and in [19], EXIT charts are used to optimize the bit mapping of LDPC coded modulation with APSK constellations. Irregular LDPC codes are also examples of EXIT charts with different check nodes. However, RCM-LDGM codes present the added difficulty of mixing analog and digital check nodes, and therefore, previous strategies cannot be directly applied.

The contribution of this paper is twofold:

1. Developing an EXIT chart analysis that is able to deal with the check node disparity encountered in parallel RCM-LDGM codes when driven by binary memoryless sources (both uniform and non-uniform) transmitted over AWGN and fast fading Rayleigh channels
2. Assessment of the time savings achieved by using the EXIT chart analysis, rather than Monte Carlo simulations, for BER predictions

The remainder of the paper is organized as follows. Section 2 briefly reviews previous work on the design of RCM and parallel RCM-LDGM codes. Section 3 presents the proposed EXIT chart analysis and BER prediction for RCM-LDGM codes. Section 4 evaluates the proposed EXIT chart-BER prediction method, comparing the predicted BER with simulation results. Finally, Section 5 concludes this paper.

2 Background: RCM and RCM-LDGM code design

Consider a point-to-point communication system where a binary memoryless source with distribution $(p_0; p_1 = 1-p_0)$ transmits K bits $\mathbf{u} = (u_1, u_2, \ldots, u_K)^\top \in \{0,1\}^{K\times 1}$,

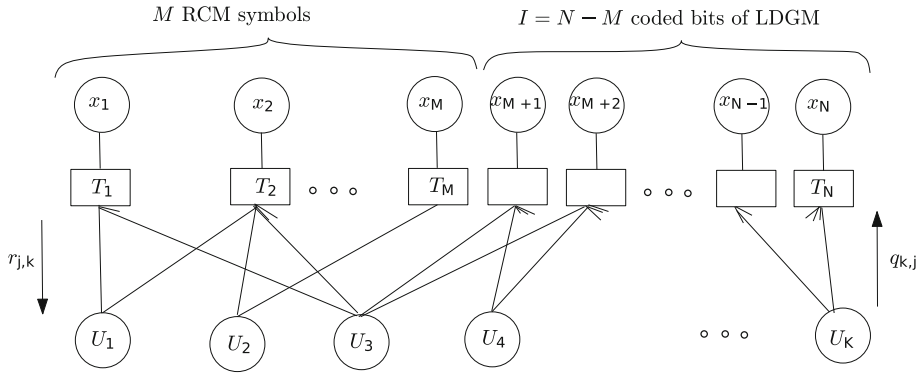

Fig. 1 Factor graph representing the parallel RCM-LDGM code

across an AWGN channel, to a far end receiver. To that end, the source symbols $\mathbf{u}$ are encoded by a rate $R = K/N$ (bits per real dimension) parallel RCM-LDGM encoder and quadrature amplitude modulation (QAM) modulated before being transmitted. Let

$$\mathbf{x} = (x_1 + \mathrm{i}x_2, x_3 + \mathrm{i}x_4, \ldots, x_{N-1} + \mathrm{i}x_N)^\top$$

be the sequence of $N/2$ complex baseband modulated symbols to be transmitted, where $x_i \in \mathbb{R}$ denotes the coded symbols at the output of the RCM-LDGM encoder. Assuming independence of the coded symbols $\{x_i\}$, a set of sufficient statistics to estimate $\mathbf{u}$ is given by the output of an equivalent discrete time AWGN channel $\mathbf{y} = (y_1, y_2, \ldots, y_N)^\top \in \mathbb{R}^{N\times 1}$,

$$y_i = x_i + n_i, \qquad i \in \{1, 2, \ldots, N\}$$

where $\{n_i\}_{i=1}^N$ are realizations of i.i.d real Gaussian random variables (RVs) with zero mean and variance $N_0/2$ (i.e., $N_i \sim \mathcal{N}(0, N_0/2)$). At the receiver side, the decoder estimates the source symbols $\mathbf{u}$ from $\mathbf{y}$.

For the sake of clarity in the exposition, we begin by providing a succinct overview of the key concepts of RCM and LDGM codes before covering parallel RCM-LDGM codes.

2.1 Rate compatible modulation (RCM) codes

An RCM code of rate K/M bits per real dimension is characterized[1] by an $M \times K$ sparse generator matrix G. Let $\mathcal{D} \subset \mathbb{N}$ be a multiset[2] with $d^{(c)}_{\mathrm{RCM}}/2$ elements where $\mathbb{N}$ is the set of natural numbers (positive integers). The entries of G belong to $\pm\mathcal{D}$. As an example, let $d^{(c)}_{\mathrm{RCM}} = 8$ and assume K to be divisible by $d^{(c)}_{\mathrm{RCM}}$. Then, the construction of matrix G is given by the following steps:

1 Construct the $K/2 \times K$ matrix G_0 as

$$G_0 = \begin{bmatrix} \Pi(D_{d_3}) & \Pi(D_{d_4}) & \Pi(D_{d_1}) & \Pi(D_{d_2}) \\ \Pi(D_{d_1}) & \Pi(D_{d_2}) & \Pi(D_{d_3}) & \Pi(D_{d_4}) \\ \Pi(D_{d_4}) & \Pi(D_{d_3}) & \Pi(D_{d_2}) & \Pi(D_{d_1}) \\ \Pi(D_{d_2}) & \Pi(D_{d_1}) & \Pi(D_{d_4}) & \Pi(D_{d_3}) \end{bmatrix},$$

where $\Pi(\cdot)$ denotes random column permutations of a matrix, and D_{d_l} is a $K/8 \times K/4$ sparse matrix given by

$$D_{d_l} = \begin{bmatrix} d_l & -d_l & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & d_l & -d_l & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & 0 & d_l & -d_l & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & d_l & -d_l \end{bmatrix},$$

with $d_l \in \mathcal{D}$, for $l \in \left\{1, \ldots, d^{(c)}_{\mathrm{RCM}}/2\right\}$.

2 Vertically stack as many G_0s as needed. Note that we should keep only as many rows as needed in the last stacked G_0 matrix so that the required $M \times K$ G matrix is obtained.

Observe that $d^{(c)}_{\mathrm{RCM}}$ gives the number of nonzero entries of any row of G. Similarly, we denote by $d^{(\nu_k)}_{\mathrm{RCM}} \geq 2$ the number of nonzero entries of column k of matrix G, and by $\overline{d}^{(\nu)}_{\mathrm{RCM}}$ its average value, i.e, $\overline{d}^{(\nu)}_{\mathrm{RCM}} = \frac{1}{K}\sum_{k=1}^{K} d^{(\nu_k)}_{\mathrm{RCM}}$.

Given the message $\mathbf{u}$, the jth RCM symbol, x_j, is obtained as

$$x_j = [G\mathbf{u}]_j = \sum_{i=1}^{K} g_{j,i}u_i, \qquad j \in 1, \ldots, M,$$

where $[\,\cdot\,]_j$ is the element at row j and $g_{j,i}$ denotes entry (j, i) of matrix G, where these operations are in the real field.

2.2 Low-density generator matrix (LDGM) codes

Low-density generator matrix codes, or LDGM codes, are a binary class of codes with a sparse generator matrix G_L. In this work, we will consider systematic LDGM codes, i.e., codes whose generator matrix is of the form $G_L = [I_K|P]$, where I_K is the identity matrix of size K and P is a $K \times I$ sparse matrix. We will consider regular matrices P, which are characterized by the pair $\left(d^{(\nu)}_{\mathrm{LDGM}}, d^{(c)}_{\mathrm{LDGM}}\right)$, denoting the number of nonzero elements of a column and of a row, respectively. LDGM codes are a subclass of LDPC codes, where the parity check is given by the sparse matrix $H = [P^\mathsf{T}|I_I]$. Since the generator matrix is sparse, the encoding of LDGM codes can be done in linear time and the codeword of length $N = K + I$ can be written as

$$\begin{aligned}\mathbf{c} &= [c_1, c_2, \ldots, c_N] = \mathbf{u}^\mathsf{T} G_L = \mathbf{u}^\mathsf{T}[I_K|P] \\ &= [u_1, u_2, \ldots, u_K, x_1, x_2, \ldots, x_I],\end{aligned}$$

where in this case the operations are in the binary field. Although LDGM codes have the advantage of linear encoding complexity, unlike general LDPC codes, they can only attain an arbitrarily low error probability by reducing the rate to zero [20]: Independently of the block length, they suffer from a high error floor. Therefore, they have been historically disregarded in favor of other LDPC codes. However, as explained in the next section, they can actually perform well as an aid to reduce the error floor of other codes.

