Parity-Check Density versus Performance of Binary Linear Block Codes: New Bounds and Applications

Gil Wiechman, *Student Member, IEEE,* Igal Sason, *Member, IEEE*

Abstract

The moderate complexity of low-density parity-check (LDPC) codes under iterative decoding is attributed to the sparseness of their parity-check matrices. It is therefore of interest to consider how sparse parity-check matrices of binary linear block codes can be as a function of the gap between their achievable rates and the channel capacity. This issue was addressed by Sason and Urbanke, and it is revisited in this paper. The remarkable performance of LDPC codes under practical and sub-optimal decoding algorithms motivates one to assess the inherent loss in performance which is attributed to the structure of the code or ensemble under maximum-likelihood (ML) decoding, and the additional loss which is imposed by the sub-optimality of the decoder. These issues are addressed by obtaining upper bounds on the achievable rates of binary linear block codes, and lower bounds on the asymptotic density of their parity-check matrices as a function of the gap between their achievable rates and the channel capacity; these bounds are valid under ML decoding, and hence, they are valid for any sub-optimal decoding algorithm. The new bounds improve on previously reported results by Burshtein et al. and by Sason and Urbanke, and they hold for the case where the transmission takes place over an arbitrary memoryless binary-input output-symmetric (MBIOS) channel. The significance of these information-theoretic bounds is in assessing the tradeoff between the asymptotic performance of LDPC codes and their decoding complexity (per iteration) under message-passing decoding. They are also helpful in studying the potential achievable rates of ensembles of LDPC codes under optimal decoding; by comparing these thresholds with those calculated by the density evolution technique, one obtains a measure for the asymptotic sub-optimality of iterative decoding algorithms.

Index Terms

Block codes, iterative decoding, linear codes, low-density parity-check (LDPC) codes, maximum-likelihood (ML) decoding, thresholds.

I. Introduction

Error-correcting codes which employ iterative decoding algorithms are now considered state of the art in the field of low-complexity coding techniques.

In [5], Khandekar and McEliece suggested to study the encoding and decoding complexity for ensembles of codes defined on graphs where the complexity is expressed in terms of the gap between the achievable rates of these ensembles and the channel capacity. They conjectured that if the achievable rate under message-passing iterative (MPI) decoding is a fraction $1-\varepsilon$ of the channel capacity, then for a wide class of channels, the encoding complexity scales like $\ln\frac{1}{\varepsilon}$ and the decoding complexity scales like $\frac{1}{\varepsilon}\ln\frac{1}{\varepsilon}$. The only exception is the binary erasure channel (BEC) where the decoding complexity behaves like $\ln\frac{1}{\varepsilon}$ (same as encoding complexity) due to the absolute reliability of the messages passed through the edges of the graph (hence, every edge can be used only once during the process of the iterative decoding).

Low-density parity-check (LDPC) codes have remarkable performance under practical iterative decoding algorithms, and they are efficiently encoded and decoded due to the sparseness of their parity-check matrices. Consider the number of ones in a parity-check matrix which represents a binary linear block code, and normalize it per information bit (i.e., with respect to the dimension of the code). This quantity (which will be later defined as the *density* of the parity-check matrix) is equal to the normalized number of left to right (or right to left) messages

The paper was published in the *IEEE Trans. on Information Theory*, vol. 53, no. 2, pp. 550 - 579, February 2007. It was submitted in May 23rd, 2005 and revised in September 21st, 2006. This work was presented in part at the Forty-third annual Allerton conference on Communication, Control and Computing, Monticello, Illinois, USA, September 28–30, 2005. Communicated by Gilles Zemor, Associate Editor for Coding Theory.

The authors are with the Department of Electrical Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel (email: {igillw@tx, sason@ee}.technion.ac.il).

per information bit which are passed in the corresponding bipartite graph during a single iteration of the MPI decoder. In [14], Sason and Urbanke considered how sparse parity-check matrices of binary linear block codes can be, as a function of their gap to capacity (where this gap depends in general on the channel and on the decoding algorithm). An information-theoretic lower bound on the asymptotic density of parity-check matrices was derived in [14, Theorem 2.1] where this bound holds for every sequence of binary linear block codes transmitted over an arbitrary memoryless binary-input output-symmetric (MBIOS) channel, and achieving a fraction $1-\varepsilon$ of the channel capacity with vanishing bit error probability. It holds for an arbitrary representation of these codes by full-rank parity-check matrices, and is of the form $\frac{K_1+K_2 \ln \frac{1}{\varepsilon}}{1-\varepsilon}$ where K_1 and K_2 are constants which only depend on the channel. Though the logarithmic behavior of this lower bound is in essence correct (due to a logarithmic behavior of the upper bound on the asymptotic parity-check density in [14, Theorem 2.2]), the lower bound in [14, Theorem 2.1] is *not* tight (with the exception of the BEC, as demonstrated in [14, Theorem 2.3], and possibly also the BSC).

In his thesis [3], Gallager proved that right-regular LDPC codes (i.e., LDPC codes with a constant degree (a_{R}) of the parity-check nodes) cannot achieve the channel capacity on a binary symmetric channel (BSC), even under maximum-likelihood (ML) decoding. This inherent gap to capacity is well approximated by an expression which decreases to zero exponentially fast in a_{R}. Richardson et al. [12] have extended this result, and proved that the same conclusion holds if a_{R} designates the *maximal right degree*. Sason and Urbanke later observed in [14] that the result still holds when considering the *average right degree*. Gallager's bound [3, Theorem 3.3] provides an upper bound on the rate of right-regular LDPC codes which achieve reliable communications over the BSC. Burshtein et al. have generalized Gallager's bound for general ensembles of LDPC codes transmitted over MBIOS channels [1]; to this end, they applied a two-level quantization to the log-likelihood ratio (LLR) of these channels which essentially turns them into a BSC. The derivation of the bounds in this paper was motivated by the desire to improve the results in [1, Theorems 1 and 2] and [14, Theorem 2.1] which are based on a two-level quantization of the LLR.

In [7], Measson et al. derived an upper bound on the thresholds under ML decoding of LDPC ensembles transmitted over the BEC. Their general approach relies on EXIT charts, having a surprising and deep connection with the maximum a posteriori (MAP) threshold due to the area theorem for the BEC. Generalized extrinsic information transfer (GEXIT) charts were recently introduced by Measson et al. [6]; GEXIT charts form a generalization of the concept of EXIT charts, and they satisfy the area theorem for an arbitrary MBIOS channel (see [13, Section 3.4.10]). This conservation law enables one to get upper bounds on the thresholds of turbo-like ensembles under bit-MAP decoding. The bound was shown to be tight for the BEC [7], and is conjectured to be tight in general for MBIOS channels [6].

A new method for analyzing LDPC codes and low-density generator-matrix (LDGM) codes under bit-MAP decoding is introduced by Montanari in [8]. The method is based on a rigorous approach to spin glasses, and allows a construction of lower bounds on the entropy of the transmitted message conditioned on the received one. The calculation of this bound is rather complicated, and its complexity grows exponentially with the maximal right and left degrees (see [8, Eqs. (6.2) and (6.3)]); this imposes a considerable difficulty on its calculation (especially, for continuous-output channels). Since the bounds in [7], [8] are derived for ensembles of codes, they are probabilistic in their nature; based on concentration arguments, they hold asymptotically in probability 1 as the block length goes to infinity. Based on heuristic statistical mechanics calculations, it was conjectured that the bounds in [8], which hold for general LDPC and LDGM ensembles over MBIOS channels, are tight.

The new bounds introduced in this paper provide upper bounds on the achievable rates of LDPC codes under ML decoding, and lower bounds on their asymptotic parity-check density. These information-theoretic bounds are exemplified in this paper in the context of providing some insight on the tradeoff between the asymptotic performance of LDPC codes and their decoding complexity (per iteration) under message-passing decoding. They are also used in this paper for studying the potential achievable rates of ensembles of LDPC codes under optimal decoding; by comparing these thresholds with the values which are calculated by the density evolution technique, one obtains a measure for the sub-optimality of iterative decoding algorithms in the limit where we let the block length tend to infinity. We note that the information-theoretic bounds in [1], [14] and this paper are valid for *every* sequence of binary linear block codes, in contrast to high probability results. As examples for the latter category of probabilistic bounds which apply to ensembles, the reader is referred to the recent bounds of Montanari [8] under

MAP decoding, the bound of Measson et al. for the BEC under MAP decoding [7], and the previously derived bound of Shokrollahi, relying on density evolution analysis for the BEC [17]. Shokrollahi proved in [17] that when the codes are communicated over a BEC, the growth rate of the average right degree (i.e., the average degree of the parity-check nodes in a bipartite Tanner graph) is at least logarithmic in terms of the gap to capacity. The statement in [17] is a high probability result which assumes a sub-optimal (iterative) decoding algorithm, whereas the statements in [1], [14] and this paper are valid even under ML decoding. As mentioned above, the bounds in [7], [8] refer to MAP decoding, but they form high probability results as the block length gets large.

The structure of the paper is the following: Preliminary material is presented in Section II, and the theorems are introduced and proved in Sections III and IV. The derivation of the bounds in Section III was motivated by the desire to generalize the results in [1, Theorems 1 and 2] and [14, Theorem 2.1]. A two-level quantization of the log-likelihood ratio (LLR), in essence replacing the arbitrary MBIOS channel by a physically degraded BSC, is modified in Section III to a quantized channel which better reflects the statistics of the original channel (though the quantized channel is still physically degraded w.r.t. the original channel). The number of quantization levels of the LLR for the new channel is an arbitrary integer power of 2, and the calculation of these bounds is subject to an optimization of the quantization levels, as to obtain the tightest bounds within their form. In Section IV, the analysis relies on the conditional pdf of the LLR at the output of the considered MBIOS channel, and operates on an equivalent channel without quantizing the LLR. This second approach leads in Section IV to bounds which are uniformly tighter than the bounds derived in Section III and are easier to calculate. The significance of the quantized and un-quantized bounds in Sections III and IV, respectively, stems from a comparison between these bounds which gives insight on the effect of the number of quantization levels of the LLR (even if they are optimally determined) on the achievable rates, as compared to the ideal case where no quantization is done. Numerical results are exemplified in Section V. Finally, in Section VI, we summarize and present interesting issues which deserve further research. Appendices provide further technical details referring to the proofs in Sections III and IV.

We note that the statements in this paper refer to the case where the parity-check matrices are full rank. At first glance, this requirement poses a problem when considering ensembles of LDPC codes; a parity-check matrix, referring to a randomly chosen bipartite graph with a given pair of degree distributions, may not be full rank (even asymptotically, as we let the block length tend to infinity).[1] As explained in Section V, the statements in this paper still hold for ensembles of LDPC codes when we replace the code rate with the design rate.

II. Preliminaries

We introduce here some definitions and theorems from [1], [14] which serve as preliminary material for the rest of the paper. Definitions 2.1 and 2.2 are taken from [14, Section 2].

Definition 2.1: (Capacity-Approaching Codes) Let $\{\mathcal{C}_m\}$ be a sequence of codes, and denote the rate of the code $\mathcal{C}_m$ by R_m. Assume that for every m, the codewords of the code $\mathcal{C}_m$ are transmitted with equal probability over a channel whose capacity is C. This sequence is said to *achieve a fraction* $1-\varepsilon$ *of the channel capacity with vanishing bit error probability* if $\lim_{m\to\infty} R_m = (1-\varepsilon)C$, and there exists a decoding algorithm under which the average bit error probability of the code $\mathcal{C}_m$ tends to zero in the limit where $m \to \infty$.

Definition 2.2: (Parity-Check Density) Let $\mathcal{C}$ be a binary linear code of rate R and block length n, which is represented by a parity-check matrix H. We define the *density* of H, call it $\Delta = \Delta(H)$, as the normalized number of ones in H *per information bit*. The total number of ones in H is therefore equal to $nR\Delta$.

Definition 2.3: (Log-Likelihood Ratio (LLR)) Let us consider an MBIOS channel whose conditional pdf is $p_{Y|X}$ where X and Y designate the channel input and output, respectively. The log-likelihood ratio (LLR) at the output of the channel is

$$\text{LLR}(y) \triangleq \ln\left(\frac{p_{Y|X}(y|0)}{p_{Y|X}(y|1)}\right).$$

Throughout the paper, we assume that all the codewords of a binary linear block code are equally likely to be transmitted. Also, the function h_2 designates the binary entropy function to base 2, i.e., $h_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$.

[1] One can construct LDPC ensembles where the design rate is strictly less than the asymptotic rate as the block length goes to infinity, e.g., by simply repeating a non-vanishing fraction of the rows of the parity-check matrix.

Theorem 2.1: (An Upper Bound on the Achievable Rates for Reliable Communication over MBIOS Channels) [1, Theorem 2]: Consider a sequence $\{\mathcal{C}_m\}$ of binary linear block codes of rate R_m, and assume that their block length tends to infinity as $m \to \infty$. Let H_m be a full-rank parity-check matrix of the code $\mathcal{C}_m$, and assume that $\Gamma_{k,m}$ designates the fraction of the parity-check equations involving k variables. Let

$$\Gamma_k \triangleq \lim_{m\to\infty} \Gamma_{k,m}, \quad R \triangleq \lim_{m\to\infty} R_m \tag{1}$$

where these limits are assumed to exist. Suppose that the transmission of these codes takes place over an MBIOS channel with capacity C bits per channel use, and let

$$w \triangleq \frac{1}{2}\int_{-\infty}^{\infty} \min\bigl(f(y), f(-y)\bigr)\, dy \tag{2}$$

where $f(y) \triangleq p_{Y|X}(y|0)$ designates the conditional pdf of the output of the MBIOS channel when zero is transmitted. Then, a necessary condition for vanishing block error probability as $m \to \infty$ is

$$R \leq 1 - \frac{1-C}{\displaystyle\sum_k \left\{\Gamma_k \, h_2\left(\frac{1-(1-2w)^k}{2}\right)\right\}}.$$

Theorem 2.2: (Lower Bounds on the Asymptotic Parity-Check Density with Two-Level Quantization) [14, Theorem 2.1]: Let $\{\mathcal{C}_m\}$ be a sequence of binary linear block codes achieving a fraction $1-\varepsilon$ of the capacity of an MBIOS channel with vanishing bit error probability. Denote Δ_m as the density of a full-rank parity-check matrix of the code $\mathcal{C}_m$. Then, the asymptotic density satisfies

$$\liminf_{m\to\infty} \Delta_m > \frac{K_1 + K_2 \ln\frac{1}{\varepsilon}}{1-\varepsilon} \tag{3}$$

where

$$K_1 = \frac{(1-C)\ \ln\left(\frac{1}{2\ln 2}\ \frac{1-C}{C}\right)}{2C\ \ln\left(\frac{1}{1-2w}\right)}\ , \quad K_2 = \frac{1-C}{2C\ \ln\left(\frac{1}{1-2w}\right)} \tag{4}$$

and w is defined in (2). For a BEC with erasure probability p, the coefficients K_1 and K_2 in (4) are improved to

$$K_1 = \frac{p\ \ln\left(\frac{p}{1-p}\right)}{(1-p)\ \ln\left(\frac{1}{1-p}\right)}\ , \quad K_2 = \frac{p}{(1-p)\ \ln\left(\frac{1}{1-p}\right)}\ . \tag{5}$$

The bounds in Theorems 2.1 and 2.2 are applied in [1] and [14], respectively, to ensembles of LDPC codes. In general, LDPC codes are linear block codes which are represented by a sparse parity-check matrix H. This matrix can be represented in an equivalent form by a bipartite graph $\mathcal{G}$ whose variable nodes (appearing on the left of $\mathcal{G}$) represent the code bits, and whose parity-check nodes (appearing on the right of $\mathcal{G}$) represent the linear constraints defined by H. In such a bipartite graph, an edge connects a variable node with a parity-check node if and only if the corresponding code bit is involved in the parity-check equation; the degree of a node is defined as the number of edges which are adjacent to it.

Following standard notation, let λ_i and ρ_i denote the fraction of *edges* attached to variable and parity-check nodes of degree i, respectively. In a similar manner, let Λ_i and Γ_i denote the fraction of variable and parity-check nodes of degree i, respectively. The LDPC ensemble is characterized by a triplet (n, λ, ρ) where n designates the block length of the codes, and the polynomials

$$\lambda(x) \triangleq \sum_{i=1}^{\infty} \lambda_i x^{i-1}, \qquad \rho(x) \triangleq \sum_{i=1}^{\infty} \rho_i x^{i-1}$$

represent, respectively, the left and right degree distributions (d.d.) from the *edge* perspective. Equivalently, this ensemble can be also characterized by the triplet (n, Λ, Γ) where the polynomials

$$\Lambda(x) \triangleq \sum_{i=1}^{\infty} \Lambda_i x^i, \qquad \Gamma(x) \triangleq \sum_{i=1}^{\infty} \Gamma_i x^i$$

represent, respectively, the left and right d.d. from the *node* perspective. We denote by LDPC(n, λ, ρ) (or LDPC(n, Λ, Γ)) the ensemble of codes whose bipartite graphs are constructed according to the corresponding pairs of degree distributions. One can switch between degree distributions w.r.t. to the nodes and edges of a bipartite graph, using the following equations [13]:

$$\Lambda(x) = \frac{\int_0^x \lambda(u)du}{\int_0^1 \lambda(u)du}\ , \qquad \Gamma(x) = \frac{\int_0^x \rho(u)du}{\int_0^1 \rho(u)du} \tag{6}$$

$$\lambda(x) = \frac{\Lambda'(x)}{\Lambda'(1)}\ , \qquad \rho(x) = \frac{\Gamma'(x)}{\Gamma'(1)}\ . \tag{7}$$

An important characteristic of an ensemble of LPDC codes is its *design rate*. For an LDPC ensemble whose codes are represented by parity-check matrices of dimension $c \times n$, the design rate is defined to be $R_{\text{d}} \triangleq 1 - \frac{c}{n}$. This serves as a lower bound on the actual rate of any code from this ensemble, and is equal to the actual rate if the parity-check matrix of a code is *full rank* (i.e., if the linear constraints which define this code are linearly independent). For an ensemble of LDPC codes, the design rate is given in terms of the degree distributions (either w.r.t. the edges or nodes of a graph), and it can be expressed in two equivalent forms:

$$R_{\text{d}} = 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx} = 1 - \frac{\Lambda'(1)}{\Gamma'(1)}\ . \tag{8}$$

A sufficient condition for the asymptotic convergence of the rate of a code, chosen uniformly at random from an LDPC ensemble, to its design rate was stated in [7, Lemma 7].