2.3 Parallel RCM-LDGM code

As shown in Fig. 1, a parallel RCM-LDGM code of rate $K/(M + I)$ bits per real dimension consists of the parallel concatenation of an RCM code of rate K/M with a high rate binary regular LDGM code that produces I non-systematic coded binary symbols from its K input

binary symbols. That is, the encoded symbol sequence is given by

$$\mathbf{x}^{\mathsf{T}} = \left[(G\mathbf{u})^{\mathsf{T}} \mid 2 \cdot \left(\left(\mathbf{u}^{\mathsf{T}} P \mod 2 \right) - \frac{1}{2} \right) \right]$$

where G is the $M \times K$ RCM matrix introduced in Section 2.1 and P is the non-systematic part of the LDGM generator matrix in Section 2.2. Recall that the objective of the LDGM code is to reduce the error floor produced by the RCM code, but without degrading the RCM waterfall region.

As explained before, the RCM symbols and the BPSK modulated LDGM coded bits are grouped two by two and transmitted using a QAM modulator, so that the spectral efficiency, ρ, is

$$\rho = \frac{2 \cdot K}{M + I}$$

bits per complex channel symbol. In the results, we will utilize the spectral efficiency, ρ, instead of the rate (given by $\frac{K}{M+I}$).

2.4 Decoder block

For decoding, the sum-product algorithm (SPA) [21] is applied to the factor graph that models the overall communications system. This factor graph is sketched in Fig. 1. Let $r_{j,k}$ and $q_{k,j}$ denote the passing log-likelihood ratio (LLR) messages from the jth check node (CN) to the kth variable node (VN), and from the kth VN to the jth CN, respectively. In what follows, we denote by $n(U_k) \setminus T_j$ and $n(T_j) \setminus U_k$ the set of CNs connected to VN U_k without considering CN T_j, and the set of VNs connected to CN T_j without considering VN k, respectively. At each iteration t, the sum-product algorithm is implemented by the sequential execution of the following steps:

- *Step 1.* $q_{k,i}^{(t)}$*: Message passing from variable nodes,* $\{U_k\}_{k=1}^{K}$*, to RCM and LDGM check nodes* $\{T_i\}_{i=1}^{M+I}$.

$$q_{k,i}^{(t)} = \sum_{j \in n(U_k) \setminus T_i} r_{j,k}^{(t-1)} + \log\left(\frac{p_1}{p_0}\right), \tag{1}$$

  where $r_{j,k}^{(0)} = 0$ for $k \in \{1, \ldots, K\}, j \in n(U_k) \setminus T_i$ and $(p_1; p_0)$ is the distribution of the memoryless binary source.
- *Step 2.* $r_{i,k}^{(t)}$*: Message passing from RCM-LDGM check nodes,* $\{T_i\}_{i=1}^{M+I}$*, to variable nodes* $\{U_k\}_{k=1}^{K}$.
  - *Computation at RCM check nodes* $\{T_i\}_{i=1}^{M}$*:* Observe that $x_j = \sum_i g_{j,i} u_i$ and define $a_{j,k} = \sum_{i \sim k} g_{j,i} u_i$, where $\sum_{i \sim k}$ means the sum over all i except k. Combining both terms, we get $x_j = a_{j,k} + g_{j,k} u_k$ for all $k \in n(T_j)$, and the received symbol $y_j = x_j + n_j$. The message $r_{i,k}^{(t)}$ is calculated as

$$r_{i,k}^{(t)} = \log\left(\frac{\sum_z P^{(t)}(a_{j,k} = z) \cdot e^{-(y_j - z - g_{j,k})^2 / N_0}}{\sum_z P^{(t)}(a_{j,k} = z) \cdot e^{-(y_j - z)^2 / N_0}} \right) \tag{2}$$

    where the sum in z is over all possible values that the RCM symbols can take. Notice that $P^{(t)}(a_{j,k} = z)$, the probability of $a_{j,k} = z$ at iteration t, is calculated in a straightforward manner by convolving the probability density functions (PDFs) of the terms in the summation, where the distribution functions of these terms are obtained from the received LLR messages $q_{k,i}^{(t)}$. An efficient way to implement these convolutions is explained in [1].
  - *Computation at LDGM check nodes* $\{T_i\}_{i=M+1}^{M+I}$*:* As in standard LDGM codes, the LLR message transmitted from the ith check node to the variable node U_k is given by

$$r_{i,k}^{(t)} = -2\text{atanh}\left(\tanh\left(-\frac{\gamma_i}{2}\right) \prod_{j \in n(T_i) \setminus U_k} \tanh\left(-\frac{q_{k,j}^{(t)}}{2}\right) \right), \tag{3}$$

$$\text{where } \gamma_i = \frac{(y_i + 1)^2 - (y_i - 1)^2}{N_0}.$$

At the end of the iterations, when $t = t_{\max}$, an estimate of u_k can be calculated as

$$\widehat{u_k} = \begin{cases} 1, & \left(\sum_{j \in n(U_k)} r_{j,k}^{(t_{\max})} \right) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

3 Methods: EXIT chart analysis for the LDGM-RCM code

To obtain codes with near Shannon achieving performance, it is crucial to select good code parameters. For instance, when considering the case of a pure RCM code, one has to find a suitable weight set $\mathcal{D}$. In [10, 11], simulations of the entire communications system were performed for a given set of design parameters. This procedure was repeated until a good combination of parameters was found. The drawback of this procedure is that it takes a large amount of computational time. To overcome this problem, the authors in [17] proposed an EXIT chart analysis, shortening in this way the parameter selection procedure for pure RCM codes. As already mentioned, this EXIT chart analysis cannot be directly applied in our scheme, since two different types of check

nodes, RCM and LDGM, have to be considered. This paper extends the analysis to parallel RCM-LDGM codes, considering as well non-uniform sources. Furthermore, it also presents a BER prediction analysis based on EXIT charts that was not previously considered in the literature for this type of codes.

As shown in Fig. 2, the model for EXIT chart analysis is composed of two types of decoders: variable node decoder (VND) composed of all variable nodes, and check node decoder (CND) composed of two different types of check nodes, RCM and LDGM working in the real and binary field, respectively. The LLR values exchanged between the two decoders are modeled as outcomes of real-valued random variables E (output from either a VND or a CND) and A (input to either a VND or a CND).

To characterize the behavior of a node decoder, either check or variable, we obtain the mutual information $I(E;U)$ between the decoder's *LLR* extrinsic output E and a binary source symbol U with distribution $(p_1;p_0)$, as a function of the mutual information $I(A;U)$ between the decoder's LLR a priori input A and U. $I(E;U)$ and $I(A;U)$ are given by

$$I(L;U)=p_0\int_{-\infty}^{\infty} f_L(\xi|0)\log_2\left(\frac{f_L(\xi|0)}{p_0 f_L(\xi|0)+p_1 f_L(\xi|1))}\right)d\xi + p_1\int_{-\infty}^{\infty} f_L(\xi|1)\log_2\left(\frac{f_L(\xi|1)}{p_0 f_L(\xi|0)+p_1 f_L(\xi|1))}\right)d\xi, \tag{5}$$

where $L\in\{A,E\}$ and $f_L(\xi|u)$, for $u=0,1$, is the conditional probability density function of L given U. As indicated before, $f_L(\xi|u)$ depends on the node decoder under consideration, that is, whether such a node is a VND or a CND, and it is calculated as indicated in Sections 3.1 and 3.2. In the sequel, we will denote $I(L;U)$ for a VND or a CND as $\mathrm{I}_{\mathrm{L,VND}}=I\left(L^{(vn)};U\right)$ or $\mathrm{I}_{\mathrm{L,CND}}=I\left(L^{(cn)};U\right)$, respectively.

In the course of deriving the EXIT chart for the parallel RCM-LDGM code, we will need the parameters

$$p_{\mathrm{RCM}}^{(vn)}\triangleq\frac{M\cdot\bar{d}_{\mathrm{RCM}}^{(v)}}{M\cdot\bar{d}_{\mathrm{RCM}}^{(v)}+I\cdot d_{\mathrm{LDGM}}^{(v)}} \tag{6}$$

$$p_{\mathrm{RCM}}^{(cn)}\triangleq\frac{d_{\mathrm{RCM}}^{(c)}}{d_{\mathrm{RCM}}^{(c)}+d_{\mathrm{LDGM}}^{(c)}}, \tag{7}$$

which denote the average percentage of edge connections arriving to a VN from an RCM check node and the percentage of edge connections arriving to an RCM check node from a VN, respectively.

3.1 VND EXIT curve for RCM-LDGM codes

The EXIT curve of the VND is given by the transfer characteristic between $\mathrm{I}_{\mathrm{E,VND}}=I(E^{(vn)};U)$ and $\mathrm{I}_{\mathrm{A,VND}}=I(A^{(vn)};U)$. Note that the realizations of RVs $E^{(vn)}$ and $A^{(vn)}$ are the messages exchanged in the sum-product algorithm, $\{r_{i,k}\}$ and $\{q_{k,i}\}$, respectively. In order to evaluate these mutual informations from (5), the conditional PDF of the a priori $A^{(vn)}$ and the extrinsic $E^{(vn)}$ at a variable node decoder, given U, have to be found.

3.1.1 *Calculation of* $\mathrm{I}_{\mathrm{A,VND}}$

Different from previous work on EXIT charts, in an RCM-LDGM code, one has to consider two types of a priori messages arriving at a VND: first, the messages arriving from an edge connected to an RCM check node, $A_{\mathrm{RCM}}^{(vn)}$, and second, the messages arriving from an edge connected to an LDGM check node, $A_{\mathrm{LDGM}}^{(vn)}$.

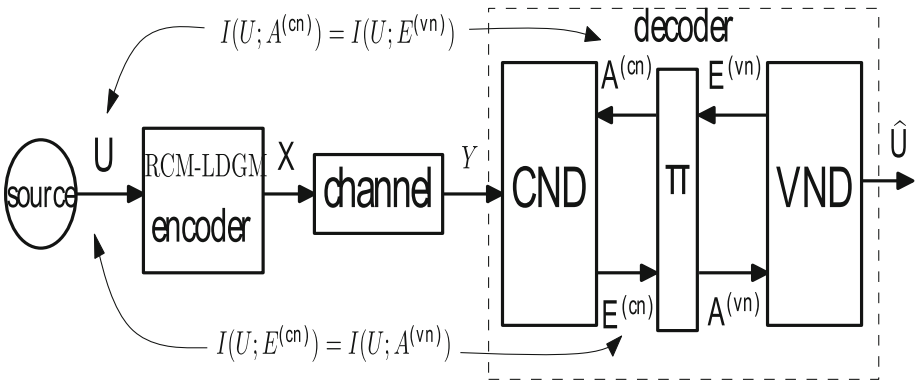

Fig. 2 Model for EXIT chart analysis. Random variable $E^{(vn)}$ models the extrinsic information produced at the VND when $A^{(vn)}$ is the a priori information arriving at the VND. Similarly, $E^{(cn)}$ is the extrinsic information produced at the CND when the channel output and $A^{(cn)}$ are the a priori information arriving at the CND. The block Π denotes the random connections of CN and VNs. The factor graph of the VND and CND is given in Fig. 1

In order to simplify calculations, the authors in [17] modeled the conditional PDF of the $A_{\text{RCM}}^{(vn)}$ message as the PDF of the LLR random variable obtained at the output of a virtual AWGN channel when its inputs are uniform[3] binary source symbols U, i.e.,

$$Y = U + N, \quad N \sim \mathcal{N}\left(0, \sigma^2\right). \tag{8}$$

Under this model, the LLR of the a priori message $A_{\text{RCM}}^{(vn)}$ at a variable node can be expressed as

$$A_{\text{RCM}}^{(vn)} = \log\left(\frac{P(u=1|Y)}{P(u=0|Y)}\right) = \frac{2U-1}{2\sigma^2} + \frac{N}{\sigma^2}.$$

Using the same assumptions for $A_{\text{LDGM}}^{(vn)}$, $A_{\text{RCM}}^{(vn)}$ and $A_{\text{LDGM}}^{(vn)}$ are defined as

$$A_{\text{RCM}}^{(vn)}|U \sim \mathcal{N}\left((2u-1)\sigma_{R,A}^2/2, \sigma_{R,A}^2\right) \tag{9}$$

$$A_{\text{LDGM}}^{(vn)}|U \sim \mathcal{N}\left((2u-1)\sigma_{L,A}^2/2, \sigma_{L,A}^2\right), \tag{10}$$

where $\sigma_{R,A}^2$ and $\sigma_{L,A}^2$ represent the inverse of the variance of the two different virtual channels.

The main challenge of having two different types of CNs rather than one, as in the case of a standard EXIT chart, is that the mutual information $\text{I}_{\text{A,VND}}$ will now depend on two variables, $\sigma_{R,A}^2$ and $\sigma_{L,A}^2$, rather than just on one. Notice, however, that although $A_{\text{RCM}}^{(vn)}|U$ and $A_{\text{LDGM}}^{(vn)}|U$ can be considered independent, their variances are coupled due to the way the SPA generates the messages (refer to Section 2.4). Therefore, if one of the variances can be expressed as a function of the other, then $\text{I}_{\text{A,VND}}$ becomes a function of only one variable, which simplifies the analysis.

Calculating the coupling between $\sigma_{R,A}^2$ and $\sigma_{L,A}^2$, as a function $\sigma_{R,A}^2 = f(\sigma_{L,A}^2, \text{SNR})$, for the range of SNR of interest (i.e., the SNRs belonging to the waterfall region of the code), is computationally expensive, which is counter to the objective of EXIT chart analysis. Fortunately, simulation results have shown that this dependency can be approximated linearly. This can be seen clearly in Fig. 3, where A, simulated in a fast Rayleigh channel, and B, simulated in a AWGN channel, denote the codes of Tables 3 and 1 for $p_0 = 0.8$ at their corresponding SNR waterfall ranges, 21–22 and 18–19 dB. Therefore, in what follows, we will assume that $\sigma_{R,A}^2 = f\left(\sigma_{L,A}^2\right)$ can be approximated in the range of SNR of interest by $\sigma_{L,A}^2 = \frac{\sigma^2{}_{R,A}}{\kappa}$. This yields

$$A_{\text{RCM}}^{(vn)}|U \sim \mathcal{N}\left((2u-1)\frac{\sigma_{R,A}^2}{2}, \sigma_{R,A}^2\right) \tag{11}$$

$$A_{\text{LDGM}}^{(vn)}|U \sim \mathcal{N}\left((2u-1)\frac{\sigma_{R,A}^2}{2\cdot\kappa}, \frac{\sigma_{R,A}^2}{\kappa}\right). \tag{12}$$

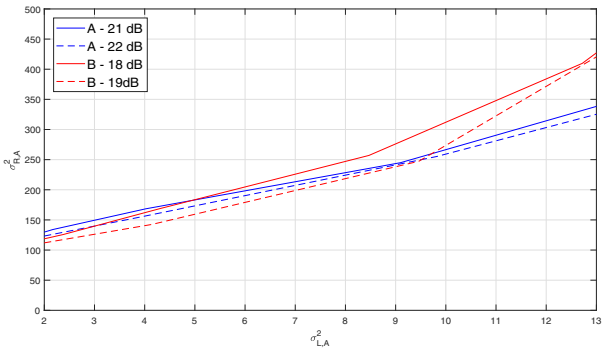

Fig. 3 $\sigma_{R,A}^2$ versus $\sigma_{L,A}^2$ curves obtained by Monte Carlo simulations for two different RCM-LDGM codes. A (fast fading Rayleigh channel) and B (AWGN channel) denote the codes of Tables 3 and 1 for $p_0 = 0.8$, respectively

The constant κ scales the variance of the distribution of $A_{\text{LDGM}}^{(vn)}$ with respect to the variance of $A_{\text{RCM}}^{(vn)}$. The steps to compute it are explained in the Appendix.

Since we have two types of a priori messages, the corresponding conditional PDF of $A^{(vn)}|U$, is obtained as

$$\begin{aligned} A^{(vn)}|U \sim f_{A^{(vn)}}(a|u) = & f_{A_{\text{RCM}}^{(vn)}}(a|u)p_{\text{RCM}}^{(vn)} \\ & + f_{A_{\text{LDGM}}^{(vn)}}(a|u)\left(1 - p_{\text{RCM}}^{(vn)}\right), \end{aligned} \tag{13}$$

where the PDFs $f_{A_{\text{RCM}}^{(vn)}}$ and $f_{A_{\text{LDGM}}^{(vn)}}$ are given in (11) and (12), respectively, and $p_{\text{RCM}}^{(vn)}$ in (6). Finally, applying (5), $\text{I}_{\text{A,VND}}$ is calculated from $f_{A_{\text{RCM}}^{(vn)}}$ as a parametric expression of $\sigma_{R,A}^2$.

3.1.2 *Calculation of* $\text{I}_{\text{E,VND}}$

Once $f_{A^{(vn)}}(a|u)$ is calculated, the conditional distribution of the extrinsic RV $E^{(vn)}$ at the variable node decoder can be computed. To this end, note that there are two types of LLR messages: messages passed on an edge connecting a VN to an RCM check node (modeled as a RV denoted as $E_{\text{RCM}}^{(vn)}$), and messages passed on an edge connecting a VN to an LDGM check node (modeled as RV $E_{\text{LDGM}}^{(vn)}$). From the corresponding connections of the factor graph, we obtain

$$E_{\text{RCM}}^{(vn)} = \underbrace{A_{\text{RCM}}^{(vn)} + \ldots + A_{\text{RCM}}^{(vn)}}_{(\bar{d}_{\text{RCM}}^{(v)}-1)\text{i.i.d RVs}} + \underbrace{A_{\text{LDGM}}^{(vn)} + \ldots + A_{\text{LDGM}}^{(vn)}}_{d_{\text{LDGM}}^{(v)}\text{i.i.d RVs}} + \log\left(\frac{p_1}{p_0}\right) \tag{14}$$