Lemma 2.1: (A sufficient condition for the equality between the design rate and asymptotic rate for ensembles of LDPC codes) [7, Lemma 7]: Let $\mathcal{C}$ be a code which is chosen uniformly at random from the ensemble LDPC(n, Λ, Γ), let R be the rate of $\mathcal{C}$, and let R_{d} be the design rate of this ensemble. Consider the function

$$\begin{aligned}\Psi_{(\Lambda,\Gamma)}(u) \quad &\triangleq \quad -\Lambda'(1)\log_2\left[\frac{1+uv}{(1+u)(1+v)}\right] \\ &\quad + \sum_{i=1}^{\infty} \Lambda_i \log_2\left[\frac{1+u^i}{2(1+u)^i}\right] \\ &\quad + \frac{\Lambda'(1)}{\Gamma'(1)} \sum_{i=1}^{\infty} \Gamma_i \log_2\left[1 + \left(\frac{1-v}{1+v}\right)^i\right]\end{aligned}$$

where

$$v \triangleq \left(\sum_{i=1}^{\infty} \frac{\lambda_i}{1+u^i}\right)^{-1} \left(\sum_{i=1}^{\infty} \frac{\lambda_i u^{i-1}}{1+u^i}\right).$$

Assume that the function $\Psi_{(\Lambda,\Gamma)}$ achieves its global maximum in the range $u \in [0, \infty)$ at $u = 1$. Then, there exists a constant $B > 0$ such that for any $\xi > 0$ and $n > n_0(\xi, \Lambda, \Gamma)$

$$\Pr\{|R - R_{\text{d}}| > \xi\} \leq e^{-Bn\xi}\ .$$

Moreover, there exists a constant $C > 0$ such that for $n > n_0(\xi, \Lambda, \Gamma)$

$$\mathbb{E}\left\{|R - R_{\text{d}}|\right\} \leq \frac{C \ln n}{n}\ .$$

In Section V, we rely on this lemma in order to verify that the asymptotic rates of codes randomly chosen (with uniform distribution) from various ensembles of LDPC codes tend in probability 1 to the design rates of these ensembles.

III. Approach I: Bounds Based on Quantization of the Log-Likelihood Ratio

In this section, we introduce bounds on the achievable rates and the asymptotic parity-check density of sequences of binary linear block codes. The bounds generalize previously reported results in [1] and [14] which were based on a symmetric two-level quantization of the LLR. This is achieved by extending the quantization to a number of levels which is equal to an arbitrary natural power of 2. In Section III-A, we particularize the results and their proofs for a four-level quantization. In Section III-B, the results are extended to a symmetric quantization with a number of levels which is an arbitrary natural power of 2. This order of presentation was chosen since many concepts which are helpful for the generalization in Section III-B are written in a simplified notation for the four-level quantization, along with all the relevant lemmas for the general case which are already introduced in the derivation of the bound with four-level quantization. This also shortens considerably the proof for the general quantization in Section III-B.

A. Bounds for Four-Levels of Quantization

As a preparatory step towards developing bounds on the parity-check density and the rate of binary linear block codes, we present a lower bound on the conditional entropy of a transmitted codeword given the received sequence at the output of an arbitrary MBIOS channel.

Proposition 3.1: Let $\mathcal{C}$ be a binary linear block code of length n and rate R, and assume that its transmission takes place over an MBIOS channel whose conditional pdf is given by $p_{Y|X}$. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ designate the transmitted codeword and received sequence, respectively. For an arbitrary positive $l \in \mathbb{R}^+$, let us define the probabilities p_0, p_1, p_2, p_3 as follows:

$$\begin{aligned} p_0 &\triangleq \Pr\{\text{LLR}(Y) > l \mid X = 0\} \\ p_1 &\triangleq \Pr\{\text{LLR}(Y) \in (0, l] \mid X = 0\} \\ &\quad + \frac{1}{2} \Pr\{\text{LLR}(Y) = 0 \mid X = 0\} \\ p_2 &\triangleq \Pr\{\text{LLR}(Y) \in [-l, 0) \mid X = 0\} \\ &\quad + \frac{1}{2} \Pr\{\text{LLR}(Y) = 0 \mid X = 0\} \\ p_3 &\triangleq \Pr\{\text{LLR}(Y) < -l \mid X = 0\}. \end{aligned} \tag{9}$$

For an arbitrary full-rank parity-check matrix of the code $\mathcal{C}$, let Γ_k designate the fraction of parity-check equations involving k variables. Then, the conditional entropy of the transmitted codeword, given the received sequence, satisfies

$$\begin{aligned} \frac{H(\mathbf{X}|\mathbf{Y})}{n} &\geq 1 - C - (1 - R) \\ &\cdot \sum_k \Bigg\{ \Gamma_k \sum_{t=0}^{k} \binom{k}{t} (p_1 + p_2)^t (p_0 + p_3)^{k-t} \\ &\quad \cdot h_2\Bigg(\frac{1 - \left(1 - \frac{2p_2}{p_1+p_2}\right)^t \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2} \Bigg) \Bigg\}. \end{aligned} \tag{10}$$

Remark 3.1: Note that the input vector $\mathbf{X}$ is chosen uniformly from the codewords of a binary linear block code. Each input bit X_i therefore either gets the values 0 or 1 with probability $\frac{1}{2}$ or is set to zero (due to the linearity of the code). In the following proof, we assume that all the code symbols get the values 0 or 1 with equal probability. By slightly modifying the proof, it is simple to show that the bound also holds for the other case where some of the code bits are set to zero. Without mentioning explicitly, the same assumption will be taken in the proofs of Propositions 3.2 and 4.1.

Proof: Considering an MBIOS channel whose conditional pdf is given by $p_{Y|X}$, we introduce a new physically degraded channel. It is a binary-input, quaternary-output symmetric channel (see Fig. 1). To this end, let $l \in \mathbb{R}^+$ be an arbitrary positive number, and let α be a primitive element of the Galois field GF(2^2) (so $\alpha^2 = 1 + \alpha$). The set of elements of this field is $\{0, 1, \alpha, 1 + \alpha\}$. Let X_i and Y_i designate the random variables referring to the input and output of the original channel at time i (where $i = 1, 2, \ldots, n$). We define the degraded channel as a channel

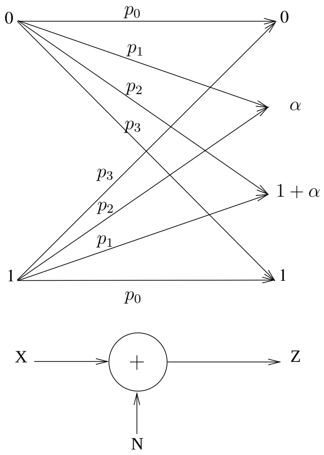

Fig. 1. The channel model in the upper plot is a physically degraded channel used for the derivation of the bound with four levels of quantization. The element α denotes a primitive element in GF(2^2). This channel model is equivalent to a channel with an additive noise in GF(2^2) (see lower plot).

with four quantization levels of the LLR. The output of the degraded channel at time i, Z_i, is calculated from the output Y_i of the original channel as follows:

- If LLR$(Y_i) > l$, then $Z_i = 0$.
- If $0 <$ LLR$(Y_i) \leq l$, then $Z_i = \alpha$.
- If $-l \leq$ LLR$(Y_i) < 0$, then $Z_i = 1 + \alpha$.
- If LLR$(Y_i) < -l$, then $Z_i = 1$.
- If LLR$(Y_i) = 0$, then $Z_i = \alpha$ or $Z_i = 1 + \alpha$ w.p. $\frac{1}{2}$.

From the definition of the degraded channel in Fig. 1, this channel has an additive noise in GF(2^2) and is also binary-input output-symmetric. It follows that the transition probabilities of the degraded channel are

$$\begin{aligned}
p_0 &= \Pr(Z = 0 \mid X = 0) = \Pr(Z = 1 \mid X = 1)\\
p_1 &= \Pr(Z = \alpha \mid X = 0) = \Pr(Z = 1 + \alpha \mid X = 1)\\
p_2 &= \Pr(Z = 1 + \alpha \mid X = 0) = \Pr(Z = \alpha \mid X = 1)\\
p_3 &= \Pr(Z = 1 \mid X = 0) = \Pr(Z = 0 \mid X = 1)
\end{aligned}$$

where p_j is introduced in (9) for $0 \leq j \leq 3$, and the symmetry in these transition probabilities holds since the original channel is MBIOS.

Since $\mathcal{C}$ is a binary linear block code of length n and rate R, and the codewords are transmitted with equal probability then

$$H(\mathbf{X}) = nR. \tag{11}$$

Also, since the channel is memoryless, then

$$H(\mathbf{Y}|\mathbf{X}) = nH(Y|X). \tag{12}$$

We designate the output sequence of the degraded channel by $\mathbf{Z} = (Z_1, \ldots, Z_n)$. Since the mapping from Y_i to the degraded output Z_i $(i = 1, 2, \cdots, n)$ is memoryless, then $H(\mathbf{Z}|\mathbf{Y}) = nH(Z|Y)$ and

$$\begin{aligned}
H(\mathbf{Y}) &= H(\mathbf{Z}) - H(\mathbf{Z}|\mathbf{Y}) + H(\mathbf{Y}|\mathbf{Z})\\
&= H(\mathbf{Z}) - nH(Z|Y) + H(\mathbf{Y}|\mathbf{Z})
\end{aligned} \tag{13}$$

$$\begin{aligned}
H(\mathbf{Y}|\mathbf{Z}) &\leq \sum_{i=1}^{n} H(Y_i|Z_i)\\
&= nH(Y|Z)\\
&= n\left[H(Y) - H(Z) + H(Z|Y)\right].
\end{aligned} \tag{14}$$

Applying the above towards a lower bound on the conditional entropy $H(\mathbf{X}|\mathbf{Y})$, we get

$$\begin{aligned} H(\mathbf{X}|\mathbf{Y}) &= H(\mathbf{X}) + H(\mathbf{Y}|\mathbf{X}) - H(\mathbf{Y}) \\ &= nR + nH(Y|X) - H(\mathbf{Y}) \\ &= nR + nH(Y|X) - H(\mathbf{Z}) - H(\mathbf{Y}|\mathbf{Z}) \\ &\quad + nH(Z|Y) \\ &\geq nR + nH(Y|X) - H(\mathbf{Z}) \\ &\quad - n\left[H(Y) - H(Z) + H(Z|Y)\right] + nH(Z|Y) \\ &= nR - H(\mathbf{Z}) + nH(Z) - n\left[H(Y) - H(Y|X)\right] \\ &= nR - H(\mathbf{Z}) + nH(Z) - nI(X;Y) \\ &\geq nR - H(\mathbf{Z}) + nH(Z) - nC \end{aligned} \tag{15}$$

where the second equality relies on (11) and (12), the third equality relies on (13), the first inequality relies on (14), and $I(X;Y) \leq C$ is used for the last transition (where C designates the capacity of the original channel).

In order to obtain a lower bound on $H(\mathbf{X}|\mathbf{Y})$ from (15), we calculate the entropy of the random variable Z, and derive an upper bound on the entropy of the random vector $\mathbf{Z}$. This finally provides the lower bound in (10). Observing that the degraded channel is additive over GF(2^2), we denote the additive noise by

$$N_i = \Theta_i + \Omega_i \alpha, \quad i \in \{1, \ldots, n\}$$

where $\mathbf{\Theta} = (\Theta_1, \ldots, \Theta_n)$ and $\mathbf{\Omega} = (\Omega_1, \ldots, \Omega_n)$ are random vectors over GF(2). Note that $\mathbf{\Theta}$ and $\mathbf{\Omega}$ are statistically independent of the transmitted codeword $\mathbf{X}$. Since the code is binary, it follows that

$$Z_i = \Phi_i + \Omega_i \alpha \tag{16}$$

where $\Phi_i \triangleq \Theta_i + X_i$. This gives

$$\begin{aligned} H(Z) &= H(\Phi, \Omega) \\ &= H(\Omega) + H(\Phi|\Omega) \\ &= H(\Omega) + 1 \end{aligned} \tag{17}$$

where the last equality follows since the input X is equally likely to be zero or one (see Remark 3.1) and the channel in Fig. 1 is MBIOS; since Ω is independent of X, then Φ is equally likely to be zero or one given the value of Ω.

We now derive an upper bound on the entropy $H(\mathbf{Z})$. Based on (16), it is easy to verify the following chain of equalities:

$$\begin{aligned} H(\mathbf{Z}) &= H(\mathbf{\Phi}, \mathbf{\Omega}) \\ &= H(\mathbf{\Omega}) + H(\mathbf{\Phi} \mid \mathbf{\Omega}) \\ &= n\, H(\Omega) + H(\mathbf{\Phi} \mid \mathbf{\Omega}) \end{aligned} \tag{18}$$

where $\mathbf{\Phi} \triangleq (\Phi_1, \ldots, \Phi_n)$, and the last equality follows since the degraded channel in Fig. 1 is memoryless. Let us define the syndrome at the output of the degraded channel as

$$\mathbf{S} \triangleq \mathbf{\Phi}\, H^T$$

where H is a full-rank parity-check matrix of the binary linear block code $\mathcal{C}$. We note that the calculation of the syndrome only takes into account the $\mathbf{\Phi}$-component of the vector $\mathbf{Z}$ in (16). Also note that since $\mathbf{X}H^T = 0$ for every codeword $\mathbf{X}$, then $\mathbf{S} = \mathbf{\Theta}\, H^T$ which is independent of the transmitted codeword. Let us define M as the index of the vector $\mathbf{\Phi}$ in the coset referring to the syndrome $\mathbf{S}$. Since each coset has exactly 2^{nR} elements which are equally likely, then $H(M) = nR$, and

$$\begin{aligned} H(\mathbf{\Phi} \mid \mathbf{\Omega}) &= H(\mathbf{S}, M \mid \mathbf{\Omega}) \\ &\leq H(M) + H(\mathbf{S} \mid \mathbf{\Omega}) \\ &= nR + H(\mathbf{S} \mid \mathbf{\Omega}). \end{aligned} \tag{19}$$

Considering a parity-check equation involving k variables, let $\{i_1, \ldots, i_k\}$ be the set of indices of the variables involved in this parity-check equation. The relevant component of the syndrome $\mathbf{S}$ which refers to this parity-check equation is equal to zero or one if and only if the Hamming weight of the sub-vector $(\Theta_{i_1}, \ldots, \Theta_{i_k})$ is even or odd, respectively. It is clear from Fig. 1 that for an index i for which $\Omega_i = 1$, Θ_i is equal to one in probability $\frac{p_2}{p_1+p_2}$. Similarly, for an index i for which $\Omega_i = 0$, then Θ_i is equal to one in probability $\frac{p_3}{p_0+p_3}$.

Given that the Hamming weight of the vector $(\Omega_{i_1}, \ldots, \Omega_{i_k})$ is t, then the probability of an even Hamming weight of the random vector $(\Theta_{i_1}, \ldots, \Theta_{i_k})$ is equal to

$$q_1(t,k)\ q_2(t,k) + \bigl(1 - q_1(t,k)\bigr)\ \bigl(1 - q_2(t,k)\bigr)$$

where $q_1(t,k)$ designates the probability that among the t indices i for which $\Omega_i = 1$, the random variable Θ_i is equal to 1 an even number of times, and $q_2(t,k)$ designates the probability that the same happens for the $k-t$ indices i for which $\Omega_i = 0$. Based on the discussion above, it follows that

$$\begin{aligned} q_1(t,k) &= \sum_{i \text{ even}} \left\{ \binom{t}{i} \left(\frac{p_1}{p_1+p_2}\right)^{t-i} \left(\frac{p_2}{p_1+p_2}\right)^{i} \right\} \\ &= \frac{1 + \left(1 - \frac{2p_2}{p_1+p_2}\right)^t}{2} \\ q_2(t,k) &= \sum_{i \text{ even}} \left\{ \binom{k-t}{i} \left(\frac{p_0}{p_0+p_3}\right)^{k-t-i} \left(\frac{p_3}{p_0+p_3}\right)^{i} \right\} \\ &= \frac{1 + \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2}\ . \end{aligned}$$

Hence, the probability that the vector $(\Theta_{i_1}, \Theta_{i_2}, \ldots, \Theta_{i_k})$ is of even Hamming weight is

$$\begin{aligned} & q_1(t,k)\ q_2(t,k) + \bigl(1 - q_1(t,k)\bigr)\ \bigl(1 - q_2(t,k)\bigr) \\ &= \frac{1 + \left(1 - \frac{2p_2}{p_1+p_2}\right)^t \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2}. \end{aligned}$$

We conclude that given a vector $\underline{\omega} \in \{0,1\}^k$ of Hamming weight t

$$\begin{aligned} & H\bigl(S_i \mid (\Omega_{i_1}, \ldots, \Omega_{i_k}) = \underline{\omega}\bigr) \\ &= h_2\left(\frac{1 + \left(1 - \frac{2p_2}{p_1+p_2}\right)^t \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2}\right). \end{aligned}$$

This yields that if the calculation of a component S_i (where $i = 1, \ldots, n(1-R)$) in the syndrome $\mathbf{S}$ relies on a parity-check equation involving k variables, then

$$\begin{aligned} & H(S_i \mid \mathbf{\Omega}) \\ &= H(S_i \,|\, \Omega_{i_1}, \ldots, \Omega_{i_k}) \\ &= \sum_{\underline{\omega} \in \{0,1\}^k} \Pr\bigl((\Omega_{i_1}, \ldots, \Omega_{i_k}) = \underline{\omega}\bigr) \\ &\qquad\qquad \cdot H\bigl(S_i \mid (\Omega_{i_1}, \ldots, \Omega_{i_k}) = \underline{\omega}\bigr) \\ &= \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \\ &\qquad \cdot h_2\left(\frac{1 + \left(1 - \frac{2p_2}{p_1+p_2}\right)^t \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2}\right) \end{aligned}$$

$$= \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \cdot h_2\left(\frac{1 - \left(1 - \frac{2p_2}{p_1+p_2}\right)^t \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2} \right)$$

where the third equality turns to averaging over the Hamming weight of $(\Omega_{i_1}, \ldots, \Omega_{i_k})$ (note that each component is Bernoulli distributed with $\Pr(\Omega_i = 0) = p_0 + p_3$), and the last equality follows from the symmetry of the binary entropy function (where $h_2(x) = h_2(1-x)$ for $x \in [0,1]$). Let Γ_k designate the fraction of parity-check equations in the full-rank parity-check matrix which involve k variables, so their total number is $n(1-R)\Gamma_k$ and

$$\begin{aligned} & H(\mathbf{S} \mid \mathbf{\Phi}) \\ & \leq \sum_{i=1}^{n(1-R)} H(S_i \,|\, \mathbf{\Phi}) \\ & = n(1-R) \sum_k \Bigg\{ \Gamma_k \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \\ & \qquad \cdot h_2\left(\frac{1 - \big(1 - \frac{2p_2}{p_1+p_2}\big)^t \big(1 - \frac{2p_3}{p_0+p_3}\big)^{k-t}}{2} \right) \Bigg\}. \end{aligned} \tag{20}$$

By combining (18)–(20), an upper bound on the entropy of the random vector $\mathbf{Z}$ follows:

$$\begin{aligned} H(\mathbf{Z}) \leq nR + nH(\Omega) + n(1-R) & \\ \cdot \sum_k \Bigg\{ \Gamma_k \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} & \\ \cdot h_2\left(\frac{1 - \big(1 - \frac{2p_2}{p_1+p_2}\big)^t \big(1 - \frac{2p_3}{p_0+p_3}\big)^{k-t}}{2} \right) \Bigg\}. & \end{aligned} \tag{21}$$

The substitution of (17) and (21) in (15) finally provides the lower bound on the conditional entropy $H(\mathbf{X} \,|\, \mathbf{Y})$ in (10). ■

The following theorem tightens the lower bound on the parity-check density of an arbitrary sequence of binary linear block codes given in [14, Theorem 2.1]. It is based on a four-level quantization of the LLR at the output of an MBIOS channel (as opposed to the two-level quantization of the LLR used in [14]).