$$E_{\text{LDGM}}^{(vn)} = \underbrace{A_{\text{RCM}}^{(vn)} + \ldots + A_{\text{RCM}}^{(vn)}}_{\overline{d}_{\text{RCM}}^{(v)}\text{i.i.d RVs}} + \underbrace{A_{\text{LDGM}}^{(vn)} + \ldots + A_{\text{LDGM}}^{(vn)}}_{(d_{\text{LDGM}}^{(v)}-1)\text{i.i.d RVs}} + \log\left(\frac{p_1}{p_0}\right) \tag{15}$$

so that the corresponding conditional PDFs of the extrinsic LLR messages, $E_{\text{RCM}}^{(vn)}$ and $E_{\text{LDGM}}^{(vn)}$, are

$$f_{\text{RCM}}(e|u) = \mathcal{N}\left((2u-1)\frac{\sigma_{R,E}^2}{2} + \log\left(\frac{p_1}{p_0}\right), \sigma_{R,E}^2\right)$$
$$f_{\text{LDGM}}(e|u) = \mathcal{N}\left((2u-1)\frac{\sigma_{L,E}^2}{2} + \log\left(\frac{p_1}{p_0}\right), \sigma_{L,E}^2\right),$$

respectively, where

$$\sigma_{R,E}^2 = \sigma_{R,A}^2\left(\overline{d}_{\text{RCM}}^{(v)} - 1 + \frac{d_{\text{LDGM}}^{(v)}}{\kappa}\right),$$
$$\sigma_{L,E}^2 = \sigma_{R,A}^2\left(\overline{d}_{\text{RCM}}^{(v)} + \frac{d_{\text{LDGM}}^{(v)} - 1}{\kappa}\right).$$

Again, since we have two types of extrinsic messages, the overall conditional PDF of the extrinsic LLR random variable $E^{(vn)}|U$ is obtained as

$$E^{(vn)}|U \sim f_{E^{(vn)}}(e|u) = f_{\text{RCM}}(e|u)p_{\text{RCM}}^{(cn)} + f_{\text{LDGM}}(e|u)\left(1 - p_{\text{RCM}}^{(cn)}\right) \tag{16}$$

where $p_{\text{RCM}}^{(cn)}$ is given in (7). Finally, applying (5), $\text{I}_{\text{E,VND}}$ is calculated from $f_{E_{\text{RCM}}^{(vn)}}$ as a parametric expression of $\sigma_{R,A}^2$.

3.2 CND EXIT curve for the RCM-LDGM codes

From the fact that the a priori information $A^{(cn)}$ at the check node decoder is equal to the extrinsic information $E^{(vn)}$ at the variable node decoder (refer to Fig. 2), the PDF $f_A^{(cn)}(a|u)$ of $A^{(cn)}$ is given by the PDF in (16), interchanging E with A

$$A^{(cn)}|U \sim f_{A^{(cn)}}(a|u) = f_{\text{RCM}}(a|u)p_{\text{RCM}}^{(cn)} + f_{\text{LDGM}}(a|u)(1 - p_{\text{RCM}}^{(cn)}), \tag{17}$$

so that $I_{\text{A,CND}} = I_{\text{E,VND}}$.

To compute $I_{\text{E,CND}}$, we need to find the conditional PDF $f_{E^{(cn)}}(e|u)$ of the extrinsic LLR $E^{(cn)}$ at the CND. This is done by running step 2 of the sum-product algorithm (see Section 2.4) and setting $q_{k,i} = a$, where a are realizations of a random variable $A^{(cn)}$ with conditional PDF (17). The empirical conditional PDF $f_{E^{(cn)}}(e|u)$ is now found by the histogram of the realizations $\{r_{i,k}\}$.

3.3 Trajectories of iterative decoding and decoding threshold

To account for the iterative nature of the decoding process, both the VND and CND transfer characteristics should be plotted into a single diagram. As long as the SNR is large enough so that both transfer curves do not intersect, the iterative process will achieve its maximum mutual information values, $(H(p_0), H(p_0))$, consequently achieving a low BER. The smallest SNR value for which both curves do not intersect is defined as the decoding threshold and represents the minimum SNR required to decode without errors an infinite length code with the given configuration. Therefore, the code design problem reduces to find a code configuration, i.e., weight sets $\mathcal{D}$, and parameters I and $d_{\text{LDGM}}^{(v)}$, such that the decoding threshold is as close as possible to the corresponding SNR Shannon limit.

Remark Note that the VND EXIT curve only depends on the values of $\overline{d}_{\text{RCM}}^{(v)}$ and $d_{\text{LDGM}}^{(v)}$. On the other hand, the CND EXIT curve depends on all the parameters, i.e., $\{\mathcal{D}, SNR, \overline{d}_{\text{RCM}}^{(v)}, d_{\text{LDGM}}^{(v)}, d_{\text{LDGM}}^{(c)}, M, I\}$. □

Remark The EXIT chart for a pure RCM code can be calculated as a particular case of the parallel LDGM-RCM by taking $p_{\text{RCM}}^{(vn)} = p_{\text{RCM}}^{(cn)} = 1$. □

3.4 Predicting the BER from the EXIT chart

For those SNR values smaller than the decoding threshold, the EXIT chart can be used to obtain an estimate on the BER after an arbitrary number of iterations. Following the sum-product algorithm, the LLR value of the decision variable, s_k, of variable node k at the end of a number of iterations, is obtained as the sum of all LLR messages $r_{i,k}$ that were passed over a single edge connecting a CN, i, with the corresponding VN, k, i.e., $s_k = \sum_i r_{i,k} + \log\left(\frac{p_1}{p_0}\right)$ for $i \in n(U_k)$. From the previous assumptions, $r_{i,k}$ can be considered to be a realization of the independent Gaussian random variables $A_{\text{RCM}}^{(vn)}$ and $A_{\text{LDGM}}^{(vn)}$. The conditional PDF of s_k given U is

$$S|U \sim \mathcal{N}\left(\mu_S(u), \sigma_S^2\right),$$

with $\sigma_S^2 = \sigma_{R,A}^2 \cdot \left(\overline{d}_{\text{RCM}}^{(v)} \cdot + \frac{d_{\text{LDGM}}^{(v)}}{\kappa}\right)$, and $\mu_S(u) = (2u-1)\frac{\sigma_S^2}{2} + \log\left(\frac{p_1}{p_0}\right)$. The BER performance is now obtained as

$$P_b = p_0 P(S > 0|U = 0) + p_1 P(S \leq 0|U = 1). \tag{18}$$

Observe that

$$P(S > 0|U = 0) = 1 - Q\left(\frac{\mu_S(0)}{\sigma_S}\right) \quad \text{and,}$$

$$P(S \leq 0|U = 1) = Q\left(\frac{\mu_S(1)}{\sigma_S}\right),$$

where $Q(\xi)$ is the Q function

$$Q(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\xi}^{\infty} \mathrm{e}^{-\frac{y^2}{2}} dy.$$

Remark The BER for a pure RCM code can be estimated as a particular case of the parallel LDGM-RCM with $d_{\text{LDGM}}^{(v)} = 0$. □

4 Results and discussion

In this section, we evaluate the proposed EXIT chart analysis and BER prediction method of Section 3 for both AWGN and fast fading Rayleigh channels. We begin by considering the AWGN channel. Section 4.1 presents some mutual information trajectories of actual codes on the corresponding EXIT charts. In Section 4.2, we compare the BER predictions obtained using these charts with the BER obtained by Monte Carlo simulations. In Section 4.3, the EXIT analysis is used to obtain codes that approach the Shannon theoretical limit. Finally, the extension to Rayleigh channels is considered in Section 4.4.

4.1 Trajectories

We begin in Fig. 4 by showing the EXIT chart of a pure RCM code with weight set $\mathcal{D} = \{2, 3, 4, 8\}$ and spectral efficiency $\rho = 7.4$ for three different SNR values, 17, 18, and 20.25 dB, and for a non-uniform source with entropy $H(p_0) = 0.72$ ($p_0 = 0.8$). Notice that the variable node curve (which is valid for all SNRs) ends at the point $(H(p_0), H(p_0))$ as it should be. Also plotted in the figure are the Monte Carlo simulated mutual information trajectories of this code with block length $K = 37000$ (and $M = 10000$). Each trajectory is plotted using the same color as their corresponding SNR's EXIT chart CN curve, and they end where the corresponding CN and VN curves intersect. In addition, the contour lines of the corresponding simulated BERs are also shown. For example, at SNR= 17 dB, the BER of the code is $5.5 \cdot 10^{-2}$ and, as observed, the blue curve of the EXIT chart intersects the VN decoder curve very close to the $5.5 \cdot 10^{-2}$ contour line. Similarly, for SNR= 18 dB and 20.25 dB, the simulated BERs are $3.3 \cdot 10^{-2}$ and $2.2 \cdot 10^{-3}$, respectively. Again, the intersections between CN and VN curves occur very close to the corresponding BER's contour lines. Note however that none of these SNRs allow the channel to be open.

As previously explained , RCM-LDGM codes substitute some RCM symbols with non-systematic LDGM QPSK modulated bits, with the goal to lower the error floor of the corresponding RCM code. Figure 5 shows the EXIT chart and mutual information trajectories for the RCM-LDGM code obtained by substituting 200 RCM symbols by 200 LDGM coded binary symbols (with $d_{\text{LDGM}}^{(v)} = 1$) in the previous RCM configuration.