Theorem 3.1: ("Four-Level Quantization" Lower Bound on the Asymptotic Parity-Check Density of Binary Linear Block Codes) Let $\{C_m\}$ be a sequence of binary linear block codes achieving a fraction $1 - \varepsilon$ of the capacity of an MBIOS channel with vanishing bit error probability. Let H_m be an arbitrary *full-rank* parity-check matrix of the code $\mathcal{C}_m$, and denote its density by Δ_m. Then, the asymptotic density satisfies

$$\liminf_{m \to \infty} \Delta_m > \frac{K_1 + K_2 \ln \frac{1}{\varepsilon}}{1 - \varepsilon} \tag{22}$$

where

$$\begin{aligned} K_1 &= K_2 \, \ln\left(\frac{1}{2\ln 2} \, \frac{1-C}{C} \right), \\ K_2 &= -\frac{1-C}{C \, \ln\left(\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3} \right)} \end{aligned} \tag{23}$$

and p_0, p_1, p_2, p_3 are defined in (9) in terms of $l \in \mathbb{R}^+$. The optimal value of l is given implicitly as a solution to the equation

$$\frac{p_2^2 + e^{-l} p_1^2}{(p_1+p_2)^2} = \frac{p_3^2 + e^{-l} p_0^2}{(p_0+p_3)^2} \tag{24}$$

where such a solution always exists.[2]

Proof: We first prove the lower bound in (22) and (23), and then, the optimization equation (24).

Derivation of the lower bound in (22) *and* (23):

Lemma 3.1: Let $\mathcal{C}$ be a binary linear block code of length n and rate R. Let P_{b} designate the average bit error probability of the code $\mathcal{C}$ which is associated with an arbitrary decoding algorithm and channel, and let $\mathbf{X}$ and $\mathbf{Y}$ designate the transmitted codeword and received sequence, respectively. Then

$$\frac{H(\mathbf{X} \mid \mathbf{Y})}{n} \leq R\, h_2(P_{\text{b}}). \tag{25}$$

Proof: The lemma is proved in Appendix I-A. ■

Lemma 3.2: $h_2(x) \leq 1 - \frac{2}{\ln 2}\left(\frac{1}{2} - x\right)^2$ for $0 \leq x \leq 1$.

Proof: The lemma is proved in [14, Lemma 3.1]; this inequality actually forms a particular case of Eq. (100) whose derivation is based on truncating the power series expansion of the binary entropy function around $\frac{1}{2}$. ■

Referring to an arbitrary sequence of binary linear block codes $\{\mathcal{C}_m\}$ which achieves a fraction $1-\varepsilon$ of capacity with vanishing bit error probability, then according to Definition 2.1, there exists a decoding algorithm (e.g., ML decoding) so that the average bit error probability of the code $\mathcal{C}_m$ tends to zero as m goes to infinity, and $\lim_{m\to\infty} R_m = (1-\varepsilon)C$. From Lemma 3.1, we get that $\lim_{m\to\infty} \frac{H(\mathbf{X}_m|\mathbf{Y}_m)}{n_m} = 0$ where $\mathbf{X}_m$ and $\mathbf{Y}_m$ designate the transmitted codeword of the code $\mathcal{C}_m$ and the received sequence, respectively, and n_m designates the block length of the code $\mathcal{C}_m$. From Proposition 3.1, we obtain

$$\begin{aligned}
&\frac{H(\mathbf{X}_m|\mathbf{Y}_m)}{n_m} \\
&\geq 1 - C - (1-R_m) \\
&\quad \cdot \sum_k \Bigg\{ \Gamma_{k,m} \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \\
&\qquad\qquad \cdot h_2\left(\frac{1 - \left(1 - \frac{2p_2}{p_1+p_2}\right)^t \left(1 - \frac{2p_3}{p_0+p_3}\right)^{k-t}}{2} \right) \Bigg\}
\end{aligned}$$

where $\Gamma_{k,m}$ designates the fraction of parity-check equations in a parity-check matrix H_m which involve k variables. The upper bound on the binary entropy function h_2 in Lemma 3.2 gives

$$\begin{aligned}
&\frac{H(\mathbf{X}_m|\mathbf{Y}_m)}{n_m} \\
&\geq 1 - C - (1-R_m) \qquad\qquad (26)\\
&\quad \cdot \sum_k \Bigg\{ \Gamma_{k,m} \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \\
&\qquad\qquad \cdot \left[1 - \frac{1}{2\ln 2} \left(\frac{p_1-p_2}{p_1+p_2}\right)^{2t} \left(\frac{p_0-p_3}{p_0+p_3}\right)^{2(k-t)} \right] \Bigg\}
\end{aligned}$$

Since $p_0+p_1+p_2+p_3 = 1$ (i.e., the transition probabilities of the degraded channel in Fig. 1 sum to 1), then

$$\begin{aligned}
&\sum_k \Bigg\{ \Gamma_{k,m} \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \\
&\qquad\qquad \cdot \left[1 - \frac{1}{2\ln 2} \left(\frac{p_1-p_2}{p_1+p_2}\right)^{2t} \left(\frac{p_0-p_3}{p_0+p_3}\right)^{2(k-t)} \right] \Bigg\} \\
&= \sum_k \Bigg\{ \Gamma_{k,m} \left[1 - \frac{1}{2\ln 2} \sum_{t=0}^{k} \binom{k}{t} \left(\frac{(p_1-p_2)^2}{p_1+p_2}\right)^t \right.
\end{aligned}$$

[2] It was observed numerically that the solution l of the optimization equation (24) is unique when considering the binary-input AWGN channel. We conjecture that the uniqueness of such a solution is a property which holds for MBIOS channels under some mild conditions.

$$\cdot \left(\frac{(p_0-p_3)^2}{p_0+p_3}\right)^{k-t}\Bigg]\Bigg\}$$

$$= 1 - \frac{1}{2\ln 2} \sum_k \Bigg\{ \Gamma_{k,m} \sum_{t=0}^{k} \binom{k}{t} \left(\frac{(p_1-p_2)^2}{p_1+p_2}\right)^{t} \cdot \left(\frac{(p_0-p_3)^2}{p_0+p_3}\right)^{k-t} \Bigg\}$$

$$= 1 - \frac{1}{2\ln 2} \sum_k \left\{ \Gamma_{k,m} \left(\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3}\right)^{k} \right\}$$

$$\leq 1 - \frac{1}{2\ln 2} \left(\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3}\right)^{a_{\mathrm{R}}(m)} \tag{27}$$

where $a_{\mathrm{R}}(m) \triangleq \sum_k k\Gamma_{k,m}$ designates the average right degree of the bipartite graph which refers to the parity-check matrix H_m, and the last transition follows from Jensen's inequality. Substituting (27) into the RHS of (26) and letting m tend to infinity gives the inequality

$$0 \geq 1 - C - \big(1-(1-\varepsilon)C\big) \left(1 - \frac{1}{2\ln 2} \left(\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3}\right)^{a_{\mathrm{R}}(\infty)}\right) \tag{28}$$

where $a_{\mathrm{R}}(\infty) \triangleq \liminf_{m\to\infty} a_{\mathrm{R}}(m)$. Note that the base of the exponent in the RHS of this inequality does not exceed unity, i.e.,

$$\begin{aligned} \frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3} &\leq \frac{(p_1+p_2)^2}{p_1+p_2} + \frac{(p_0+p_3)^2}{p_0+p_3} \\ &= p_0+p_1+p_2+p_3 \\ &= 1. \end{aligned}$$

Therefore, the inequality in (28) yields the following lower bound on the asymptotic average right degree:

$$a_{\mathrm{R}}(\infty) \geq K_1' + K_2' \ln\left(\frac{1}{\varepsilon}\right) \tag{29}$$

where

$$\begin{aligned} K_1' &= -\frac{\ln\left(\frac{1}{2\ln 2}\frac{1-C}{C}\right)}{\ln\left(\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3}\right)}, \\ K_2' &= -\frac{1}{\ln\left(\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3}\right)}. \end{aligned} \tag{30}$$

According to Definition 2.2, the density (Δ) of a parity-check matrix is equal to the number of edges in the corresponding bipartite graph normalized per information bit, while the average right degree (a_{R}) is equal to the same number of edges normalized per parity-check node. Since the parity-check matrix H is full rank, then the above scalings of the number of edges in a bipartite graph imply

$$\Delta = \left(\frac{1-R}{R}\right) a_{\mathrm{R}} \tag{31}$$

where R is the rate of a binary linear block code. By our assumption, the asymptotic rate of the sequence of codes $\{\mathcal{C}_m\}$ is equal to a fraction $1-\varepsilon$ of the capacity. Therefore, by combining (29) and (31) with $R=(1-\varepsilon)C$, we obtain a lower bound on the asymptotic parity-check density which gets the form

$$\liminf_{m\to\infty} \Delta_m \geq \frac{K_1 + K_2 \ln\left(\frac{1}{\varepsilon}\right)}{1-\varepsilon}$$

where

$$K_{1,2} = \frac{1-C}{C} \cdot K'_{1,2} \tag{32}$$

and $K'_{1,2}$ are introduced in (30) (note that $1-R \geq 1-C$). This completes the proof of the lower bound in (22) with the coefficients $K_{1,2}$ in (23).

Derivation of the optimization equation (24): The reader is referred to Appendix I-B, where the existence of such a solution is also shown. ■

Discussion: It is required to show that we achieve an improved lower bound on the parity-check density, as compared to the one in [14, Theorem 2.1]. To this end, it suffices to show that

$$\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3} \geq (1-2w)^2 \tag{33}$$

where w is introduced in (2). For a proof of (33), the reader is referred to Appendix I-C. This therefore proves that the new lower bound which is based on a four-level quantization is tighter than the original bound in [14, Theorem 2.1] which corresponds to a two-level quantization of the LLR. A more general proof which relies on the information processing inequality is later presented (see Section III-B) where it shows that the bound is improved by adding quantization levels to the existing ones; this enables to show that the lower bound on the parity-check density which corresponds to 2^d optimized quantization levels (so as to get the tightest bound where the number of quantization levels is fixed) is monotonically increasing w.r.t. d.

Based on Proposition 3.1, we introduce an upper bound on the asymptotic rate of every sequence of binary linear block codes for which reliable communication is achievable. The bound refers to soft-decision ML decoding, and it is therefore valid for any suboptimal decoding algorithm. Hence, the following result also provides an upper bound on the achievable rates of ensembles of LDPC codes under iterative decoding where the transmission takes places over an MBIOS channel. The following bound improves the bounds stated in [1, Theorems 1 and 2]:

Corollary 3.1: ("Four-Level Quantization" Upper Bound on the Asymptotic Achievable Rates of Sequences of Binary Linear Block Codes) Let $\{C_m\}$ be a sequence of binary linear block codes whose codewords are transmitted with equal probability over an MBIOS channel, and suppose that the block length of this sequence of codes tends to infinity as $m \to \infty$. Let $\Gamma_{k,m}$ be the fraction of the parity-check nodes of degree k in an arbitrary representation of the code $\mathcal{C}_m$ by a bipartite graph which corresponds to a full-rank parity-check matrix. Then a necessary condition for this sequence to achieve vanishing bit error probability as $m \to \infty$ is that the asymptotic rate R of this sequence satisfies

$$\begin{aligned}1 - \max\Biggl\{&(1-C)\Biggl[\sum_k \Bigl\{\Gamma_k \sum_{t=0}^{k} \binom{k}{t} (p_1+p_2)^t (p_0+p_3)^{k-t} \\ &\cdot h_2\biggl(\frac{1-\bigl(1-\frac{2p_2}{p_1+p_2}\bigr)^t\bigl(1-\frac{2p_3}{p_0+p_3}\bigr)^{k-t}}{2}\biggr)\Bigr\}\Biggr]^{-1}, \\ &\frac{2(p_2+p_3)}{1-\sum_k \Gamma_k \Bigl(1-2(p_2+p_3)\Bigr)^k}\Biggr\} \geq R\end{aligned} \tag{34}$$

where p_0, p_1, p_2, p_3 are introduced in (9), and Γ_k and R are introduced in (1).

Proof: The first term in the maximization on the LHS of (34) is based on (10) and (25); it follows directly by combining these inequalities, and letting the bit error probability P_b tend to zero. The second term in the maximization on the LHS of (34) follows from the proof of [14, Corollary 3.1] which is based on the erasure decomposition Lemma [12]; in the context considered here, $p = 2(p_2+p_3)$ is the erasure probability of the corresponding BEC (see the upper plot in Fig. 1). ■

B. Extension of the Bounds to 2^d Quantization Levels

Following the method introduced in Section III-A, we commence by deriving a lower bound on the conditional entropy of a transmitted codeword given the received sequence.

Proposition 3.2: Let $\mathcal{C}$ be a binary linear block code of length n and rate R. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ designate the transmitted codeword and received sequence, respectively, when the communication takes place over an MBIOS channel with conditional pdf $p_{Y|X}(\cdot|\cdot)$. For an arbitrary $d \geq 2$ and $0 < l_{2^{d-1}-1} < \ldots < l_1 < l_0 \triangleq \infty$, let us define the set of probabilities $\{p_s\}_{s=0}^{2^d-1}$ as follows:

$$p_s \triangleq \begin{cases} \Pr\{l_{s+1} < \text{LLR}(Y) \leq l_s \mid X = 0\} \\ \text{if } s = 0, \ldots, 2^{d-1} - 2 \\[1ex] \Pr\{0 < \text{LLR}(Y) \leq l_{2^{d-1}-1} \,|\, X = 0\} \\ +\frac{1}{2}\Pr\{\text{LLR}(Y) = 0 \,|\, X = 0\} \\ \text{if } s = 2^{d-1} - 1 \\[1ex] \Pr\{-l_{2^{d-1}-1} \leq \text{LLR}(Y) < 0 \mid X = 0\} \\ +\frac{1}{2}\Pr\{\text{LLR}(Y) = 0 \mid X = 0\} \\ \text{if } s = 2^{d-1} \\[1ex] \Pr\{-l_{2^d-(s+1)} \leq \text{LLR}(Y) < -l_{2^d-s} \,|\, X = 0\} \\ \text{if } s = 2^{d-1} + 1, \ldots, 2^d - 1. \end{cases} \tag{35}$$

For an arbitrary full-rank parity-check matrix of the code $\mathcal{C}$, let Γ_k designate the fraction of the parity-checks involving k variables. Then, the conditional entropy of the transmitted codeword given the received sequence satisfies

$$\begin{aligned} &\frac{H(\mathbf{X}|\mathbf{Y})}{n} \\ &\geq 1 - C - (1 - R) \\ &\quad \cdot \sum_k \Bigg\{ \Gamma_k \sum_{\substack{k_0, \ldots, k_{2^{d-1}-1} \\ \sum_i k_i = k}} \binom{k}{k_0, \ldots, k_{2^{d-1}-1}} \\ &\qquad \cdot \prod_{i=0}^{2^{d-1}-1} (p_i + p_{2^d-1-i})^{k_i} \\ &\qquad \cdot h_2\left(\frac{1}{2}\left[1 - \prod_{i=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-i}}{p_i + p_{2^d-1-i}}\right)^{k_i}\right]\right) \Bigg\}. \end{aligned} \tag{36}$$

Proof: Following the proof of Proposition 3.1, we introduce a new physically degraded channel. It is a memoryless binary-input 2^d-ary output symmetric channel (see Fig. 1 for $d = 2$). To this end, let $l_{2^{d-1}-1} < \ldots < l_1$ be arbitrary positive numbers, and denote $l_0 \triangleq \infty$. The output alphabet of the degraded channel is defined to be GF(2^d) whose elements form the set

$$\left\{ \sum_{j=0}^{d-1} a_j \, \alpha^j \; : \quad (a_0, a_1, \ldots, a_{d-1}) \in \{0, 1\}^d \right\}$$

where α is a primitive element of the Galois field GF(2^d).

For $s = 0, 1, \ldots, 2^{d-1} - 1$, denote the $(d-1)$–bit binary representation of s by $(a_1^{(s)}, \ldots, a_{d-1}^{(s)})$, i.e.,

$$s = \sum_{j=1}^{d-1} a_j^{(s)} \, 2^{j-1}.$$

Let X_i and Y_i designate the random variables referring to the input and output of the original channel $p_{Y|X}$ at time i (where $i = 1, \ldots, n$). As a natural generalization of the channel model in Fig. 1, we introduce a physically degraded channel with 2^d quantization levels of the LLR. The output of this channel at time i, Z_i, is calculated from the output Y_i of the original channel as follows:

$$Z_i = \Phi_i + \Omega_i \alpha. \tag{37}$$

The component Φ_i in (37) depends on the sign of LLR(Y_i); it is set to zero or one, if the LLR is positive or negative, respectively; if LLR$(Y_i) = 0$, then Φ_i is either zero or one with equal probability. The value of Ω_i is calculated based on the absolute value of LLR(Y_i) as follows:

- If $l_{s+1} < |\text{LLR}(Y_i)| \leq l_s$ for some $0 \leq s < 2^{d-1} - 1$, then
$$\Omega_i = \sum_{j=1}^{d-1} a_j^{(s)} \, \alpha^{j-1}. \tag{38}$$
- If $0 \leq |\text{LLR}(Y_i)| \leq l_{2^{d-1}-1}$, then
$$\Omega_i = \sum_{j=1}^{d-1} \alpha^{j-1}. \tag{39}$$

From (35), the transition probabilities of the degraded channel are given by

$$\begin{aligned} p_s &= \Pr(Z = \sum_{j=1}^{d-1} a_j^{(s)} \, \alpha^j \mid X = 0) \\ &= \Pr(Z = 1 + \sum_{j=1}^{d-1} a_j^{(s)} \, \alpha^j \mid X = 1) \\ p_{2^d-1-s} &= \Pr(Z = 1 + \sum_{j=1}^{d-1} a_j^{(s)} \, \alpha^j \mid X = 0) \\ &= \Pr(Z = \sum_{j=1}^{d-1} a_j^{(s)} \, \alpha^j \mid X = 1) \end{aligned} \tag{40}$$

where $s = 0, 1, \ldots, 2^{d-1} - 1$. The symmetry in these equalities holds since the channel is MBIOS.

Equations (11)–(15) hold also for the case of 2^d-level quantization. Thus, we will calculate the entropy of the random variable Z, and an upper bound on the entropy of the random vector $\mathbf{Z}$. This will finally provide the lower bound in (36).

Analogously to the proof of Proposition 3.1, the degraded channel is additive over GF(2^d). We denote the additive noise by

$$N_i = \Theta_i + \Omega_i \alpha. \tag{41}$$

Note that since the code $\mathcal{C}$ is binary, then $\Phi_i = \Theta_i + X_i$, and the value of Ω_i stays the same in (37) and (41). Let $\boldsymbol{\Theta} \triangleq (\Theta_1, \ldots, \Theta_n)$ and $\boldsymbol{\Omega} \triangleq (\Omega_1, \ldots, \Omega_n)$. Due to the symmetry of the communication channel, it follows that $\boldsymbol{\Theta}$ and $\boldsymbol{\Omega}$ are statistically independent of the transmitted codeword $\mathbf{X}$. This gives

$$H(Z) = H(\Omega) + 1 \tag{42}$$

which follows from the same argument which validates (17).

We now derive an upper bound on the entropy of the random vector $\mathbf{Z}$. From the same chain of equalities leading to (18), it follows that

$$H(\mathbf{Z}) = nH(\Omega) + H(\boldsymbol{\Phi} \mid \boldsymbol{\Omega}) \tag{43}$$

where $\boldsymbol{\Phi} \triangleq (\Phi_1, \ldots, \Phi_n)$. As in the proof of Proposition 3.1, we define the syndrome as $\mathbf{S} \triangleq \boldsymbol{\Phi} H^T$ where H is a full-rank parity-check matrix of the code $\mathcal{C}$. As before, the calculation of the syndrome $\mathbf{S}$ only takes into account the Φ-components of the vector $\mathbf{Z}$. Since $\mathbf{X}H^T = 0$ and $\boldsymbol{\Phi} = \mathbf{X} + \boldsymbol{\Theta}$, then $\mathbf{S} = \boldsymbol{\Theta} H^T$ which is independent of the transmitted codeword. In parallel to (19), we obtain

$$H(\boldsymbol{\Phi} \mid \boldsymbol{\Omega}) \leq nR + H(\mathbf{S} \mid \boldsymbol{\Omega}). \tag{44}$$

Consider a parity-check equation which involves k variables, and let $\{i_1, \ldots, i_k\}$ be the set of indices of the variables involved in this parity-check equation. The component of the syndrome $\mathbf{S}$ which refers to this parity-check equation is zero if and only if the binary sub-vector $(\Theta_{i_1}, \ldots, \Theta_{i_k})$ has an even Hamming weight.

Lemma 3.3: Given that $(\Omega_{i_1}, \ldots, \Omega_{i_k})$ has k_s elements equal to $\sum_{j=1}^{d-1} a_j^{(s)} \; \alpha^{j-1}$ $(s = 0, \ldots, 2^{d-1} - 1)$, the probability that the corresponding component of the syndrome S_l is equal to 1 is given by

$$\frac{1}{2}\left[1 - \prod_{s=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-s}}{p_s + p_{2^d-1-s}}\right)^{k_s}\right].$$

Proof: From the probabilities which are associated with the quantized values of the LLR in (40), it follows that for $0 \leq s \leq 2^{d-1} - 1$

$$\Pr\Bigl(\Theta_i = 1 \,|\, \Omega_i = \sum_{j=1}^{d-1} a_j^{(s)} \alpha^{j-1}\Bigr) = \frac{p_{2^d-1-s}}{p_s + p_{2^d-1-s}}.$$

Since there are k_s indices i in the set $\{i_1, \ldots, i_k\}$ for which $\Omega_i = \sum_{j=1}^{d-1} a_j^{(s)} \; \alpha^{j-1}$, the lemma follows from [3, Lemma 4.1]. ■

Based on Lemma 3.3 and the discussion above, it follows that for any vector $\underline{\omega} = (\omega_1, \ldots, \omega_k)$ which has k_s elements equal to $\sum_{j=1}^{d-1} a_j^{(s)} \; \alpha^{j-1}$ (where $s = 0, \ldots, 2^{d-1} - 1$)

$$\begin{aligned} & H\bigl(S_i \,|\, (\Omega_{i_1}, \ldots, \Omega_{i_k}) = \underline{\omega}\bigr) \\ & = h_2\left(\frac{1}{2}\left[1 - \prod_{s=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-s}}{p_s + p_{2^d-1-s}}\right)^{k_s}\right]\right). \end{aligned} \quad (45)$$

For a component S_i $(1 \leq i \leq n(1-R))$ of the syndrome $\mathbf{S}$ which refers to a parity-check equation involving k variables

$$\begin{aligned} & H(S_i \,|\, \mathbf{\Omega}) \\ & = H(S_i \,|\, \Omega_{i_1}, \ldots, \Omega_{i_k}) \\ & = \sum_{\underline{\omega}} \Pr\bigl((\Omega_{i_1}, \ldots, \Omega_{i_k}) = \underline{\omega}\bigr) \; H\bigl(S_i | (\Omega_{i_1}, \ldots, \Omega_{i_k}) = \underline{\omega}\bigr) \\ & = \sum_{\substack{k_0, \ldots, k_{2^{d-1}-1} \\ \sum_s k_s = k}} \left\{ \binom{k}{k_0, \ldots, k_{2^{d-1}-1}} \prod_{s=0}^{2^{d-1}-1} (p_s + p_{2^d-1-s})^{k_s} \right. \\ & \qquad \left. \cdot \, h_2\left(\frac{1}{2}\left[1 - \prod_{s=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-s}}{p_s + p_{2^d-1-s}}\right)^{k_s}\right]\right) \right\} \end{aligned}$$

where the last equality follows since there are $\binom{k}{k_0, \ldots, k_{2^{d-1}-1}}$ vectors $\underline{\omega}$ which have k_s elements of the type $\sum_{j=1}^{d-1} a_j^{(s)} \; \alpha^{j-1}$ for $s \in \{0, \ldots, 2^{d-1} - 1\}$, and it also follows from the statistical independence of the components of the vector $\mathbf{\Omega}$, and from Eqs. (40) and (45). The number of parity-check equations involving k variables is $n(1-R)\Gamma_k$, hence

$$\begin{aligned} & H(\mathbf{S} \,|\, \mathbf{\Omega}) \\ & \leq \sum_{i=1}^{n(1-R)} H(S_i \,|\, \mathbf{\Omega}) \\ & = n(1-R) \\ & \quad \cdot \sum_k \left\{ \Gamma_k \sum_{\substack{k_0, \ldots, k_{2^{d-1}-1} \\ \sum_s k_s = k}} \binom{k}{k_0, \ldots, k_{2^{d-1}-1}} \right. \end{aligned} \quad (46)$$

$$\cdot \prod_{s=0}^{2^{d-1}-1} (p_s + p_{2^d-1-s})^{k_s}$$
$$\cdot h_2\left(\frac{1}{2}\left[1 - \prod_{s=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-s}}{p_s + p_{2^d-1-s}}\right)^{k_s}\right]\right)\Bigg\}.$$

By combining (43)–(46), an upper bound on the entropy of the random vector $\mathbf{Z}$ follows:

$$\begin{aligned} & H(\mathbf{Z}) \\ & \leq nR + nH(\Omega) + n(1-R) \qquad (47) \\ & \cdot \sum_k \Bigg\{ \Gamma_k \sum_{\substack{k_0,\dots,k_{2^{d-1}-1} \\ \sum_s k_s = k}} \binom{k}{k_0, \dots, k_{2^{d-1}-1}} \\ & \quad \cdot \prod_{s=0}^{2^{d-1}-1} (p_s + p_{2^d-1-s})^{k_s} \\ & \quad \cdot h_2\left(\frac{1}{2}\left[1 - \prod_{s=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-s}}{p_s + p_{2^d-1-s}}\right)^{k_s}\right]\right)\Bigg\}. \end{aligned}$$

The substitution of (42) and (47) in (15) finally provides the lower bound on the conditional entropy $H(\mathbf{X} \mid \mathbf{Y})$ in (36). ■

Discussion: In the proof of Proposition 3.2, the upper bound on the entropy of the output $\mathbf{Z}$ is of the form

$$H(\mathbf{Z}) \leq nH(\Omega) + nR + \sum_{i=1}^{n(1-R)} H(S_i|\mathbf{\Omega})$$

which follows directly from (43), (44) and (46). Here, Ω is defined according to the conditions stated in (38) and (39), and S_i is the i^{th} component of the syndrome $\mathbf{S}$. Substituting (42) and the above inequality in (15) yields

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq 1 - C - \frac{1}{n} \sum_{i=1}^{n(1-R)} H(S_i|\mathbf{\Omega})$$

and the lower bound in (36) is in fact an explicit expression for the above inequality. The calculation of the lower bound in the RHS of (36) becomes more complex as the value of d is increased. However, when the quantization levels $l_1, \dots, l_{2^{d-1}-1}$ are set to maximize the lower bound in the RHS of (36), we show that the bound is monotonically increasing with d.

To this end, let $d \geq 2$ be an arbitrary integer, $(l_1^{(d)}, \dots, l_{2^{d-1}-1}^{(d)})$ with their symmetric values around zero be the optimal choice of 2^d quantization levels, and denote the random variable Ω for this setting by $\Omega^{(d)}$ (see (38) and (39)). Consider any set of 2^{d+1} quantization levels $(l_1^{(d+1)}, \dots, l_{2^d-1}^{(d+1)})$ with their symmetric values around zero such that $l_{2i}^{(d+1)} = l_i^{(d)}$ for $i = 1, \dots, 2^{d-1} - 1$. Denote the random variable Ω for this choice of quantization levels by $\Omega^{(d+1)}$. Clearly, since the former set of 2^d quantization levels is a subset of the latter set of 2^{d+1} levels, then $\Omega^{(d)}$ can be calculated from $\Omega^{(d+1)}$. Let $\mathbf{\Omega}^{(k)} \triangleq (\Omega_1^{(k)}, \dots, \Omega_n^{(k)})$ for $k = d$ and $d + 1$. By the information processing inequality

$$\begin{aligned} & 1 - C - \frac{1}{n} \sum_{i=1}^{n(1-R)} H(S_i|\mathbf{\Omega}^{(d)}) \\ & \leq 1 - C - \frac{1}{n} \sum_{i=1}^{n(1-R)} H(S_i|\mathbf{\Omega}^{(d+1)}). \end{aligned}$$

Therefore, the (possibly sub-optimal) set of 2^{d+1} quantization levels $(l_1^{(d+1)}, \dots, l_{2^d-1}^{(d+1)})$ with their symmetric values around zero provides a tighter lower bound than the *optimal* choice of 2^d quantization levels. Hence, this proves that the lower bound is monotonically increasing with the number of quantization levels when these levels are set optimally.

Theorem 3.2: ("2^d-Level Quantization" Lower Bound on the Asymptotic Parity-Check Density of Binary Linear Block Codes) Let $\{C_m\}$ be a sequence of binary linear block codes achieving a fraction $1-\varepsilon$ of the capacity of an MBIOS channel with vanishing bit error probability. Let H_m be an arbitrary *full-rank* parity-check matrix of the code $\mathcal{C}_m$, and denote its density by Δ_m. Then, the asymptotic density satisfies

$$\liminf_{m\to\infty} \Delta_m > \frac{K_1 + K_2 \ln \frac{1}{\varepsilon}}{1-\varepsilon} \tag{48}$$

where

$$\begin{aligned} K_1 &= K_2 \ \ln\left(\frac{1}{2\ln 2} \ \frac{1-C}{C}\right) , \\ K_2 &= -\frac{1-C}{C \ \ln\left(\displaystyle\sum_{i=0}^{2^{d-1}-1} \frac{(p_i - p_{2^d-1-i})^2}{p_i + p_{2^d-1-i}}\right)} . \end{aligned} \tag{49}$$

Here, $d \geq 2$ is an arbitrary integer and the probabilities $\{p_i\}$ are introduced in (35) in terms of $l_1 > \ldots > l_{2^{d-1}-1} \in \mathbb{R}^+$. The optimal vector of quantization levels $(l_1, \ldots, l_{2^{d-1}-1})$ is given implicitly by solving the set of $2^{d-1}-1$ equations

$$\frac{p_{2^d-1-i}^2 + e^{-l_i} p_i^2}{(p_i + p_{2^d-1-i})^2} = \frac{p_{2^d-i}^2 + e^{-l_i} p_{i-1}^2}{(p_{i-1} + p_{2^d-i})^2} \ , \ \ i = 1, \ldots, 2^{d-1}-1 \tag{50}$$

where such a solution always exists.[3]

Proof: For an arbitrary sequence of binary linear block codes $\{\mathcal{C}_m\}$ which achieves a fraction $1-\varepsilon$ to capacity with vanishing bit error probability, we get from Lemma 3.1 that

$$\lim_{m\to\infty} \frac{H(\mathbf{X}_m \mid \mathbf{Y}_m)}{n_m} = 0$$

where $\mathbf{X}_m$ and $\mathbf{Y}_m$ designate the transmitted codeword in the code $\mathcal{C}_m$ and the received sequence, respectively, and n_m designates the block length of the code $\mathcal{C}_m$. From Proposition 3.2, we obtain

$$\begin{aligned} &\frac{H(\mathbf{X}_m|\mathbf{Y}_m)}{n_m} \\ &\geq 1 - C - (1-R_m) \\ &\quad \cdot \sum_k \Bigg\{ \Gamma_{k,m} \sum_{\substack{k_0,\ldots,k_{2^{d-1}-1} \\ \sum_s k_s = k}} \binom{k}{k_0, \ldots, k_{2^{d-1}-1}} \\ &\qquad\qquad \cdot \prod_{s=0}^{2^{d-1}-1} (p_s + p_{2^d-1-s})^{k_s} \\ &\qquad\qquad \cdot h_2\left(\frac{1}{2}\left[1 - \prod_{s=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-s}}{p_s + p_{2^d-1-s}}\right)^{k_s}\right]\right) \Bigg\} \end{aligned}$$

[3] See the footnote to Theorem 3.1 on p. 10.

where $\Gamma_{k,m}$ designates the fraction of parity-check equations in a parity-check matrix H_m which involve k variables. The upper bound on the binary entropy function h_2 in Lemma 3.2 gives

$$\begin{aligned}
&\frac{H(\mathbf{X}_m|\mathbf{Y}_m)}{n_m} \\
&\geq 1 - C - \big(1 - R_m\big) \qquad (51)\\
&\quad \cdot \sum_k \Bigg\{ \Gamma_{k,m} \sum_{\substack{k_0,\ldots,k_{2^{d-1}-1} \\ \sum_s k_s = k}} \binom{k}{k_0,\ldots,k_{2^{d-1}-1}} \\
&\qquad \cdot \prod_{s=0}^{2^{d-1}-1} (p_s + p_{2^d-1-s})^{k_s} \\
&\qquad \cdot \left[1 - \frac{1}{2\ln 2} \prod_{s=0}^{2^{d-1}-1} \left(\frac{p_s - p_{2^d-1-s}}{p_s + p_{2^d-1-s}} \right)^{2k_s} \right] \Bigg\}.
\end{aligned}$$

Since $\sum_k \Gamma_{k,m} = 1$ and $\sum_{s=0}^{2^d-1} p_s = 1$, we get

$$\begin{aligned}
&1 - \frac{1}{2\ln 2} \sum_k \Bigg\{ \Gamma_{k,m} \sum_{\substack{k_0,\ldots,k_{2^{d-1}-1} \\ \sum_s k_s = k}} \binom{k}{k_0,\ldots,k_{2^{d-1}-1}} \\
&\qquad\qquad \cdot \prod_{s=0}^{2^{d-1}-1} \left(\frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}} \right)^{k_s} \Bigg\} \\
&= 1 - \frac{1}{2\ln 2} \sum_k \left\{ \Gamma_{k,m} \left(\sum_{s=0}^{2^{d-1}-1} \frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}} \right)^k \right\} \\
&\leq 1 - \frac{1}{2\ln 2} \left(\sum_{s=0}^{2^{d-1}-1} \frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}} \right)^{a_{\mathrm{R}}(m)} \qquad (52)
\end{aligned}$$

where $a_{\mathrm{R}}(m) \triangleq \sum_k k\Gamma_{k,m}$ designates the average right degree of the bipartite graph which refers to the parity-check matrix H_m, and the last transition follows from Jensen's inequality.

Substituting (52) into the RHS of (51) and letting m tend to infinity gives the inequality

$$\begin{aligned}
&0 \geq 1 - C - \big(1 - (1-\varepsilon)C\big) \\
&\quad \cdot \left[1 - \frac{1}{2\ln 2} \left(\sum_{s=0}^{2^{d-1}-1} \frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}} \right)^{a_{\mathrm{R}}(\infty)} \right] \qquad (53)
\end{aligned}$$

where $a_{\mathrm{R}}(\infty) \triangleq \liminf_{m\to\infty} a_{\mathrm{R}}(m)$. Note that the validity of (53) follows since the base of the exponent in this inequality does not exceed unity, i.e.,

$$\begin{aligned}
&\sum_{s=0}^{2^{d-1}-1} \frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}} \\
&\leq \sum_{s=0}^{2^{d-1}-1} \frac{(p_s + p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}} \\
&= \sum_{s=0}^{2^{d-1}-1} (p_s + p_{2^d-1-s}) = 1.
\end{aligned}$$

Inequality (53) gives the following lower bound on the asymptotic average right degree:

$$a_{\mathrm{R}} \geq K_1' + K_2' \ln\left(\frac{1}{\varepsilon}\right) \tag{54}$$

where

$$\begin{aligned} K_1' &= \ln\left(\frac{1}{2\ln 2}\,\frac{1-C}{C}\right)\, K_2'\,, \\ K_2' &= -\frac{1}{\ln\left(\displaystyle\sum_{s=0}^{2^{d-1}-1} \frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}}\right)}. \end{aligned}$$

By combining (31) and (54) with the asymptotic rate $R = (1-\varepsilon)C$, we obtain a lower bound on the asymptotic parity-check density which is of the form

$$\liminf_{m\to\infty} \Delta_m \geq \frac{K_1 + K_2 \ln\left(\frac{1}{\varepsilon}\right)}{1-\varepsilon}$$

where

$$K_{1,2} = \frac{1-C}{C} \cdot K_{1,2}'.$$

This completes the proof of the lower bound in (48) and (49). The derivation of the set of optimization equations in (50) follows along the lines of the derivation of (24). In the general case of 2^d quantization levels, it follows from (49) that we need to maximize

$$\sum_{s=0}^{2^{d-1}-1} \frac{(p_s - p_{2^d-1-s})^2}{p_s + p_{2^d-1-s}}\,.$$

To this end, we set to zero all the partial derivatives w.r.t. l_s where $s = 1, \ldots, 2^{d-1}-1$. Since from (35) only p_s, p_{s-1}, p_{2^d-s} and p_{2^d-s-1} depend on l_s, then

$$\frac{\partial}{\partial l_s}\left\{\frac{(p_{s-1} - p_{2^d-s})^2}{p_{s-1} + p_{2^d-s}} + \frac{(p_s - p_{2^d-s-1})^2}{p_s + p_{2^d-s-1}}\right\} = 0.$$

We express now the probabilities p_s, p_{s-1}, p_{2^d-s} and p_{2^d-s-1} as integrals of the conditional pdf a of the LLR, and rely on the symmetry property which states that $a(l) = e^l a(-l)$ for $l \in \mathbb{R}$. In a similar manner to the derivation of (24), this gives the set of equations in (50). Their solution provides the quantization levels $l_1, \ldots, l_{2^{d-1}-1}$ (where according to Proposition III-B, the other $2^{d-1}-1$ levels are set to be symmetric w.r.t. zero). ■

Based on the proof of Theorem 3.2, we derive an upper bound on the asymptotic rate of every sequence of binary linear codes for which reliable communication is achievable. The bound refers of soft-decision ML decoding, and it is therefore valid for any sub-optimal decoding algorithm.