Observe that by introducing these 200 LDGM coded bits, we avoid the previous intersection of the curves at SNR 20.25 dB, improving in this way the BER at 20.25 dB. The corresponding mutual information trajectory at

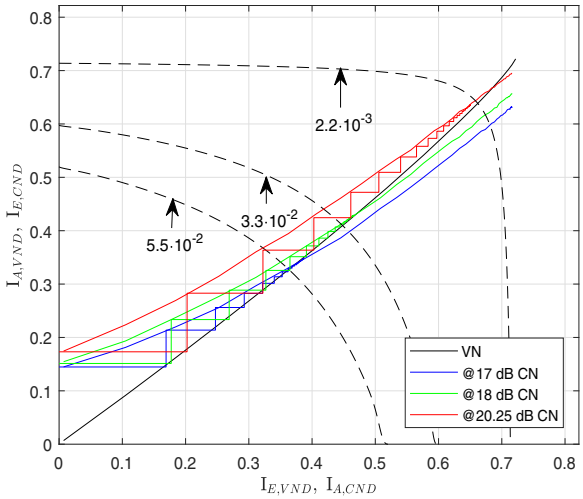

Fig. 4 EXIT chart, BER contour lines, and mutual information trajectory for a pure RCM code of $\rho = 7.4$ when transmitting a non-uniform source with entropy $H(p_0) = 0.72$ over an AWGN channel

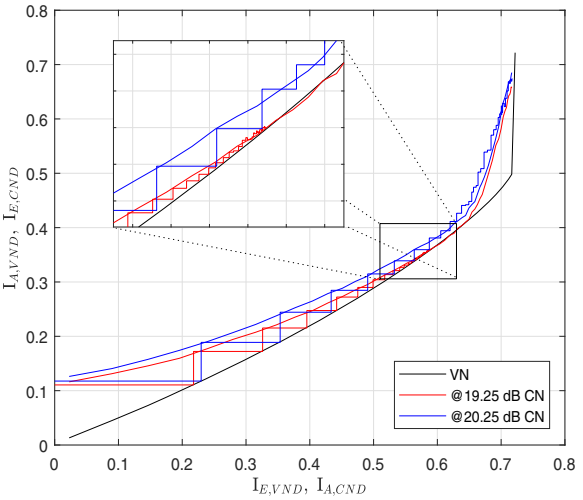

Fig. 5 EXIT chart and mutual information trajectory of a RCM-LDGM code when transmitting a non-uniform source with entropy $H(p_0) = 0.72$ over an AWGN channel

SNR 20.25 is shown in Fig. 5. Since the channel is open, it reaches its maximal value, i.e., (0.72,0.72). It turns out that SNR = 20.25 is the smallest SNR that allows the channel to remain open, and as such, it is the corresponding decoding threshold of the given code. In the same figure, the trajectory at SNR = 19.25 dB is also shown, but in this case the channel is closed and does not reach the maximum value.

4.2 Bit error rate from the EXIT charts

As explained in Section 3.4, an estimated BER can be assigned to each point of the variable node (VN) curve of the EXIT chart. Therefore, the BER of a particular code at a given SNR is obtained from the value of the VN point where the CN and VN curves intersect.

In this section, we will consider two different RCM-LDGM configurations with $\rho = 4$ given by $K = 25000$, $M = 12365$, and $I = 135$, with $d^{(v)}_{\text{LDGM}}$ 1 and 2. Moreover, we will consider three different sources with $p_0 = 0.5$, $p_0 = 0.8$, and $p_0 = 0.95$ and three different weight sets $\mathcal{D} = \{1,1,1,1,2,2\}, \{1,1,2,2,4,4\}$, and $\{2,2,3,3,4,4\}$. Recall that the VN curve of the EXIT chart depends on M, I, $d^{(v)}_{\text{LDGM}}$, and $d^{(c)}_{\text{RCM}}$, whereas the CND curve depends also on the actual values of $\mathcal{D}$ and on the SNR.

Figure 6a shows the EXIT chart of the configuration with $d^{(v)}_{\text{LDGM}} = 1$ for the different weight sets and two SNR values, 10 and 12 dB. The plot shows the BER estimated values at the intersecting points. For example, for the configuration described above with $\{1,1,1,1,2,2\}$, the estimated BERs are $1.5 \cdot 10^{-2}$ at SNR= 10 dB and $5 \cdot 10^{-4}$ at SNR= 12 dB. When $\{1,1,2,2,4,4\}$, the estimated BERs are $2.2 \cdot 10^{-2}$ at SNR= 10 dB and $3 \cdot 10^{-3}$ at SNR= 12 dB. Finally, for $\{2,2,3,3,4,4\}$, we obtain $9.5 \cdot 10^{-3}$ and $4 \cdot 10^{-4}$, respectively. Similarly, Fig. 6b shows the EXIT curves at SNR= 4 and 6 dB for the configuration with $d^{(v)}_{\text{LDGM}} = 2$ with their estimated BER values when transmitting a source with $p_0 = 0.95$. Since $H(0.95) = 0.28$, the VN curve ends at point (0.28,0.28).

In order to corroborate our BER predictions, Fig. 7 compares the BER curves obtained by Monte Carlo simulation with those obtained by the EXIT chart BER estimation, as it is done in Fig. 6. As it can be seen in the figure, the predictions are accurate for both uniform and non-uniform sources.

The parameters of these codes have not been optimized, and therefore, they present a large gap to the corresponding Shannon limits given by $10 \cdot \log_{10}\left(2^{\rho \cdot H(S)} - 1\right)$, which correspond to 0.91, 8.06, and 11.76 dB for $p_0 = 0.95$, $p_0 = 0.8$, and $p_0 = 0.5$, respectively. In the next section, we will obtain near capacity high spectral efficiency codes using the EXIT chart analysis.

4.3 Code design based on the decoding threshold for AWGN channels

The idea behind the design method is to start with a pool of possible codes having the required rate and then obtain the EXIT charts for the source of interest. The codes with the lower decoding threshold or those whose curves

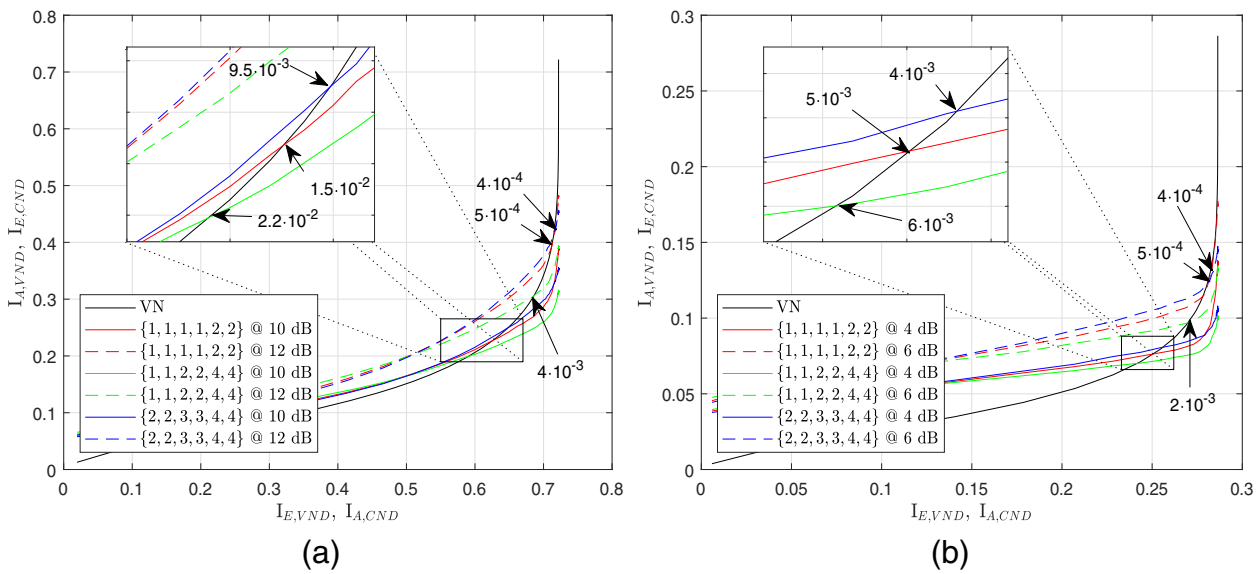

Fig. 6 EXIT chart and predicted BER at the crossing points for different SNR values. **a** For a non-uniform source with $p_0 = 0.8$ and the configuration with $d^{(v)}_{\text{LDGM}} = 1$. **b** For a non-uniform source with $p_0 = 0.95$ and the configuration with $d^{(v)}_{\text{LDGM}} = 2$. An AWGN channel is considered