Corollary 3.2: ("2^d-Level Quantization" Upper Bound on the Asymptotic Achievable Rates of Sequences of Binary Linear Block Codes) Let $\{C_m\}$ be a sequence of binary linear block codes whose codewords are transmitted with equal probability over an MBIOS channel, and suppose that the block length of this sequence of codes tends to infinity as $m \to \infty$. Let $\Gamma_{k,m}$ be the fraction of the parity-check nodes of degree k in an arbitrary representation of the code $\mathcal{C}_m$ by a bipartite graph which corresponds to a full-rank parity-check matrix. Then a necessary condition for this sequence to achieve vanishing bit error probability as $m \to \infty$ is that the asymptotic rate R of this sequence satisfies

$$
\begin{aligned}
1 - \max \Bigg\{ (1-C) \Bigg[\sum_k \Gamma_k \sum_{\substack{k_0,\ldots,k_{2^{d-1}-1} \\ \sum_i k_i = k}} \binom{k}{k_0,\ldots,k_{2^{d-1}-1}} \\
\cdot \prod_{i=0}^{2^{d-1}-1} (p_i + p_{2^d-1-i})^{k_i} \\
\cdot h_2 \Bigg(\frac{1}{2} \Bigg\{ 1 - \prod_{i=0}^{2^{d-1}-1} \left(1 - \frac{2p_{2^d-1-i}}{p_i + p_{2^d-1-i}} \right)^{k_i} \Bigg\} \Bigg) \Bigg]^{-1}, \\
\frac{2 \displaystyle\sum_{i=2^{d-1}}^{2^d-1} p_i}{1 - \displaystyle\sum_k \Gamma_k \left(1 - 2 \sum_{i=2^{d-1}}^{2^d-1} p_i \right)^k} \Bigg\} \geq R
\end{aligned}
\qquad (55)
$$

where $d \geq 2$ is an integer, the probabilities $\{p_i\}$ are introduced in (35), and Γ_k and R are introduced in (1).

Proof: The concept of the proof here is the same as the one of Corollary 3.1, except that the first term in the maximization on the RHS of (55) relies on (36). The second term in this maximization follows from [14, Corollary 3.1], relying on the erasure decomposition lemma; from (35), the erasure probability of the corresponding BEC is equal to $p = 2 \sum_{i=2^{d-1}}^{2^d-1} p_i$. ■

IV. Approach II: Bounds without Quantization of the LLR

Similarly to the previous section, we derive bounds on the asymptotic achievable rate and the asymptotic parity-check density of an arbitrary sequence of binary, linear block codes transmitted over an MBIOS channel. As in Section III, the derivation of these two bounds is based on a lower bound on the conditional entropy of a transmitted codeword given the received sequence at the output of the channel.

Proposition 4.1: Let $\mathcal{C}$ be a binary linear block code of length n and rate R transmitted over an MBIOS channel. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ designate the transmitted codeword and the received sequence, respectively. For an arbitrary representation of the code $\mathcal{C}$ by a full-rank parity-check matrix, let Γ_k designate the fraction of the parity-check equations of degree k, and $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$ be the degree distribution of the parity-check nodes in the corresponding bipartite graph. Then, the conditional entropy of the transmitted codeword given the received sequence satisfies

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq 1 - C - (1-R) \left(1 - \frac{1}{2 \ln 2} \sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)} \right) \qquad (56)$$

where

$$g_p \triangleq \int_0^{\infty} a(l)(1 + e^{-l}) \tanh^{2p}\left(\frac{l}{2}\right) dl \,, \quad p \in \mathbb{N} \qquad (57)$$

and a denotes the conditional pdf of the LLR given that the channel input is 0.

Proof: We consider a binary linear block code $\mathcal{C}$ of length n and rate R whose transmission takes place over an MBIOS channel. For the continuation of the proof, we move from the mapping of the MBIOS channel $X \rightarrow Y$ to the channel $X \rightarrow \widetilde{Y}$ where $\widetilde{Y}$ represents the LLR of the channel output Y. These channels are equivalent in the sense that $H(X|Y) = H(X|\widetilde{Y})$. The basic idea for showing the equivalence between the original channel and the one which will be introduced shortly is based on the fact that the LLR forms a sufficient statistics of the channel.

For the characterization of the equivalent channel, let the function a designate the conditional pdf of the LLR given that the channel input is 0. We randomly generate an i.i.d. sequence $\{L_i\}_{i=1}^n$ w.r.t. the conditional pdf a, and define

$$\Omega_i \triangleq |L_i|, \quad \Theta_i \triangleq \begin{cases} 0 & \text{if} \quad L_i > 0 \\ 1 & \text{if} \quad L_i < 0 \\ 0 \text{ or } 1 \text{ w.p. } \frac{1}{2} & \text{if} \quad L_i = 0 \end{cases} . \qquad (58)$$

The output of the equivalent channel is defined to be the sequence $\widetilde{\mathbf{Y}} = (\widetilde{Y}_1, \ldots, \widetilde{Y}_n)$ where

$$\widetilde{Y}_i = (\Phi_i, \Omega_i), \quad i = 1, \ldots, n$$

and $\Phi_i = \Theta_i + X_i$ where this addition is modulo-2. The output of this equivalent channel at time i is therefore the pair (Φ_i, Ω_i) where $\Phi_i \in \{0, 1\}$ and $\Omega_i \in \mathbb{R}^+$. This defines the memoryless mapping

$$X \rightarrow \widetilde{Y} \triangleq (\Phi, \Omega)$$

where Φ is a binary random variable which is affected by X, and Ω is a non-negative random variable which is not affected by X. Note that due to the symmetry of the communication channel, the joint distribution of the pair (Φ, Ω) is equal to the one which corresponds to the pair representing the sign and magnitude of LLR(Y). Hence,

$$f_\Omega(\omega) = \begin{cases} a(\omega) + a(-\omega) = (1 + e^{-\omega})\, a(\omega) & \text{if } \omega > 0 \\ a(0) & \text{if } \omega = 0 \end{cases} \tag{59}$$

where we rely on the symmetry property of the pdf a.

Following the lines which lead to (15), we obtain

$$H(\mathbf{X}|\mathbf{Y}) \geq nR - H(\widetilde{\mathbf{Y}}) + nH(\widetilde{Y}) - nC. \tag{60}$$

In order to get a lower bound on $H(\mathbf{X}|\mathbf{Y})$, we calculate the entropy of $\widetilde{Y}$ and also obtain an upper bound on the entropy of $\widetilde{\mathbf{Y}}$. The calculation of the first entropy is direct

$$\begin{aligned} H(\widetilde{Y}) &= H(\Phi, \Omega) \\ &= H(\Omega) + H(\Phi|\Omega) \\ &= H(\Omega) + E_\omega\left[H(\Phi|\Omega = \omega)\right] \\ &= H(\Omega) + 1 \end{aligned} \tag{61}$$

where the last transition is due to the fact that given the absolute value of the LLR, its sign is equally likely to be positive or negative. The entropy $H(\Omega)$ is not expressed explicitly as it will cancel out later.

We now derive an upper bound on $H(\widetilde{\mathbf{Y}})$

$$\begin{aligned} H(\widetilde{\mathbf{Y}}) &= H\big(\Phi_1, \Omega_1, \ldots, \Phi_n, \Omega_n\big) \\ &= H(\Omega_1, \ldots, \Omega_n) + H\big(\Phi_1, \ldots, \Phi_n \mid \Omega_1, \ldots, \Omega_n\big) \\ &= nH(\Omega) + H\big(\Phi_1, \ldots, \Phi_n \mid \Omega_1, \ldots, \Omega_n\big). \end{aligned} \tag{62}$$

Define the syndrome vector

$$\mathbf{S} = (\Phi_1, \ldots, \Phi_n)H^T$$

where H is an arbitrary full-rank parity-check matrix of the binary linear block code $\mathcal{C}$, and let M be the index of the vector $(\widetilde{\Phi}_1, \ldots, \widetilde{\Phi}_n)$ in the coset which corresponds to $\mathbf{S}$. Since each coset has exactly 2^{nR} elements which are equally likely then $H(M) = nR$, and we get

$$\begin{aligned} & H\big((\Phi_1, \ldots, \Phi_n) \,|\, (\Omega_1, \ldots, \Omega_n)\big) \\ &= H(\mathbf{S}, M \,|\, (\Omega_1, \ldots, \Omega_n)) \\ &\leq H(M) + H\big(\mathbf{S} \,|\, (\Omega_1, \ldots, \Omega_n)\big) \\ &= nR + H\big(\mathbf{S} \,|\, (\Omega_1, \ldots, \Omega_n)\big) \\ &\leq nR + \sum_{j=1}^{n(1-R)} H\big(S_j \,|\, (\Omega_1, \ldots, \Omega_n)\big)\,. \end{aligned} \tag{63}$$

Since $\mathbf{X}H^T = \mathbf{0}$ for any codeword $\mathbf{X}$, then

$$\mathbf{S} = (\Theta_1, \ldots, \Theta_n)H^T$$

which is independent of the transmitted codeword. Consider the j^{th} parity-check equation, and assume that it involves k variables whose indices are $i_1, \ldots, i_k$. Then, the component S_j of the syndrome is equal to 1 if and only if there is an odd number of ones in the random vector $(\Theta_{i_1}, \ldots, \Theta_{i_k})$.

Lemma 4.1: If the j-th component of the syndrome $\mathbf{S}$ involves k variables whose indices are $i_1, i_2, \ldots, i_k$, then

$$\begin{aligned}&\Pr\big(S_j = 1 \,|\, (\Omega_{i_1}, \ldots, \Omega_{i_k}) = (\alpha_1, \ldots, \alpha_k)\big)\\&= \frac{1}{2}\left[1 - \prod_{m=1}^{k} \tanh\left(\frac{\alpha_m}{2}\right)\right] .\end{aligned} \tag{64}$$

Proof: Due to the symmetry of the channel

$$\begin{aligned}P_m &\triangleq \Pr(\Theta_{i_m} = 1 \mid \Omega_{i_m} = \alpha_m)\\&= \frac{a(-\alpha_m)}{a(\alpha_m) + a(-\alpha_m)}\\&= \frac{1}{1 + e^{\alpha_m}} .\end{aligned}$$

Substituting this result in [3, Lemma 4.1] gives

$$\begin{aligned}&\Pr\big(S_j = 1 \,|\, (\Omega_{i_1}, \ldots, \Omega_{i_k}) = (\alpha_1, \ldots, \alpha_k)\big)\\&= \frac{1}{2}\left[1 - \prod_{m=1}^{k} (1 - 2P_m)\right]\\&= \frac{1}{2}\left[1 - \prod_{m=1}^{k} \tanh\left(\frac{\alpha_m}{2}\right)\right], \quad m = 1, \ldots, k.\end{aligned}$$

■

We therefore obtain from Lemma 4.1 that

$$\begin{aligned}&H\big(S_j | (\Omega_{i_1}, \ldots, \Omega_{i_k}) = (\alpha_1, \ldots, \alpha_k)\big)\\&= h_2\left(\frac{1}{2}\left[1 - \prod_{m=1}^{k} \tanh\left(\frac{\alpha_m}{2}\right)\right]\right)\end{aligned}$$

and by taking the statistical expectation over the k random variables $\Omega_{i_1}, \ldots, \Omega_{i_k}$, we get

$$\begin{aligned}&H\big(S_j | \Omega_{i_1}, \ldots, \Omega_{i_k}\big)\\&= \int_0^\infty \ldots \int_0^\infty h_2\left(\frac{1}{2}\left[1 - \prod_{m=1}^{k} \tanh\left(\frac{\alpha_m}{2}\right)\right]\right)\\&\qquad \cdot \prod_{m=1}^{k} f_\Omega(\alpha_m)\, d\alpha_1 d\alpha_2 \ldots d\alpha_k\\&= 1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \Bigg\{ \frac{1}{p(2p-1)}\\&\qquad \cdot \left(\int_0^\infty f_\Omega(\alpha)\, \tanh^{2p}\left(\frac{\alpha}{2}\right)\, d\alpha\right)^k \Bigg\}\end{aligned} \tag{65}$$

where the equality in the last transition is proved in Appendix II-B. Hence, since $n(1-R)\Gamma_k$ designates the number of parity-check equations of degree k, then

$$\begin{aligned}
&\sum_{j=1}^{n(1-R)} H\big(S_j|\Omega_1,\ldots,\Omega_n\big)\\
&= n(1-R)\Bigg[1-\frac{1}{2\ln 2}\sum_k\Bigg\{\Gamma_k\sum_{p=1}^{\infty}\frac{1}{p(2p-1)}\\
&\qquad\qquad\cdot\left(\int_0^{\infty} f_\Omega(\alpha)\,\tanh^{2p}\left(\frac{\alpha}{2}\right)\,d\alpha\right)^k\Bigg\}\Bigg]\\
&= n(1-R)\left[1-\frac{1}{2\ln 2}\sum_k\left\{\Gamma_k\sum_{p=1}^{\infty}\frac{g_p^k}{p(2p-1)}\right\}\right] \qquad (66)
\end{aligned}$$

where the last equality follows from (57) and (59). By combining (62), (63) and (66), we get the following upper bound on $H(\widetilde{\mathbf{Y}})$:

$$\begin{aligned}
&H(\widetilde{\mathbf{Y}})\\
&\leq nH(\Omega)+nR+n(1-R)\\
&\quad\cdot\left[1-\frac{1}{2\ln 2}\sum_k\left\{\Gamma_k\sum_{p=1}^{\infty}\frac{g_p^k}{p(2p-1)}\right\}\right]\\
&= nH(\Omega)+nR+n(1-R)\\
&\quad\cdot\left[1-\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\left\{\frac{1}{p(2p-1)}\sum_k\Gamma_k g_p^k\right\}\right]\\
&= nH(\Omega)+nR+n(1-R)\left(1-\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{\Gamma(g_p)}{p(2p-1)}\right). \qquad (67)
\end{aligned}$$

Finally, the equality in (61) and the upper bound on $H(\widetilde{\mathbf{Y}})$ given in (67) are substituted in the RHS of (60). This provides the lower bound on the conditional entropy $H(\mathbf{X}|\mathbf{Y})$ given in (56), and completes the proof of this proposition. ■

Remark 4.1: For the particular case of a BEC with erasure probability p, the capacity is $C = 1-p$ bits per cannel use. The conditional pdf of the LLR, given that the 0 is transmitted, is equal to

$$a(l) = p\delta_0(l) + (1-p)\delta_\infty(l)$$

where the function δ_a designates the Dirac Delta function at the point a, i.e., $\delta_a(x) \triangleq \delta(x-a)$. We obtain from (57) that $g_m = 1-p$ for all $m \in \mathbb{N}$, so (56) gives

$$\begin{aligned}
\frac{H(\mathbf{X}|\mathbf{Y})}{n} &\geq p-(1-R)\left[1-\frac{1}{2\ln 2}\sum_{m=1}^{\infty}\frac{\Gamma(1-p)}{m(2m-1)}\right]\\
&= p-(1-R)\Big[1-\Gamma(1-p)\Big] \qquad (68)
\end{aligned}$$

where the last transition follows from the equality

$$\sum_{m=1}^{\infty}\frac{1}{2m(2m-1)} = \ln 2\,.$$

The lower bound on the conditional entropy for the BEC, as given in (68), coincides with the result proved in [14, Eqs. (33) and (34)]. The result there was obtained by the derivation of an upper bound on the rank of H_{E} which is a sub-matrix of H whose columns correspond to the variables erased by the BEC.

Discussion: The proof of Proposition 4.1 relies on the analysis of an equivalent channel rather than a degraded (quantized) channel. We therefore expect the lower bound in the RHS of (56) to be tighter than the one in the RHS of (36). By following the derivation in (60)–(63), one gets

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq 1 - C - \frac{1}{n} \sum_{j=1}^{n(1-R)} H(S_j|\Omega_1, \ldots, \Omega_n) \tag{69}$$

where the random variables $\Omega_1, \ldots, \Omega_n$ are defined in (58) and S_j is the j-th component of the syndrome. The lower bound in (56) is in fact an explicit expression for the above inequality where the side information Ω is the absolute value of the LLR without quantization. From the discussion following Proposition 3.2, the bound in (36) is of the same form, except that the side information $\Omega_1, \ldots, \Omega_n$ is a *quantized version* of the absolute value of the LLR. Hence, from the information processing inequality, it follows that indeed (56) is a tighter lower bound on the conditional entropy than (36) for any number of quantization levels.

Theorem 4.1: ("Un-Quantized" Lower Bound on the Asymptotic Parity-Check Density of Binary Linear Block Codes) Let $\{C_m\}$ be a sequence of binary linear block codes achieving a fraction $1-\varepsilon$ of the capacity C of an MBIOS channel with vanishing bit error probability. Let H_m be an arbitrary *full-rank* parity-check matrix of the code $\mathcal{C}_m$, and denote its density by Δ_m. Then, the asymptotic density satisfies

$$\liminf_{m\to\infty} \Delta_m \geq \frac{K_1 + K_2 \ln \frac{1}{\varepsilon}}{1-\varepsilon} \tag{70}$$

where

$$K_1 = K_2 \ \ln\left(\frac{\xi\,(1-C)}{C}\right) \ , \quad K_2 = \frac{1-C}{C} \, \frac{1}{\ln\left(\frac{1}{g_1}\right)} \tag{71}$$

g_1 is introduced in (57), and

$$\xi \triangleq \begin{cases} 1 & \text{for a BEC} \\ \frac{1}{2\ln 2} & \text{otherwise} \end{cases} . \tag{72}$$

Proof: From the lower bound on $\frac{H(\mathbf{X}\mid\mathbf{Y})}{n}$ in Eq. (56) and Lemma 3.1 (see p. 11), we obtain that if $\{\mathcal{C}_m\}$ is a sequence of binary linear block codes which achieves a fraction $1-\varepsilon$ of the channel capacity with vanishing bit error probability, then

$$1 - C - \big(1-(1-\varepsilon)C\big) \left(1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)}\right) \leq 0$$

where $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$ is the asymptotic right degree distribution from the node perspective which corresponds to the sequence of parity-check matrices $\{H_m\}$. Since $\sum_k k\Gamma_k = a_{\mathrm{R}}$ is the average right degree, then from the convexity of the exponential function, we obtain by invoking Jensen's inequality that $\Gamma(x) \geq x^{a_{\mathrm{R}}}$ for all $x \geq 0$, and therefore

$$1 - C - \big(1-(1-\varepsilon)C\big) \left[1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{{g_p}^{a_{\mathrm{R}}}}{p(2p-1)}\right] \leq 0. \tag{73}$$

We derive now two different lower bounds on the infinite sum in the RHS of (73), and compare them later. For the derivation of the lower bound in the first approach, let us define the positive sequence

$$\alpha_p \triangleq \frac{1}{2\ln 2} \, \frac{1}{p(2p-1)} \ , \quad p \in \mathbb{N}. \tag{74}$$

From (99) (see Appendix II-A), the substitution of $x = 0$ in both sides of the equality gives that $\sum_{p=1}^{\infty} \alpha_p = 1$, so the sequence $\{\alpha_p\}$ forms a probability distribution. We therefore obtain that

$$\begin{aligned}
&\frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{g_p^{a_{\mathrm{R}}}}{p(2p-1)} \\
&\overset{\text{(a)}}{=} \sum_{p=1}^{\infty} \left\{ \alpha_p \left(\int_0^{\infty} a(l)(1+e^{-l}) \tanh^{2p}\left(\frac{l}{2}\right) dl \right)^{a_{\mathrm{R}}} \right\} \\
&\overset{\text{(b)}}{\geq} \left(\int_0^{\infty} a(l)(1+e^{-l}) \sum_{p=1}^{\infty} \alpha_p \tanh^{2p}\left(\frac{l}{2}\right) dl \right)^{a_{\mathrm{R}}} \\
&\overset{\text{(c)}}{=} \left(\int_0^{\infty} a(l)(1+e^{-l}) \left[1 - h_2\left(\frac{1-\tanh\left(\frac{l}{2}\right)}{2} \right) \right] dl \right)^{a_{\mathrm{R}}} \\
&\overset{\text{(d)}}{=} \left(\int_0^{\infty} a(l)(1+e^{-l}) \left[1 - h_2\left(\frac{1}{1+e^{l}} \right) \right] dl \right)^{a_{\mathrm{R}}} \\
&\overset{\text{(e)}}{=} C^{a_R}
\end{aligned} \tag{75}$$

where equality (a) follows from (57) and (74), inequality (b) follows from Jensen's inequality, equality (c) follows from (74) and (99), equality (d) follows from the identity $\tanh(x) = \frac{e^{2x}-1}{e^{2x}+1}$, and equality (e) follows from the relation between the capacity of an MBIOS channel and the pdf of the absolute value of the LLR (see [13, Lemma 4.30]).

For a derivation of an alternative lower bound on the infinite series above, we truncate the infinite sum in the RHS of (73) and take into account only the first term in this series. This gives

$$\frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{g_p^{a_{\mathrm{R}}}}{p(2p-1)} \geq \frac{g_1^{a_{\mathrm{R}}}}{2\ln 2} \tag{76}$$

which follows from (57) since $g_p \geq 0$ for all $p \in \mathbb{N}$.

In order to compare the tightness of the two lower bounds in (75) and (76), we first compare the bases of their exponents (i.e., g_1 and C). To this end, it is easy to verify that

$$\tanh^2\left(\frac{l}{2}\right) \geq 1 - h_2\left(\frac{1}{1+e^l}\right) \quad l \in [0, \infty)$$

with an equality if and only if $l = 0$ or $l \to \infty$. To show this, we start from equality (99), use the inequality $(1-2x)^{2p} \leq (1-2x)^2$ for $p \in \mathbb{N}$ and $0 \leq x \leq 1$, and the equality $\displaystyle\sum_{p=1}^{\infty} \frac{1}{2p(2p-1)} = \ln 2$ to finally get

$$h_2(x) \geq 1 - (1-2x)^2 \, , \quad 0 \leq x \leq 1.$$

Hence, (57) and (75) give $g_1 \geq C$ with equality if and only if the MBIOS channel is a BEC. Therefore, up to the multiplicative constant $\frac{1}{2\ln 2}$, the lower bound in (76) is tighter than the first one in (75). However, for the BEC, (75) is tighter than (76); it provides an improvement by a factor of $2\ln 2 \approx 1.386$.

We will therefore continue the analysis based on the second bound in (76), and then give the potential improvement which follows from the first bound in (75) for a BEC. From (73) and (76), we obtain that

$$1 - C - \big(1 - (1-\varepsilon)C\big)\left(1 - \frac{g_1^{a_{\mathrm{R}}}}{2\ln 2}\right) \leq 0.$$

Since $g_1 \leq 1$ (where equality is achieved for a noiseless channel), then the asymptotic average right degree (a_{R}) satisfies the lower bound

$$a_{\mathrm{R}} \geq \frac{\ln\left(\frac{1}{2\ln 2}\left(1 + \frac{1-C}{\varepsilon C}\right)\right)}{\ln\left(\frac{1}{g_1}\right)}.$$

By dropping the 1 inside the logarithm in the numerator, we obtain that

$$a_{\mathrm{R}} > K_1' + K_2' \ln\left(\frac{1}{\varepsilon}\right) \tag{77}$$

where

$$K_1' = \frac{\ln\left(\frac{1}{2\ln 2}\,\frac{1-C}{C}\right)}{\ln\left(\frac{1}{g_1}\right)}, \quad K_2' = \frac{1}{\ln\left(\frac{1}{g_1}\right)}\,.$$

Finally, since for a full-rank parity-check matrix, the parity-check density and average right degree are directly linked by the equality $\Delta = \left(\frac{1-R}{R}\right) a_{\mathrm{R}}$, one obtains the following lower bound on the asymptotic parity-check density:

$$\begin{aligned}\liminf_{m\to\infty} \Delta_m \quad &> \quad \frac{1-(1-\varepsilon)C}{(1-\varepsilon)C}\left(K_1' + K_2' \ln\left(\frac{1}{\varepsilon}\right)\right) \\ &> \quad \frac{K_1 + K_2 \ln\left(\frac{1}{\varepsilon}\right)}{1-\varepsilon} \end{aligned} \tag{78}$$

where $K_{1,2} \triangleq \frac{1-C}{C}\, K_{1,2}'$. For the BEC, this lower bound can be improved by using the first bound in (75). In this case, $g_1 = C = 1-p$ where p designates the erasure probability of the BEC, so the additive coefficient K_1 in the RHS of (70) is improved to

$$K_1 = \frac{p}{1-p}\,\frac{\ln\left(\frac{p}{1-p}\right)}{\ln\left(\frac{1}{1-p}\right)}\,.$$

This concludes the proof of this theorem. ■

Remark 4.2: For a BEC, the lower bound on the asymptotic parity-check density stated in Theorem 4.1 coincides with the bound for the BEC in [14, Eq. (3)]. This lower bound was demonstrated in [14, Theorem 2.3] to be tight. This is proved by showing that the sequence of right-regular LDPC ensembles of Shokrollahi [17] is optimal in the sense that it achieves (up to a small additive coefficient) the lower bound on the asymptotic parity-check density for the BEC.

For a general MBIOS channel (other than the BEC), we show in the proof above that the preferable logarithmic growth rate of the lower bound on the parity-check density is achieved by using the bound which follows from (76). However, we note that the lower bound on the parity-check density which follows from (75) is *universal* for all MBIOS channels with the same capacity.

Remark 4.3: The lower bound on the parity-check density in Theorem 4.1 is uniformly tighter than the one in [14, Theorem 2.1] (except for the BSC and BEC where they coincide). For a proof of this claim, the reader is referred to Appendix III-A.

Based on the proof of Theorem 4.1, we prove and discuss an upper bound on the asymptotic rate of an arbitrary sequence of binary linear block codes for which reliable communication is achievable. The bound refers to ML decoding, and it is therefore valid for any sub-optimal decoding algorithm. Hence, the following result also provides an upper bound on the achievable rates of ensembles of LDPC codes under iterative decoding, where the transmission takes places over an MBIOS channel.

Corollary 4.1: (Upper Bound on Achievable Rates) Let $\{C_m\}$ be a sequence of binary linear block codes whose codewords are transmitted with equal probability over an MBIOS channel, and assume that the block lengths of these codes tend to infinity as $m \to \infty$. Let $\Gamma_{k,m}$ be the fraction of the parity-check nodes of degree k for arbitrary representations of the codes $\mathcal{C}_m$ by bipartite graphs which correspond to full-rank parity-check matrices, and assume the limit $\Gamma_k \triangleq \lim_{m\to\infty} \Gamma_{k,m}$ exists. Then, in the limit where $m \to \infty$, a necessary condition on the asymptotic achievable rate (R) for obtaining vanishing bit error probability is

$$R \le 1 - \frac{1-C}{1 - \frac{1}{2\ln 2}\displaystyle\sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)}} \tag{79}$$

where $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$, and g_p is given in (57).

Proof: This upper bound on the achievable rate follows immediately from Lemma 3.1 and the lower bound on the conditional entropy in Proposition 4.1. The upper bound on R follows since the bit error probability of the sequence of codes $\{\mathcal{C}_m\}$ vanishes as we let m tend to infinity. ■

Remark 4.4: We note that the upper bound on the achievable rate in the RHS of (79) doesn't involve maximization, in contrast to the bound in the RHS of (55). The second term of the maximization in the latter bound follows from considerations related to the BEC where such an expression is not required in the RHS of (79). The reader is referred to Appendix III-B for a proof of this claim.

Corollary 4.2: (Lower Bounds on the Bit Error Probability of LDPC Codes) Let $\mathcal{C}$ be a binary linear block code of rate R whose transmission takes place over an MBIOS channel with capacity C. For an arbitrary full-rank parity-check matrix H of the code $\mathcal{C}$, let Γ_k designate the fraction of parity-check equations that involve k variables, and $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$ be the right degree distribution from the node perspective which refers to the corresponding bipartite graph of $\mathcal{C}$. Then, under ML decoding (or any other decoding algorithm), the bit error probability (P_b) of the code satisfies

$$h_2(P_\text{b}) \geq 1 - \frac{C}{R} + \frac{1-R}{2R\ln 2} \sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)} \tag{80}$$

where g_p is introduced in (57).

Proof: This follows directly by combining (25) and (56). ■

We now introduce the definition of a normalized parity-check density, as given in [14], and derive an improved lower bound on the bit error probability (as compared to [14, Theorem 2.5]) in terms of this quantity.

Definition 4.1: (Normalized parity-check density [14]) Let $\mathcal{C}$ be a binary linear block code of rate R, which is represented by a parity-check matrix H whose density is Δ. The *normalized density* of H, call it $t = t(H)$, is defined to be $t = \frac{R\Delta}{2-R}$.

In the following, we clarify the motivation for the definition of a normalized parity-check density. Let us assume that $\mathcal{C}$ is a binary linear block code of length n and rate R, and suppose that it can be represented by a bipartite graph which is *cycle-free*. From [14, Lemma 2.1], since this bipartite graph contains $(2-R)n-1$ edges, connecting n variable nodes with $(1-R)n$ parity-check nodes without any cycles, then the parity-check density of such a cycle-free code is $\Delta = \frac{2-R}{R} - \frac{1}{nR}$. Hence, in the limit where we let n tend to infinity, the normalized parity-check density of a cycle-free code tends to 1. For codes which are represented by bipartite graphs with cycles, the normalized parity-check density is above 1. As shown in [14, Corollary 2.5], the number of fundamental cycles in a bipartite graph which represents an arbitrary linear block $\mathcal{C}$ grows linearly with the normalized parity-check density. The normalized parity-check density therefore provides a measure for the number of cycles in bipartite graphs representing linear block codes. It is well known that cycle-free codes are not good in terms of performance, even under ML decoding [18]; hence, good error-correcting codes (e.g., LDPC codes) should be represented by bipartite graphs with cycles. Following [14], a lower bound on the asymptotic normalized parity-check density is expressed in terms of the gap (in rate) to capacity (even under ML decoding); this bound provides a quantitative measure for the number of fundamental cycles of bipartite graphs representing good error correcting codes. In the following, we provide such an improved bound as compared to the bound given in [14, Theorem 2.5]. In the continuation (see Section V-B), the resulting improvement is exemplified.

From Definition 4.1, it follows that the relation between the normalized density of a full-rank parity-check matrix and the corresponding average right degree is $t = \left(\frac{1-R}{2-R}\right) a_\text{R}$ so the normalized parity-check density grows linearly with the average right degree (which is directly linked to the decoding complexity per iteration of LDPC codes under MPI decoding) where the scaling factor depends on the code rate R.

Since $\sum_k k\Gamma_k = a_\text{R}$, then by applying Jensen's inequality to the RHS of (80), we get the following lower bound on the bit error probability:

$$h_2(P_\text{b}) \geq 1 - \frac{C}{R} + \frac{1-R}{2R\ln 2} \sum_{p=1}^{\infty} \frac{g_p^{\frac{(2-R)t}{1-R}}}{p(2p-1)}. \tag{81}$$

This lower bound on the bit error probability is tighter than the bound given in [14, Eq. (23)] because of two reasons: Firstly, by combining inequality (76) with Lemma III.1 (see Appendix III-A), we obtain that

$$\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{g_p^{\frac{(2-R)t}{1-R}}}{p(2p-1)} \geq \frac{(1-2w)^{\frac{2(2-R)t}{1-R}}}{2\ln 2}\ .$$

Secondly, the further improvement in the tightness of the new bound is obtained by dividing the RHS of (81) by R (where $R \leq 1$), as compared to the RHS of [14, Eq. (23)].

The bounds in (80) and (81) become trivial when the RHS of these inequalities are non-positive. Let the (multiplicative) gap to capacity be defined as $\varepsilon \triangleq 1-\frac{R}{C}$. Analysis shows that the bounds in (80) and (81) are useful unless $\varepsilon \geq \varepsilon_0$. For the bound in the RHS of (80), ε_0 gets the form

$$\varepsilon_0 = \frac{(1-C)B}{C(1-B)}\ , \quad B \triangleq \frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{\Gamma(g_p)}{p(2p-1)} \tag{82}$$

and for the bound in the RHS of (81), ε_0 is the unique solution of the equation

$$-\varepsilon_0 C + \frac{1-(1-\varepsilon_0)C}{2\ln 2}\sum_{p=1}^{\infty}\frac{g_p^{\frac{(2-(1-\varepsilon_0)C)t}{1-(1-\varepsilon_0)C}}}{p(2p-1)} = 0. \tag{83}$$

For a proof of (82) and (83), the reader is referred to Appendices III-C and III-D, respectively. Similarly to [14, Eq. (25)], we note that ε_0 in (83) forms a lower bound on the gap to capacity for an arbitrary sequence of binary linear block codes achieving vanishing bit error probability over an MBIOS channel; the bound is expressed in terms of their asymptotic rate R and normalized parity-check density t. It follows from the transition from (80) to (81) that the lower bound on the gap to capacity in (83) is looser as compared to the one given in (82). However, the bound in (83) solely depends on the normalized parity-check density, while the bound in (82) requires full knowledge of the degree distribution for the parity-check nodes.

Discussion: Due to the symmetry property for an arbitrary MBIOS channel, the conditional probability density function of the random variable $T = \Pr(X=1|Y) - \Pr(X=-1|Y)$ satisfies the property (see [16, Proposition 3.1])

$$f_T(t) = f_T(-t)\left(\frac{1+t}{1-t}\right), \quad -1 \leq t \leq 1.$$

The relation $T = \tanh\left(\frac{L}{2}\right)$ forms a one-to-one correspondence between the value of the random variable T and the value of the random variable L for the LLR. This implies that $I(X;L) = I(X;T)$. It is shown in [16, Lemma 3.2] that the moments of T satisfy the property

$$\mathbb{E}[T^{2p}] = \mathbb{E}[T^{2p-1}], \quad \forall\, p \in \mathbb{N}\ .$$

Based on the above equality, it follows (see [16, Proposition 3.3]) that the mutual information between X and L can be expressed in the following form:

$$\begin{aligned} I(X;L) &= I(X;T) \\ &= \int_{-1}^{1} f_T(t)\log_2(1+t)dt \\ &= \frac{1}{\ln 2}\sum_{p=1}^{\infty}\frac{1}{2p(2p-1)}\ \mathbb{E}[T^{2p}] \\ &= \frac{1}{\ln 2}\sum_{p=1}^{\infty}\frac{g_p}{2p(2p-1)} \end{aligned}$$

where the last transition follows from (57) and the symmetry property, giving the equality

$$g_p = \int_{-\infty}^{\infty} a(l)\tanh^{2p}\left(\frac{l}{2}\right)\, dl = \mathbb{E}[T^{2p}].$$

Therefore, the mutual information is expressed as a sum of even moments of T (exactly like the lower bound on the conditional entropy in (56)). Similar properties also appear in [8, Lemma 3]. This gives insight to the reason for being able to express the bound on the entropy in (56) as a sum of one-dimensional integrals with *even* moments of the tangent hyperbolic function.

V. Numerical Results

We present here numerical results for the information-theoretic bounds derived in Sections III and IV. As expected, the new bounds improve the numerical results presented in [1, Section 4] and [14, Section 4]. This improvement is attributed to the fact that, in contrast to [1] and [14], the derivation of the new bounds does not rely on a two-level quantization of the LLR; this quantization converts the arbitrary MBIOS channel (whose output may be continuous) to a BSC, and it therefore loosens the bounds. Throughout this section, we assume that the transmission of the codes is over the binary-input AWGN channel. Note that the ”quantized bounds” rely on quantized values of the LLR as side information used for the derivation of these bounds, but the channel itself is not quantized. Hence, by increasing the number of quantization levels used for these bounds, the corresponding values of the $\frac{E_{\text{b}}}{N_0}$ thresholds under ML decoding get farther from the channel capacity (see Tables I–III).

We note that the statements in Sections II–IV refer to the case where the parity-check matrices are full rank. Though it seems like a feasible requirement for specific linear codes, this poses a problem when considering ensembles of LDPC codes. In the latter case, a parity-check matrix, referring to a randomly chosen bipartite graph with a given pair of degree distributions, may not be full rank (one can even construct LDPC ensembles where the design rate is strictly less than their asymptotic rate as the block length goes to infinity). Considering ensembles of LDPC codes, it follows from the proofs of Propositions 3.1, 3.2 and 4.1 that the statements stay valid with the following modifications: the actual code rate R of a code which is randomly picked from the ensemble is replaced by the design rate (R_{d}) of the ensemble, and $\{\Gamma_k\}$ becomes the degree distribution of the parity-check nodes referring to the original bipartite graph which represents a parity-check matrix, possibly not of full rank. The reason for the validity of the bound with the suggested modification is that instead of bounding the entropy of the syndrome by the sum of the entropies of its $n(1-R)$ independent components, we sum over *all* $n(1-R_{\text{d}})$ components (where since $R_{\text{d}} \leq R$, some of these components are possibly linearly dependent). It follows from the proofs that the entropy of the transmitted codeword $(\mathbf{X})$ cancels out with the entropy of the index M of the received vector in the appropriate coset (see e.g. (19)), regardless of the rank of H. By doing this modification, the bound becomes looser when the asymptotic rate of the codes is strictly above the design rate of the ensemble. In light of this modification, we note that the fraction Γ_k of nodes of degree k is calculated in terms of the degree distribution ρ by equation (6), which gives

$$\Gamma_k = \frac{\rho_k}{k} \frac{1}{\displaystyle\int_0^1 \rho(x)\, dx} \, .$$

Though the bounds in Sections III–IV improve on the bounds in [1], the possible replacement of the code rate with the design rate in case that the parity-check matrices are not full rank was also noted in [1, p. 2439].

Based on [7, Lemma 7] (see Lemma 2.1 on p. 5), it was verified that the design rates of the LDPC ensembles presented in this section are equal with probability 1 to the asymptotic rates of codes chosen uniformly at random from these ensembles. This allows one to consider the Shannon capacity limit for these ensembles by referring to the capacity values of $\frac{E_{\text{b}}}{N_0}$ which correspond to their design rates (see Tables I–III).

A. Thresholds of LDPC Ensembles under ML Decoding

Tables I–III provide bounds on the thresholds of various ensembles of LDPC codes under ML decoding. By comparing these bounds with the exact thresholds under iterative decoding (based on the density evolution (DE) technique), Tables I–III provide a quantitative measure for the sub-optimality of MPI decoding (as compared to ML decoding).

The bounds on the achievable rates derived in [1] and Corollaries 3.1, 3.2 and 4.1 provide lower bounds on the $\frac{E_{\text{b}}}{N_0}$ thresholds under ML decoding. For Gallager’s ensembles of regular LDPC codes, the gap between the thresholds under ML decoding and the exact thresholds under the iterative sum-product decoding algorithm is rather large.