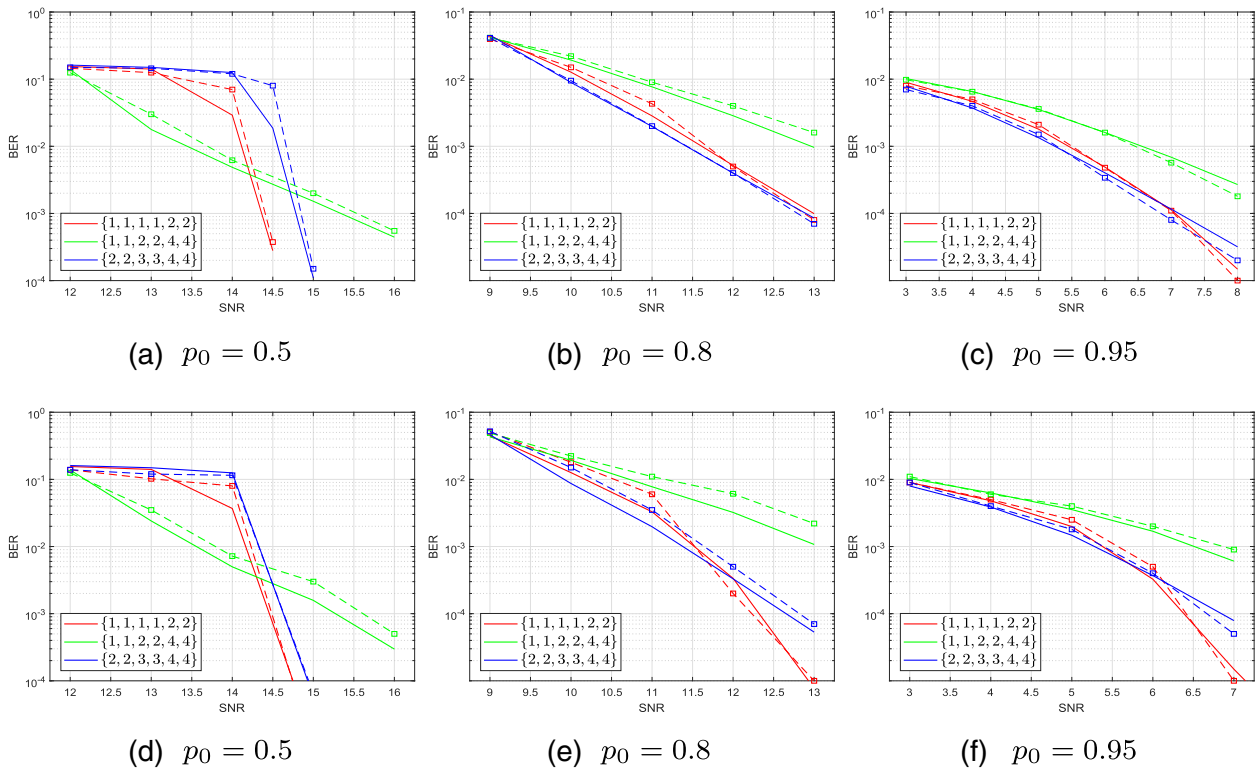

Fig. 7 Predicted (dashed line) and Monte Carlo simulated (continuous line) BER vs SNR curves for the RCM-LDGM codes. **a**–**c** The configuration with $d^{(v)}_{\text{LDGM}} = 1$. **d**–**f** The configuration with $d^{(v)}_{\text{LDGM}} = 2$. An AWGN channel is considered

intersect closer to the maximum point $(H(p_0), H(p_0)$ are kept. The resulting subset of codes are then tuned-up by slightly changing their designed parameters. We have observed the following trends:

1 For sources with smaller entropy, larger RCM weight sets, D, tend to work better, since the sum-product algorithm is aided by the a priori probability.
2 When designing the LDGM part of the code, there is a trade-off regarding the number I of LDGM binary symbols. By increasing I, more residual errors are corrected in the waterfall region, making it steeper. However, larger SNRs are required to reach this waterfall region.
3 The range for parameter $d^{(v)}_{\text{LDGM}}$ is between 1 and 5. The larger parameter I is, the larger value for $d^{(v)}_{\text{LDGM}}$ can be selected.

Next, we provide EXIT chart design examples for an RCM-LDGM code with a spectral efficiency $\rho = 7.4$ for the transmission over AWGN channels of three memoryless sources with a priori probabilities $p_0 = 0.5$, $p_0 = 0.8$, and $p_0 = 0.95$. The corresponding Shannon limits are at 22.25 dB, 15.97 dB, and 5.24 dB, respectively. Table 1 shows the best codes obtained by the EXIT chart analysis for a code length K = 37000 bits.

The corresponding EXIT charts and real trajectories of the designed codes are plotted in Fig. 8. Note from Fig. 8c that when transmitting the uniform source ($p_0 = 0.5$), the channel between both EXIT curves remains open at SNR = 24.15 dB (1.9 dB from the Shannon limit) and, consequently, the received blocks should be decoded with a low probability of error. However, if this same code was used to transmit the symbols generated by the non-uniform source ($p_0 = 0.8$) by only modifying the a

Table 1 Best configurations obtained by the EXIT chart analysis for AWGN channels

p_0	K	M	I	$d^{(v)}_{\text{LDGM}}$	D	Decoding threshold (dB)
0.5	37000	9800	200	1	{2, 3, 4, 8}	24.15
0.8	37000	9720	280	4	{2, 2, 3, 3, 4, 8}	18.2
0.95	37000	9200	800	3	{1, 1, 1, 1, 1, 1, 1, 1, 1, 1}	7.25

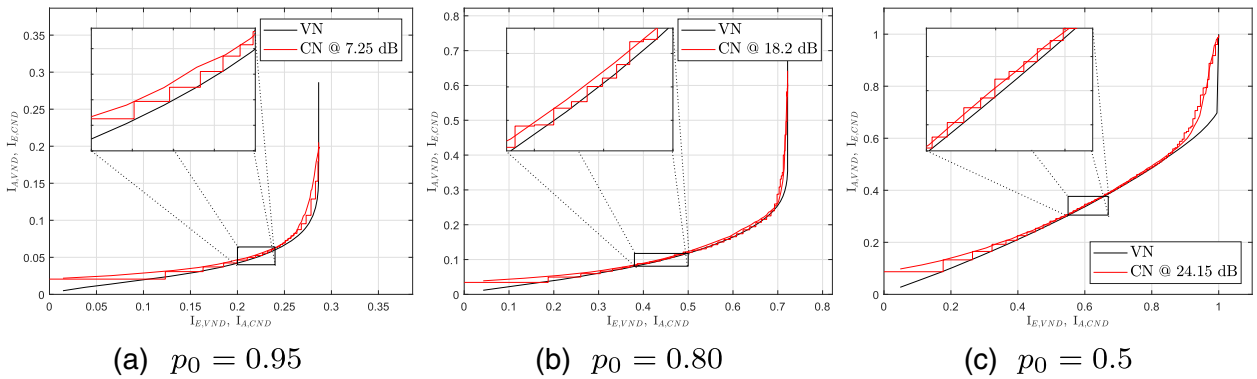

Fig. 8 EXIT chart and real trajectories of the designed codes for AWGN channels and sources with **c** $p_0 = 0.5$, **b** $p_0 = 0.8$, and **a** $p_0 = 0.95$

priori probability to 0.8 in the SP algorithm, the decoding threshold would decrease to 20.15 dB (refer Fig. 5). This is still 4 dB away from the Shannon limit. The optimized code for this source is given in Table 1. Notice that the gap is reduced to 2.23 dB (refer to Fig. 8b). This clearly shows that when transmitting binary symbols generated by a non-uniform source, the channel code behaves like a joint source-channel code and, therefore, it has to be designed according to the source. Finally, Fig. 8a plots the EXIT chart of the best found configuration when transmitting the non-uniform source with $p_0 = 0.95$. The decoding threshold is only 2.01 dB away from the Shannon limit.

In order to corroborate that the codes obtained from the EXIT chart analysis perform as expected, Fig. 9 plots the BER vs SNR curves obtained by Monte Carlo simulations, as well as the theoretical decoding thresholds for the designed codes (shown as vertical dashed lines), and the corresponding Shannon limits (vertical black lines). Note from the Monte Carlo simulations that the code designed for $p_0 = 0.5$ is 1.9 dB away from its Shannon limit for a BER $= 10^{-5}$, while the codes optimized for the sources with $p_0 = 0.8$ and $p_0 = 0.95$ present both a gap of 2.3 dB with respect to the Shannon limits. The figure indicates that the decoding threshold obtained from the EXIT chart analysis very accurately predicts the waterfall region. For the source with $p_0 = 0.5$, the gap between the decoding threshold and the waterfall region at BER = 10^{-5} is not appreciable, whereas for the sources with $p_0 = 0.8$ and $p_0 = 0.95$, these gaps are 0.1 dB and 0.3 dB, respectively. An explanation for the gap increase is that for non-uniform sources longer blocks are required to maintain stationarity.