TABLE I

Comparison of thresholds for Gallager's ensembles of regular LDPC codes transmitted over the binary-input AWGN channel. The 2-level lower bound on the threshold of $\frac{E_\mathrm{b}}{N_0}$ refers to ML decoding, and is based on [1, Theorem 1] (see also [14, Table II]). The 4-level, 8-level and un-quantized lower bounds refer to ML decoding, and are based on Corollaries 3.1, 3.2 and 4.1, respectively. The upper bound on the threshold of $\frac{E_\mathrm{b}}{N_0}$ holds under 'typical pairs' decoding [4] (and hence, also under ML decoding). The DE thresholds are based on the density evolution analysis, providing exact thresholds under the iterative sum-product decoding algorithm [11].

LDPC Ensemble	Capacity Limit	Lower Bounds				Upper Bound [4]	DE Threshold
		2-Level	4-Level	8-Level	Un-Quantized		
(3,6)	$+0.187$ dB	$+0.249$ dB	$+0.332$ dB	$+0.361$ dB	$+0.371$ dB	$+0.673$ dB	$+1.110$ dB
(4,6)	-0.495 dB	-0.488 dB	-0.472 dB	-0.463 dB	-0.463 dB	-0.423 dB	$+1.674$ dB
(3,4)	-0.794 dB	-0.761 dB	-0.713 dB	-0.694 dB	-0.687 dB	-0.510 dB	$+1.003$ dB

TABLE II

Comparison of thresholds for rate one-half ensembles of irregular LDPC codes transmitted over the binary-input AWGN channel. The Shannon capacity limit corresponds to $\frac{E_\mathrm{b}}{N_0} = 0.187$ decibels. The 2-level, 4-level, 8-level and un-quantized lower bounds on the threshold refer to ML decoding, and are based on [1, Theorem 2], Corollaries 3.1, 3.2 and 4.1, respectively. The degree distributions of the ensembles and their DE thresholds are based on density evolution analysis under iterative sum-product decoding [11], and are taken from [12, Tables 1 and 2].

$\lambda(x)$	$\rho(x)$	Lower Bounds				DE Threshold
		2-Level	4-Level	8-Level	Un-Quantized	
$0.38354x + 0.04237x^2 + 0.57409x^3$	$0.24123x^4 + 0.75877x^5$	0.269 dB	0.370 dB	0.404 dB	0.417 dB	0.809 dB
$0.23802x + 0.20997x^2 + 0.03492x^3 + 0.12015x^4 + 0.01587x^6 + 0.00480x^{13} + 0.37627x^{14}$	$0.98013x^7 + 0.01987x^8$	0.201 dB	0.226 dB	0.236 dB	0.239 dB	0.335 dB
$0.21991x + 0.23328x^2 + 0.02058x^3 + 0.08543x^5 + 0.06540x^6 + 0.04767x^7 + 0.01912x^8 + 0.08064x^{18} + 0.22798x^{19}$	$0.64854x^7 + 0.34747x^8 + 0.00399x^9$	0.198 dB	0.221 dB	0.229 dB	0.232 dB	0.310 dB
$0.19606x + 0.24039x^2 + 0.00228x^5 + 0.05516x^6 + 0.16602x^7 + 0.04088x^8 + 0.01064x^9 + 0.00221x^{27} + 0.28636x^{29}$	$0.00749x^7 + 0.99101x^8 + 0.00150x^9$	0.194 dB	0.208 dB	0.214 dB	0.216 dB	0.274 dB

For this reason, we also compare the lower bounds on the $\frac{E_\mathrm{b}}{N_0}$ thresholds under ML decoding with upper bounds on the $\frac{E_\mathrm{b}}{N_0}$ thresholds which rely on "typical pairs decoding" [4]; an upper bound on the $\frac{E_\mathrm{b}}{N_0}$ thresholds under an arbitrary sub-optimal decoding algorithm (e.g., "typical pairs decoding") also forms an upper bound on these thresholds under ML decoding. It is shown in Table I that for Gallager's ensembles of regular LDPC codes, the gap between the thresholds under iterative sum-product decoding and ML decoding is rather large (this follows by comparing the columns referring to the DE threshold and the upper bound based on "typical pairs decoding"). This large gap is attributed to the sub-optimality of belief propagation decoding for regular LDPC codes. On the other hand, it is also demonstrated in Table I that the gap between the upper and lower bounds on the thresholds under ML decoding is much smaller. For example, according to the numerical results in Table I, the inherent loss in the asymptotic performance due to the sub-optimality of belief propagation for Gallager's ensemble of $(4, 6)$ regular LDPC codes (whose design rate is $\frac{1}{3}$ bits per channel use) ranges between 2.097 and 2.137 dB.

For carefully chosen ensembles of LDPC codes, it is shown in Tables II and III that the gap between the DE

TABLE III

Comparison of thresholds for rate-$\frac{3}{4}$ ensembles of irregular LDPC codes transmitted over the binary-input AWGN channel. The Shannon capacity limit corresponds to $\frac{E_\mathrm{b}}{N_0} = 1.626$ decibels. The 2-level, 4-level, 8-level and un-quantized lower bounds on the threshold refer to ML decoding, and are based on [1, Theorem 2], Corollaries 3.1, 3.2 and 4.1, respectively. The degree distributions of the ensembles and their DE thresholds are based on density evolution analysis under iterative sum-product decoding [11], and are taken from [19].

$\lambda(x)$	$\rho(x)$	Lower Bounds				DE
		2-Level	4-Level	8-Level	Un-Quantized	Threshold
$0.302468x + 0.319447x^2 + 0.378085x^4$	x^{11}	1.698 dB	1.786 dB	1.815 dB	1.825 dB	2.049 dB
$0.244067x + 0.292375x^2 + 0.463558x^6$	x^{13}	1.664 dB	1.718 dB	1.736 dB	1.742 dB	1.874 dB
$0.205439x + 0.255432x^2 + 0.0751187x^4 + 0.1013440x^5 + 0.3626670x^{11}$	x^{15}	1.647 dB	1.680 dB	1.691 dB	1.695 dB	1.763 dB

thresholds under the sum-product decoding algorithm and the new lower bounds on the $\frac{E_\mathrm{b}}{N_0}$ thresholds under ML decoding is rather small. This indicates that for the degree distributions which are provided by the LDPC optimizer [19], the asymptotic degradation in performance due to the sub-optimality of belief propagation is marginal (it is observed from Tables II and III that for several LDPC ensembles, this degradation in the asymptotic performance is at most in the order of hundredths of a decibel).

The plots in Figure 2 compare different lower bounds on the $\frac{E_\mathrm{b}}{N_0}$-threshold under ML decoding of right-regular LDPC ensembles. The plots refer to a right degree of 6 (upper plot) or 10 (lower plot). The following lower bounds are depicted in these plots: the Shannon capacity limit, the 2-level quantization lower bound in [1, Theorem 1], the 4 and 8-level quantization bounds of the LLR in Section III, and finally, the bound in Section IV where no quantization of the LLR is performed. It can be observed from the two plots in Figure 2 that the range of code rates where there exists a visible improvement with the new lower bounds depends on the degree of the parity-check nodes. In principle, the larger the value of the right-degree is, then the improvement obtained by these bounds is more pronounced starting from a higher rate code rate (e.g., for a right degree of 6 or 10, the improvement obtained by the new ’un-quantized’ bounds is observed for code rates above 0.35 and 0.55 bits per channel use, respectively).

B. Lower Bounds on the Bit Error Probability of LDPC Codes

Corollary 4.2 provides an improved lower bound on the bit error probability of binary linear block codes, as compared to the one given in [14, Theorem 2.5]. The plot of Fig. 3 presents a comparison of these lower bounds for binary linear block codes where the new bound relies on (81) and the previously reported bound by Sason and Urbanke is introduced in [14, Theorem 2.5]. The two bounds are plotted as a function of the normalized density of an arbitrary parity-check matrix (see Definition 4.1). In our setting, the capacity of the channel is $\frac{1}{2}$ bit per channel use, and the bounds are depicted for binary linear block codes whose rate is a fraction $1-\varepsilon$ of the channel capacity. To demonstrate the advantage of the lower bound on the bit error probability in (81) over the lower bound derived in [14, Theorem 2.5], let us assume that one wishes to design a binary LDPC code which achieves a bit-error probability of 10^{-6} at a rate which is 99% of the channel capacity. The curve of the lower bound from [14] for a fractional gap to capacity of $\varepsilon = 0.01$ implies that the normalized density of an arbitrary parity-check matrix which represents the code (see Definition 4.1 on p. 28) should be at least $t_{\min} = 4.33$, while the curve depicting the bound from (81) strengthens this requirement to a normalized density (of each parity-check matrix) of at least $t_{\min} = 5.68$. Translating this into terms of parity-check density, which is also the complexity per iteration under MPI decoding, yields minimal parity-check densities of 13.16 and 17.27, respectively (the minimal parity-check density is given by $\Delta_{\min} = \frac{(2-R)t_{\min}}{R}$). It is reflected from Fig. 3 that as the gap to capacity ε tends to zero, the lower bound on the normalized density of an arbitrary parity-check matrix (t), representing a code which achieves low error probability for a rate of $R = (1-\varepsilon)C$, grows significantly.

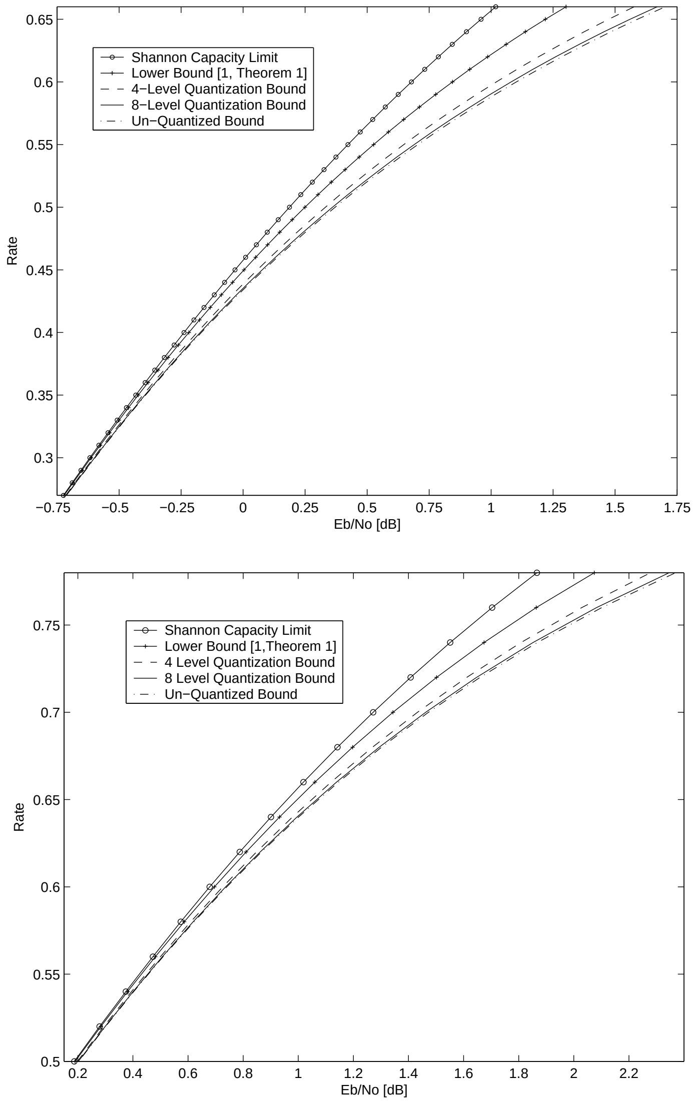

Fig. 2. Comparison between lower bounds on the $\frac{E_{\text{b}}}{N_0}$–thresholds under ML decoding for right-regular LDPC ensembles with $a_{\text{R}} = 6$ (upper plot) and $a_{\text{R}} = 10$ (lower plot). The transmission takes place over the binary-input AWGN channel.

C. Lower Bounds on the Asymptotic Parity-Check Density

The lower bound on the parity-check density derived in Theorem 4.1 enables to assess the tradeoff between asymptotic performance and asymptotic decoding complexity (per iteration) of an MPI decoder. This bound tightens the lower bound on the asymptotic parity-check density derived in [14, Theorem 2.1]. Fig. 4 compares these bounds for codes of rate $\frac{1}{2}$ (upper plot) and $\frac{3}{4}$ (lower plot) where the bounds are plotted as a function of $\frac{E_{\text{b}}}{N_0}$. It can be observed from Fig. 4 that as $\frac{E_{\text{b}}}{N_0}$ increases, the advantage of the bound in Theorem 4.1 over the bound in [14, Theorem 2.1] diminishes. This follows from the fact that as the value of $\frac{E_{\text{b}}}{N_0}$ is increased, the two-level quantization of the LLR used in [1] and [14, Theorem 2.1] better captures the true behavior of the MBIOS channel. It is also reflected in this figure that as ε tends to zero (i.e., when the gap to capacity vanishes), the slope of the bounds becomes very sharp. This is due to the logarithmic behavior of the bounds.

VI. Summary and Outlook

The outstanding performance of low-density parity-check (LDPC) codes under iterative decoding is attributed to the sparseness of the parity-check matrices of these codes. Motivated to consider how sparse parity-check matrices of binary linear block codes can be as a function of their achievable rates and their gap to capacity, we derive in

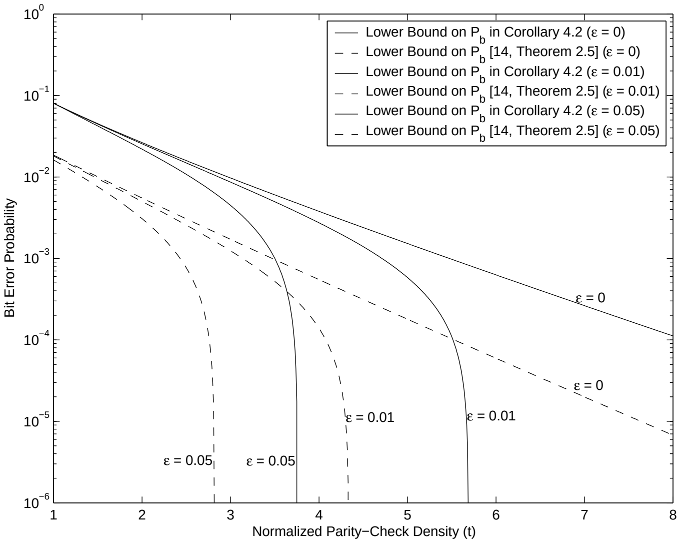

Fig. 3. Lower bounds on the bit error probability for any binary linear block code transmitted over a binary-input AWGN channel whose capacity is $\frac{1}{2}$ bits per channel use. The bounds are depicted in terms of the normalized density of an arbitrary parity-check matrix which represents the code, and the curves correspond to code rates which are a fraction $1 - \varepsilon$ of the channel capacity (for different values of ε). The bounds depicted in dashed lines are based on [14, Theorem 2.5], and the bounds in solid lines are given in Corollary 4.2.

this paper two kinds of bounds. The first category is some improved lower bounds on the asymptotic density of parity-check matrices in terms of the achievable gap (in rate) to capacity, and the second category is upper bounds on the achievable rates of binary linear block codes. These bounds refer to the case where the transmission takes place over memoryless binary-input output-symmetric (MBIOS) channels, and improve the tightness of the bounds given in [1], [14] (as exemplified in Section V). The information-theoretic bounds derived in this paper are valid for *every* sequence of binary linear block codes, in contrast to high probability results which follow from probabilistic tools (e.g., density evolution (DE) analysis under message-passing iterative (MPI) decoding). The bounds hold under maximum-likelihood (ML) decoding, and hence, they also hold under any sub-optimal decoding algorithm.

The bounds in this paper are applied to ensembles of LDPC codes where the significance of these bounds is as follows: Firstly, by comparing the new upper bounds on the achievable rates with thresholds provided by DE analysis, we obtain rigorous bounds on the asymptotic loss in performance of various LDPC ensembles due to the sub-optimality of MPI decoding (as compared to ML decoding). Secondly, the parity-check density of binary linear block codes which are represented by standard bipartite graphs can be interpreted as the complexity per iteration under MPI decoding. Therefore, by tightening the reported lower bound on the asymptotic parity-check density (see [14, Theorem 2.1]), the new bounds provide better insight on the tradeoff between the asymptotic performance and the asymptotic decoding complexity of iteratively decoded LDPC codes. Thirdly, the new lower bound on the bit error probability of binary linear block codes tightens the reported lower bound in [14, Theorem 2.5] and provides a quantitative measure to the number of fundamental cycles in the graph which should exist in terms of the achievable rate (even under ML decoding) and its gap to capacity. It is well known that cycle-free codes have poor performance [18]; hence, the lower bound on the minimal number of fundamental cycles in the graph, as a function of the gap to capacity, strengthens the result in [18].

The derivation of the bounds in Section III was motivated by the desire to generalize the results in [1, Theorems 1 and 2] and [14, Theorem 2.1]. The two-level quantization of the log-likelihood ratio (LLR) which in essence replaces the arbitrary MBIOS channel by a physically degraded binary symmetric channel (BSC), is modified in Section III to a quantized channel which better reflects the statistics of the original channel (though the quantized channel is still physically degraded w.r.t. the original channel). The number of quantization levels at the output of the new channel is an arbitrary integer power of 2. The calculation of the bounds in Section III is subject to an optimization of the quantization levels of the LLR, as to get the tightest bounds within their form. In Section IV, we rely on the conditional pdf of the LLR at the output of the MBIOS channel, and operate on an equivalent channel without quantizing the LLR. This second approach finally leads to bounds which are uniformly tighter than the bounds in

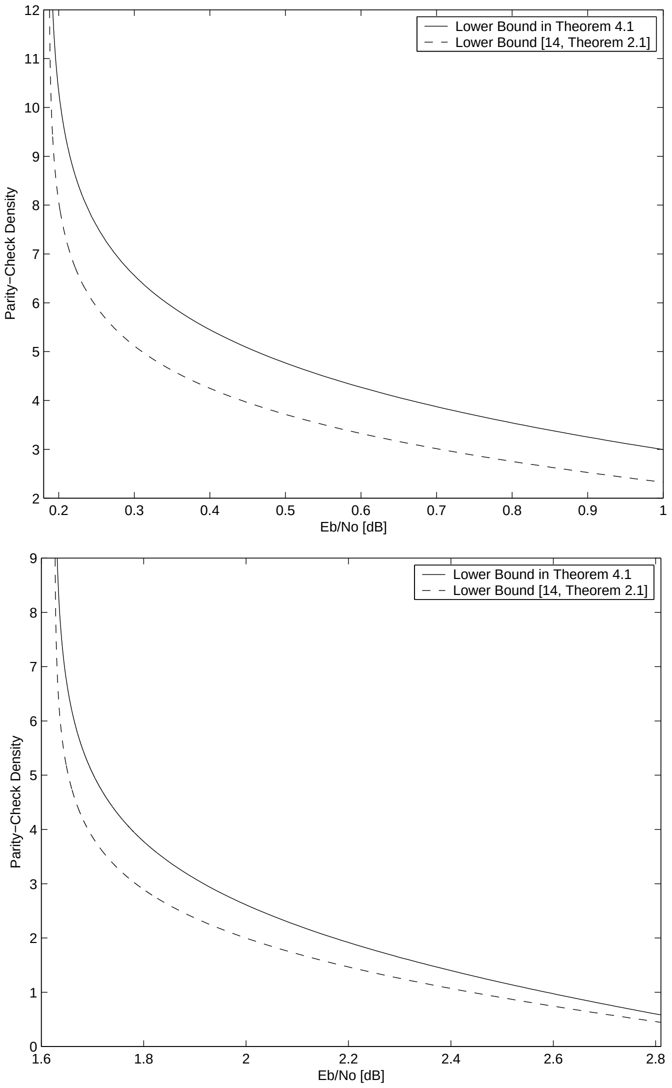

Fig. 4. Comparison between lower bounds on the asymptotic parity-check density of binary linear block codes where the transmission takes place over a binary-input AWGN channel. The dashed line refers to [14, Theorem 2.1], and the solid line refers to Theorem 4.1. The upper and lower plots refer to code rates of $\frac{1}{2}$ and $\frac{3}{4}$, respectively. The Shannon capacity limit for these code rates corresponds to $\frac{E_b}{N_0}$ of 0.187 dB and 1.626 dB, respectively.