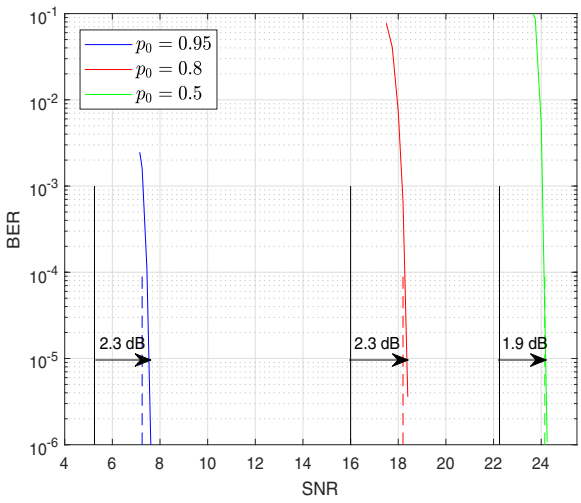

Fig. 9 BER vs SNR behavior of the obtained codes for $p_0 = 0.5$, $p_0 = 0.8$, and $p_0 = 0.95$. The Shannon limits (continuous lines) and the predicted decoding threshold (dashed lines) are plotted as vertical lines. An AWGN channel is considered

Table 2 summarizes the simulation time, as well as the computational time of the EXIT chart analysis, required to obtain the BER vs SNR points of the code $K = 37000$, $M = 9800$, $I = 200$, $d^{(v)}_{\text{LDGM}} = 1$, and $D = \{2, 3, 4, 8\}$. As shown in the table, the EXIT chart analysis is much faster

Table 2 Computational time required to predict a BER vs SNR point in Fig. 9

BER	EXIT chart	Monte Carlo simulation		Reduction factor
		Average convergence time per block	Blocks for 10 errors	
10^{-3}	10s	229s	1	22.9
10^{-4}	10s	113s	3	33.9
10^{-5}	10s	96s	27	259.2
10^{-6}	10s	85s	270	2295

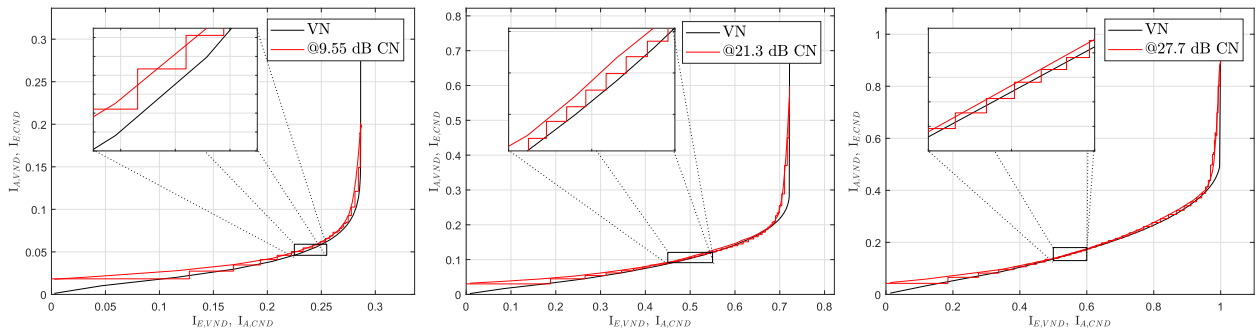

Fig. 10 EXIT chart and real trajectories of the designed codes for fast Rayleigh channel and sources with **c** $p_0 = 0.5$, **b** $p_0 = 0.8$, and **a** $p_0 = 0.95$

than the simulations, making the search by trial and error feasible.

4.4 Extension to fast fading Rayleigh channels

We now look at the behavior of the EXIT chart analysis when considering fast fading Rayleigh channels. Note that the only modification that has to be introduced in this case is in step 2 of the SP algorithm (see Section 3.2). Specifically, since we are assuming perfect channel state information (CSI) at the receiver, the fading factor that multiplies the coded RCM-LDGM symbols (i.e., realizations of i.i.d. exponential random variables) has to be provided to the decoder.

As in the previous AWGN case, we focus on the EXIT chart design for codes of spectral efficiency $\rho = 7.4$ bits per complex dimension and with sources having a priori probabilities $p_0 = 0.5$, $p_0 = 0.8$, and $p_0 = 0.95$. The corresponding SNR Shannon limits are 24.7 dB, 18.3 dB, and 6.8 dB, respectively. Figure 10 is similar to Fig. 8, except that we now consider fast fading. It plots the EXIT charts and real trajectories of the good codes, specified in Table 3, which have been selected by our EXIT chart analysis. The EXIT chart channels are open at SNRs close to the Shannon limits. This is shown in Fig. 11, where the BER predictions and the actual Monte Carlo simulations are presented for different values of SNR. Note that the gaps to the Shannon limits are within 3 dB and that the BER predictions are very close to the results obtained by simulations, corroborating that the proposed EXIT chart analysis is also well suited for fast fading Rayleigh channels.

Table 3 Best configurations obtained by the EXIT chart analysis for fast fading Rayleigh channels

p_0	K	M	I	$d^{(v)}_{\text{LDGM}}$	D	Decoding threshold (dB)
0.5	37000	9600	300	3	{2, 3, 4, 4, 8}	27.7
0.8	37000	9240	760	5	{2, 2, 3, 3, 4, 8}	21.3
0.95	37000	9480	520	4	{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1}	9.55

5 Conclusion

Parallel RCM-LDGM codes are very well suited for implementing smooth high rate adaptation when transmitting uniform and non-uniform binary memoryless sources. However, when long block lengths are considered, the search of good design parameters using a brute force approach is time consuming. To speed up the design process, we have successfully developed an EXIT chart analysis for these codes, which presents the challenge of the combination of analog and digital check nodes,

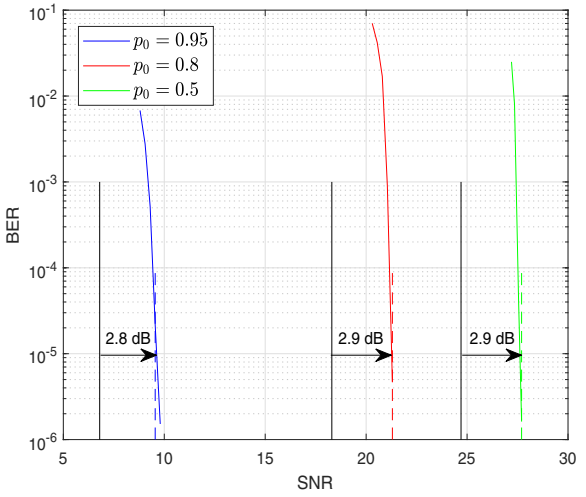

Fig. 11 BER vs SNR obtained by Monte Carlo simulations for fast fading Rayleigh channels. The Shannon limits (continuous lines) and the predicted decoding threshold (dashed lines) are plotted as vertical lines

something not encountered in other works. By assuming a linear relationship between the variances of the LLR messages in both types of CNs, very precise EXIT charts are obtained for the case of AWGN and fast fading Rayleigh channels. The predicted BER vs SNR curves are very close to the results obtained through simulations.

Endnotes

[1] Without loss of generality, it is assumed that K is an even integer.

[2] A multiset is a generalization of the concept of a set that, unlike a set, allows multiple instances of the multiset's elements.

[3] The a priori probability of the source symbols is already considered in step 1 of the SPA.

Appendix

Obtaining κ

The constant κ is computed by Monte Carlo simulation through the following iterative procedure:

1. Start with an initial value of κ in (12) (say $\kappa = 1$), and choose a value for $\sigma^2_{R,A}$ so that the corresponding value of the mutual information computed by the PDF in (11) is in the range (0.5,0.9). For the value of $\sigma^2_{R,A}$ under consideration, generate the extrinsic messages passed from the VN to the RCM and LDGM check nodes according to (14) and (15), respectively.
2. Run one iteration of the sum-product algorithm to obtain the extrinsic LLR messages passed from each LDGM and RCM check nodes to the VN, and obtain their empirical conditional PDFs.
3. Define κ_1 as the ratio between the variances of the empirical conditional distributions of RCM and LDGM check nodes obtained in step 2.
4. Repeat the previous 3 steps, using κ_1 as the initial value for κ, until the value of κ_1 in step 3 is close enough to the value of κ_1 in the previous iteration.
5. Set $\kappa = \kappa_1$ in the distribution (12).

Figure 12 shows a graphical example of the steps followed to find κ. The initial empirical conditional PDFs (i.e., when $\kappa = 1$) are shown in Fig. 12a. As it can be observed, none of the LLR messages is appropriately modeled at this point, since the initial value for κ ($\kappa = 1$) was chosen arbitrarily. Note that since $\kappa = 1$, the modeled $A^{(VN)}_{\text{RCM}}$ is equal to $A^{(VN)}_{\text{LDGM}}$. The value of κ obtained in step 3 is 43, and the corresponding empirical conditional distributions are shown in the Fig. 12b. Notice that for this value of κ, the messages are better modeled by (11) and (12). However, the process is not finished yet. The second iteration results in $\kappa = 26$. The corresponding empirical conditional distributions are shown in the Fig. 12c. If we perform an additional iteration, it will result in a κ close to 26, which will indicate that the procedure has concluded. We have found that in all simulated cases the number of iterations required for $\kappa_0 \approx \kappa_1$ is around three.

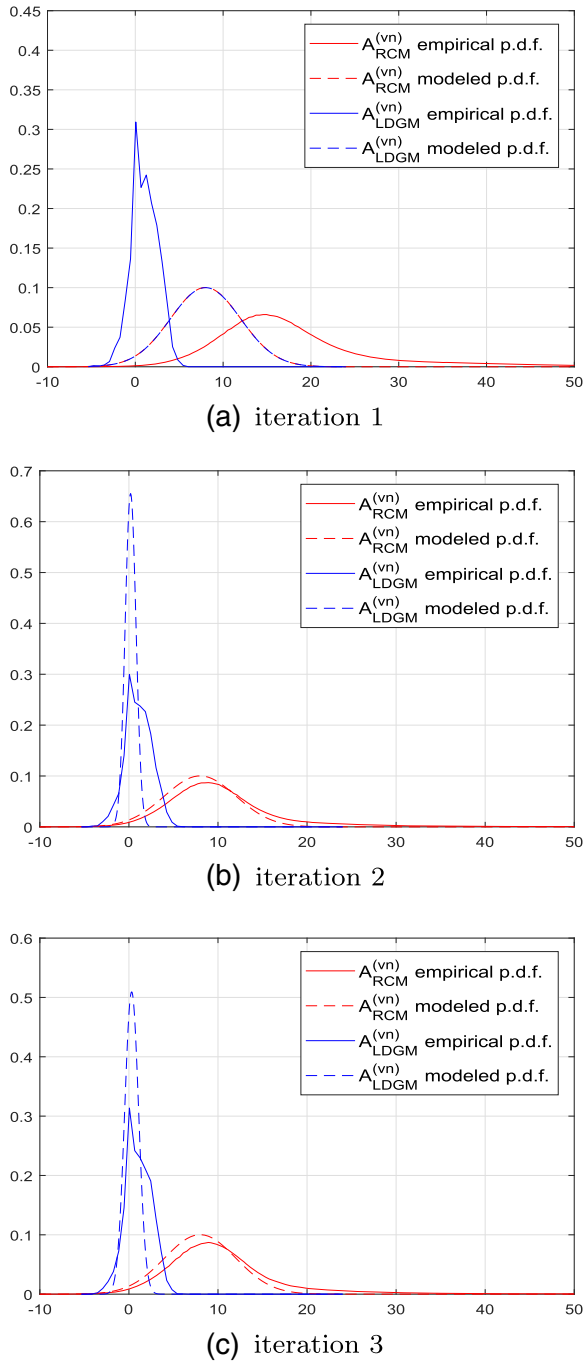

Fig. 12 Example of the iterative procedure to find κ for the EXIT chart of Fig. 8c. The histograms (conditioned to $U = 1$) obtained by Monte Carlo simulations are plotted in blue and the corresponding modeled conditional PDFs in red (refer to (11) and (12)). **a** Iteration 1. **b** Iteration 2. **c** Iteration 3

Abbreviations
ACM: Adaptive coded modulation; APSK: Amplitude and phase-shift keying; AWGN: Additive white Gaussian noise; BER: Bit error rate; BICM:Bit-interleaved coded modulation; CN(D): Check node (decoder); EXIT: Extrinsic information transfer; LDGM: Low-density generator matrix; LDPC: Low-density parity check; LLR: Log-likelihood ratio; PDF: Probability density function; QAM: Quadrature amplitude modulation; RCM: Rate compatible modulation; RV: Random variable; SNR: Signal to noise ratio; SP(A): Sum-product (algorithm); VN(D): Variable node (decoder)

Funding
This work was supported in part by the Spanish Ministry of Economy and Competitiveness through the CARMEN project (TEC2016-75067-C4-3-R), the COMONSENS network (TEC2015-69648-REDC) and by NSF Award CCF-1618653.

Availability of data and materials
Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Authors' contributions
JG-F and PMC conceived the research question. IG and PMC proved the main results. IG, PMC, and JG-F wrote the paper. All authors have read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Department of Basic Science, University of Navarra, Mikeletegi Pasealekua, 48, 20018 San Sebastian, Spain. [2]Department of Electrical and Computer Engineering, University of Delaware, 307 Evans Hall, 19716 Newark, USA.

Received: 30 July 2018 Accepted: 14 December 2018
Published online: 14 January 2019

References

1. H. Cui, C. Luo, J. Wu, C. W. Chen, F. Wu, Compressive coded modulation for seamless rate adaptation. IEEE Trans. Wirel. Commun. **12**(10), 4892–4904 (2013)
2. H. Cui, C. Luo, K. Tan, F. Wu, C. W. Chen, in *Proceedings of the 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*. Seamless rate, adaptation for wireless networking (ACM, New York, 2011), pp. 437–446. http://doi.acm.org/10.1145/2068897.2068971.
3. W. Zhong, J. Garcia-Frias, LDGM codes for channel coding and joint source-channel coding of correlated sources. EURASIP J. Appl. Sig. Process. **2005**, 942–953 (2005)
4. W. Zhong, H. Chai, J. Garcia-Frias, in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium On*. Approaching the Shannon limit through parallel concatenation of regular LDGM codes (IEEE, 2005), pp. 1753–1757. https://doi.org/10.1109/ISIT.2005.1523646
5. J. M. Kroll, N. Phamdo, Source-channel optimized trellis codes for bitonal image transmission over AWGN channels. IEEE Trans. Image Process. **8**(7), 899–912 (1999)
6. P. Burlina, F. Alajaji, An error resilient scheme for image transmission over noisy channels with memory. IEEE Trans. Image Process. **7**(4), 593–600 (1998)
7. G.-C. Zhu, F. Alajaji, Turbo codes for nonuniform memoryless sources over noisy channels. IEEE Commun. Lett. **6**(2), 64–66 (2002)
8. G.-C. Zhu, F. Alajaji, J. Bajcsy, P. Mitran, Transmission of nonuniform memoryless sources via nonsystematic turbo codes. IEEE Trans. Commun. **52**(8), 1344–1354 (2004)
9. J. Hagenauer, Source-controlled channel decoding. IEEE Trans. Commun. **43**(9), 2449–2457 (1995)
10. L. Li, J. Garcia-Frias, in *Information Sciences and Systems (CISS), 2014 48th Annual Conference On*. Hybrid analog-digital coding scheme based on parallel concatenation of linear random projections and LDGM codes (IEEE, 2014), pp. 1–6. https://doi.org/10.1109/CISS.2014.6814118
11. L. Li, J. Garcia-Frias, in *Information Sciences and Systems (CISS), 2015 49th Annual Conference On*. Hybrid analog-digital coding for nonuniform memoryless sources (IEEE, 2015), pp. 1–5. https://doi.org/10.1109/CISS.2015.7086861
12. G. Caire, G. Taricco, E. Biglieri, Bit-interleaved coded modulation. IEEE Trans. Inf. Theory. **44**(3), 927–946 (1998)
13. F. Cabarcas, R. D. Souza, J. Garcia-Frias, Turbo coding of strongly nonuniform memoryless sources with unequal energy allocation and PAM signaling. IEEE Trans. Sig. Process. **54**(5), 1942–1946 (2006)
14. I. Ochoa, P. M. Crespo, M. Hernaez, LDPC codes for non-uniform memoryless sources and unequal energy allocation. IEEE Commun. Lett. **14**(9), 794–796 (2010)
15. S. Ten Brink, Convergence behavior of iteratively decoded parallel concatenated codes. IEEE Trans. Commun. **49**(10), 1727–1737 (2001)
16. S. Ten Brink, G. Kramer, A. Ashikhmin, Design of low-density parity-check codes for modulation and detection. IEEE Trans. Commun. **52**(4), 670–678 (2004)
17. J. Wu, Z. Teng, H. Cui, C. Luo, C. xHuang, H.-H. Chen, Arithmetic-BICM for seamless rate adaptation for wireless communication systems. IEEE Syst. J. **10**(1), 228–239 (2016)
18. J. Du, L. Yang, J. Yuan, L. Zhou, X. He, Bit mapping design for LDPC coded BICM schemes with multi-edge type exit chart. IEEE Commun. Lett. **21**(4), 722–725 (2017)
19. T. Cheng, K. Peng, J. Song, K. Yan, EXIT-aided bit mapping design for LDPC coded modulation with APSK constellations. IEEE Commun. Lett. **16**(6), 777–780 (2012)
20. D. J. MacKay, Good error-correcting codes based on very sparse matrices. IEEE Trans. Inf. Theory. **45**(2), 399–431 (1999)
21. F. R. Kschischang, B. J. Frey, H.-A. Loeliger, Factor graphs and the sum-product algorithm. IEEE Trans. Inf. Theory. **47**(2), 498–519 (2001)