Section III. It appears to be even simpler to calculate the un-quantized bounds in Section IV, as their calculation does not involve the solution of any optimization equation related to the quantization levels. The comparison between the quantized and un-quantized bounds gives insight on the effect of the number of quantization levels of the LLR (even if they are chosen optimally) on the achievable rates, as compared to the ideal case where no quantization is done. The results of such a comparison are shown in Tables I–III (see Section V-A), and indicate that the improvement in the tightness of the bounds when more than 8 levels of quantization are used is marginal (in the case that the quantization levels are optimally determined). We also note that practically, the possibility to calculate un-quantized bounds which are uniformly better than the quantized bounds was facilitated due to an efficient transformation of the multi-dimensional integral in Appendix II-B into an infinite series of one-dimensional integrals whose convergence rate is fast.

In [8], a new method for analyzing LDPC and low-density generator-matrix (LDGM) codes under MAP decoding is introduced, based on tools borrowed from statistical physics. Considering ensembles of codes, one gets from [8] high probability results as the block length gets large, but the performance of specific codes of finite length deviates from the average ensemble performance. Since the bounding techniques which rely on statistical physics [8] do not allow for a bound which is valid for every linear block code, it would be interesting to get some theory that unifies the information-theoretic and statistical physics approaches and provides bounds that are tight on average and valid code by code.

For ensembles of LDPC codes, the introduced bounds on the thresholds under ML decoding only depend on the degree distribution of their parity-check nodes and their design rate. For a given parity-check degree distribution (ρ) and design rate (R), the bounds provide an indication on the inherent gap to capacity which is independent of the choice of the left degree distribution λ (as long as the pair of degree distributions (λ, ρ) yield the design rate R). Sections III and IV give *universal* bounds on the gap to capacity for general LDPC ensembles over MBIOS channels, no matter what the degree of the variable nodes is (as long as the design rate is fixed). These bounds can be exploited to gain insight on how good a specific design of degree distributions is in terms of the design rate and the average right degree where this is done by comparing the exact thresholds under iterative decoding with the lower bounds on the thresholds under ML decoding (see Tables II and III in Section V). On the other hand, the bounds are not necessarily tight for LDPC ensembles with a given pair of degree distributions (λ, ρ) since the explicit influence of λ is not taken into account except through the design rate of the ensemble. As a topic for further research, it is suggested to examine the possibility of tightening the bounds for specific ensembles by explicitly taking into account the exact characterization of λ. The numerical results shown in Section V indicate, however, that these bounds are useful for assessing the inherent gap to capacity of various LDPC ensembles. The gap to capacity is attributed to the finite average right degree of these LDPC ensembles [14].

As a topic for further research, we also suggest to study a possible generalization of the bounds to non-binary linear block codes. These generalized bounds can be applied to the analysis of the ML performance of non-binary LDPC ensembles whose transmission takes place over arbitrary discrete memoryless channels with possibly different types of quantization [2].

The lower bound on the asymptotic parity-check density in [14, Theorem 2.1] and its improvements in Sections III and IV grow like the log of the inverse of the gap (in rate) to capacity. The result in [14, Theorem 2.2] shows that a logarithmic growth rate of the parity-check density is achievable for Gallager's ensembles of regular LDPC codes under ML decoding when the transmission takes place over an arbitrary MBIOS channel. These results show that for any iterative decoder which is based on the representation of the codes by Tanner graphs, there exists a tradeoff between asymptotic performance and complexity which cannot be surpassed. Recently, it was shown in [10] that a better tradeoff can be achieved by allowing more complicated graphical models which involve a sufficient number of state nodes in the graph; for the BEC, the encoding and decoding complexity of properly designed codes on graphs remains bounded as the gap to capacity vanishes (see [10]).

In [15], the authors consider the achievable rates and decoding complexity of LDPC codes over statistically independent *parallel channels*, and generalize in a non-trivial way the un-quantized bounds introduced in Section IV. The bounds in [15] are applied to randomly and intentionally punctured LDPC codes, and improved puncturing theorems are derived as compared to those introduced in [10, Theorems 3 and 4].

APPENDIX I

A. Proof of Lemma 3.1

Since there is a one to one correspondence between the codewords and the set of information bits used to encode them, then $H(\mathbf{X}|\mathbf{Y}) = H(\mathbf{U}|\mathbf{Y})$ where the random vector $\mathbf{U} = (U_1, \ldots, U_{nR})$ denotes the sequence of information bits used to encode the codeword $\mathbf{X}$. Let $P_{\text{b}}^{(i)}$ denote the probability of decoding the bit U_i erroneously given the received sequence at the output of the MBIOS channel, then the bit error probability is given by

$$P_{\text{b}} = \frac{1}{nR} \sum_{i=1}^{nR} P_{\text{b}}^{(i)}. \tag{84}$$

This therefore gives

$$\begin{aligned} \frac{H(\mathbf{X}|\mathbf{Y})}{n} &= \frac{H(\mathbf{U}|\mathbf{Y})}{n} \\ &\overset{(a)}{\leq} \frac{1}{n} \sum_{i=1}^{nR} H(U_i|\mathbf{Y}) \end{aligned}$$

$$\begin{aligned}
&\overset{(b)}{\leq} \frac{1}{n}\sum_{i=1}^{nR} h_2\left(P_{\text{b}}^{(i)}\right)\\
&\overset{(c)}{\leq} R\,h_2\left(\frac{1}{nR}\sum_{i=1}^{nR} P_{\text{b}}^{(i)}\right)\\
&\overset{(d)}{=} R\,h_2\big(P_{\text{b}}\big)
\end{aligned}$$

where inequality (a) holds from the chain rule of the entropy and since conditioning reduces entropy, inequality (b) follows from Fano's inequality and since the code is binary, inequality (c) is based on Jensen's inequality and the concavity of the binary entropy function (h_2), and equality (d) follows from (84).

B. *Derivation of the Optimization Equation in* (24) *and Proving the Existence of its Solution*

Derivation of the optimization equation (24): We derive here the optimization equation (24) which refers to the ”four-level quantization” lower bound on the parity-check density.

Let the function a designate the conditional pdf of the LLR at the output of the original MBIOS channel, given the zero symbol is transmitted. In the following, we express the transition probabilities of the degraded channel in Fig. 1 (see p. 7) in terms of the pdf a and the value of l:

$$\begin{aligned}
p_0 &= \Pr(Z=0\,|\,X=0) = \int_l^\infty a(u)\,du && (85)\\
p_1 &= \Pr(Z=\alpha\,|\,X=0)\\
&= \int_{0^+}^{l} a(u)\,du + \frac{1}{2}\int_{0^-}^{0^+} a(u)\,du && (86)\\
p_2 &= \Pr(Z=1+\alpha\,|\,X=0)\\
&= \int_{-l}^{0^-} a(u)\,du + \frac{1}{2}\int_{0^-}^{0^+} a(u)\,du && (87)\\
p_3 &= \Pr(Z=1\,|\,X=0) = \int_{-\infty}^{-l} a(u)\,du. && (88)
\end{aligned}$$

We note that the integration of the function a from $u=0^-$ to $u=0^+$ gives a non-zero value if and only if there is a non-vanishing probability that the value of the LLR at the output of the original channel is zero (e.g., a BEC). Otherwise, the contribution of this integral to (86) and (87) vanishes. Since the channel is MBIOS, the symmetry property [11] gives

$$a(u) = e^u\,a(-u), \quad \forall\, u\in\mathbb{R}. \tag{89}$$

Based on the expressions for the coefficients K_1 and K_2 in the lower bound on the asymptotic parity-check density (22), then in order to find the tightest lower bound then we need to maximize

$$\frac{(p_1-p_2)^2}{p_1+p_2} + \frac{(p_0-p_3)^2}{p_0+p_3} \tag{90}$$

w.r.t. the free parameter $l\in\mathbb{R}^+$. From Eqs. (85)–(88) and the symmetry property in (89)

$$\begin{aligned}
&p_0-p_3 = \int_l^\infty a(u)(1-e^{-u})\,du\\
&\quad\Rightarrow\quad \frac{\partial}{\partial l}(p_0-p_3) = -a(l)(1-e^{-l}) && (91)\\
&p_0+p_3 = \int_l^\infty a(u)(1+e^{-u})\,du\\
&\quad\Rightarrow\quad \frac{\partial}{\partial l}(p_0+p_3) = -a(l)(1+e^{-l}) && (92)\\
&p_1-p_2 = \int_{0^+}^{l} a(u)(1-e^{-u})\,du
\end{aligned}$$

$$\Rightarrow \quad \frac{\partial}{\partial l}(p_1 - p_2) = a(l)(1 - e^{-l}) \tag{93}$$

$$p_1 + p_2 = \int_{0^+}^{l} a(u)(1 + e^{-u})\, du$$

$$\Rightarrow \quad \frac{\partial}{\partial l}(p_1 + p_2) = a(l)(1 + e^{-l}) \tag{94}$$

so the calculation of the partial derivative of (90) w.r.t. l gives

$$\begin{aligned}&\frac{\partial}{\partial l}\left\{\frac{(p_1 - p_2)^2}{p_1 + p_2} + \frac{(p_0 - p_3)^2}{p_0 + p_3}\right\}\\ &= -4\, a(l)\Biggl\{\left[\left(\frac{p_2}{p_1 + p_2}\right)^2 - \left(\frac{p_3}{p_0 + p_3}\right)^2\right]\\ &\qquad\qquad + e^{-l}\left[\left(\frac{p_1}{p_1 + p_2}\right)^2 - \left(\frac{p_0}{p_0 + p_3}\right)^2\right]\Biggr\}.\end{aligned}$$

Since the first derivative of a function changes its sign at a neighborhood of any local maxima or minima point, and since the pdf a is always non-negative, then the second multiplicative term above is the one which changes its sign at a neighborhood of l maximizing (90). For this value of l, the second multiplicative term vanishes, which gives the optimization equation for l in (24).

Proof of existence of a solution to (24)*:* In order to show that a solution to (24) always exists, we will see how the LHS and the RHS of this equation behave as $l \to 0^+$ and $l \to \infty$. From (85)–(88), it follows that in the limit where $l \to \infty$

$$p_1 \to 1 - w - \Pr(\mathrm{LLR}(Y) = \infty \mid X = 0), \quad p_2 \to w$$

where w is introduced in (2), and therefore

$$\lim_{l \to \infty} \frac{p_2^2 + e^{-l} p_1^2}{(p_1 + p_2)^2} = \left(\frac{w}{1 - \Pr(\mathrm{LLR}(Y) = \infty \mid X = 0)}\right)^2. \tag{95}$$

Since from the symmetry property

$$\begin{aligned} p_3 &= \int_{l}^{\infty} a(-u) du\\ &= \int_{l}^{\infty} e^{-u} a(u) du\\ &\le e^{-l} \int_{l}^{\infty} a(u) du\\ &= e^{-l} p_0\end{aligned}$$

then the fraction $\frac{p_3}{p_0}$ tends to zero as $l \to \infty$, so

$$\lim_{l \to \infty} \frac{p_3^2 + e^{-l} p_0^2}{(p_0 + p_3)^2} = \lim_{l \to \infty} \frac{\left(\frac{p_3}{p_0}\right)^2 + e^{-l}}{\left(1 + \frac{p_3}{p_0}\right)^2} = 0. \tag{96}$$

It therefore follows from (95) and (96) that for large enough values of l, the LHS of (24) is larger than the RHS of this equation. On the other hand, in the limit where $l \to 0^+$, we get

$$p_1, p_2 \to \frac{1}{2} \int_{0^-}^{0^+} a(u) du$$

and therefore

$$\lim_{l \to 0^+} \frac{p_2^2 + e^{-l} p_1^2}{(p_1 + p_2)^2} = \frac{1}{2}. \tag{97}$$

In the limit where $l \to 0^+$, one also gets

$$p_0 \to \int_{0^+}^{\infty} a(u) du, \quad p_3 \to \int_{-\infty}^{0^-} a(u) du, \quad p_0 + p_3 \to \beta$$

where $\beta \triangleq 1 - \displaystyle\int_{0^-}^{0^+} a(u) du$. By denoting $u \triangleq \displaystyle\int_{0^+}^{\infty} a(u) du$,
we get $0 \le u \le \beta$, and

$$\lim_{l \to 0^+} \frac{p_3^2 + e^{-l} p_0^2}{(p_0 + p_3)^2} = \frac{u^2 + (\beta - u)^2}{\beta^2} \ge \frac{1}{2}\,, \quad \forall\, u \in [0, \beta]. \tag{98}$$

We note that the last inequality holds in equality if and only if $u = \frac{\beta}{2}$. But if this condition holds, then this implies that

$$\int_{-\infty}^{0^-} a(u)\ du = \int_{0^+}^{\infty} a(u)\ du$$

which from the symmetry property cannot be satisfied unless $a(u) = \delta(u)$. The latter condition corresponds to a BEC with erasure probability 1 (whose capacity is equal to zero).

From (97) and (98), we obtain that for small enough (and non-negative) values of l, the LHS of (24) is less or equal to the RHS of this equation. Since we also obtained that for large enough l, the LHS of (24) is larger than the RHS of this equation, the existence of a solution to (24) follows from continuity considerations.

C. *Proof of Inequality* (33)

We prove here the inequality (33) (see p. 13) which implies that the ”four-level quantization” lower bound on the parity-check density (see p. 10) is tighter than what can be interpreted as the ”two levels quantization” bound in [14, Theorem 2.1]. Based on (2), we get

$$w = \Pr\{\text{LLR}(Y) < 0 \,|\, X = 0\} + \frac{1}{2} \Pr\{\text{LLR}(Y) = 0 \,|\, X = 0\}$$

so from (9), $w = p_2 + p_3$. By invoking Jensen’s inequality, we get

$$\begin{aligned}
&\frac{(p_1 - p_2)^2}{p_1 + p_2} + \frac{(p_0 - p_3)^2}{p_0 + p_3} \\
&= (p_1 + p_2) \left(\frac{p_1 - p_2}{p_1 + p_2}\right)^2 + (p_0 + p_3) \left(\frac{p_0 - p_3}{p_0 + p_3}\right)^2 \\
&\ge \left[(p_1 + p_2) \left(\frac{p_1 - p_2}{p_1 + p_2}\right) + (p_0 + p_3) \left(\frac{p_0 - p_3}{p_0 + p_3}\right)\right]^2 \\
&= (p_0 + p_1 - p_2 - p_3)^2 \\
&= (1 - 2p_2 - 2p_3)^2 \\
&= (1 - 2w)^2.
\end{aligned}$$

An equality is achieved if and only if $\frac{p_1 - p_2}{p_1 + p_2} = \frac{p_0 - p_3}{p_0 + p_3}$. From (91)–(94), we get

$$\frac{p_1 - p_2}{p_1 + p_2} = \frac{\displaystyle\int_{0^+}^{l} a(u)(1 - e^{-u})\ du}{\displaystyle\int_{0^+}^{l} a(u)(1 + e^{-u})\ du} \le \frac{1 - e^{-l}}{1 + e^{-l}}$$

and

$$\frac{p_0 - p_3}{p_0 + p_3} = \frac{\displaystyle\int_{l}^{\infty} a(u)(1 - e^{-u})\ du}{\displaystyle\int_{l}^{\infty} a(u)(1 + e^{-u})\ du} \ge \frac{1 - e^{-l}}{1 + e^{-l}}\ .$$

The two fractions $\frac{p_1-p_2}{p_1+p_2}$ and $\frac{p_0-p_3}{p_0+p_3}$ cannot be equal unless the LLR is either equal to l or $-l$. This makes the four-level quantization of the LLR identical to the two-level quantization used for the derivation of the original bound in [1, Theorem 2]. Equality can be also achieved if $p_1+p_2=0$ or $p_0+p_3=0$ which converts the channel model in Fig. 1 (see p. 7) to a BSC.

Appendix II

This appendix provides further mathematical details related to the proof of Proposition 4.1. We note that Appendix II-A serves here as a preparatory step for the derivation in Appendix II-B.

A. Power Series Expansion of the Binary Entropy Function

Lemma II.1:

$$h_2(x) = 1 - \frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{(1-2x)^{2p}}{p(2p-1)}\,, \quad 0 \le x \le 1. \tag{99}$$

Proof: We prove this by expanding the binary entropy function into a power series around $\frac{1}{2}$. The first order derivative is

$$h_2'(x) = \frac{\ln\left(\frac{1-x}{x}\right)}{\ln 2}$$

and the higher order derivatives get the form

$$h_2^{(n)}(x) = -\frac{(n-2)!}{\ln 2}\left(\frac{(-1)^n}{x^{n-1}} + \frac{1}{(1-x)^{n-1}}\right)$$

for $n = 2, 3, \ldots$. The derivatives of odd degree therefore vanish at $x=\frac{1}{2}$, and for an even value of $n \ge 2$

$$h_2^{(n)}\left(\frac{1}{2}\right) = -\frac{(n-2)!\;2^n}{\ln 2}.$$

This yields the following power series expansion of the binary entropy function around the point $\frac{1}{2}$:

$$\begin{aligned} h_2(x) &= 1 - \sum_{n\ge 2 \text{ even}}\left\{\frac{\frac{(n-2)!\;2^n}{\ln 2}}{n!}\cdot\left(x-\frac{1}{2}\right)^n\right\} \\ &= 1 - \frac{1}{\ln 2}\sum_{n\ge 2 \text{ even}}\frac{(2x-1)^n}{n(n-1)} \\ &= 1 - \frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{(2x-1)^{2p}}{p(2p-1)} \end{aligned}$$

and this power series converges for all $x \in [0,1]$. ■

We note that since the power series in (99) has always non-negative coefficients, then its truncation always gives an upper bound on the binary entropy function, i.e.,

$$h_2(x) \le 1 - \frac{1}{2\ln 2}\sum_{p=1}^{m}\frac{(1-2x)^{2p}}{p(2p-1)} \qquad \forall\, x\in[0,1],\; m\in\mathbb{N}. \tag{100}$$

The case where $m=1$ gives the upper bound in Lemma 3.2 which is used in this paper for the derivation of the lower bounds on the parity-check density. The reason for not using a tighter version of the binary entropy function for this case was because otherwise we would get a polynomial equation for a_{R} whose solution cannot be given necessarily in closed form. As shown in Fig. 5, the upper bound on the binary entropy function h_2 over the whole interval $[0,1]$ is improved considerably by taking even a moderate value for m (e.g., $m=10$ gives already a very tight upper bound on h_2 which deviates from the exact values only at a small neighborhood near the two endpoints of this interval).

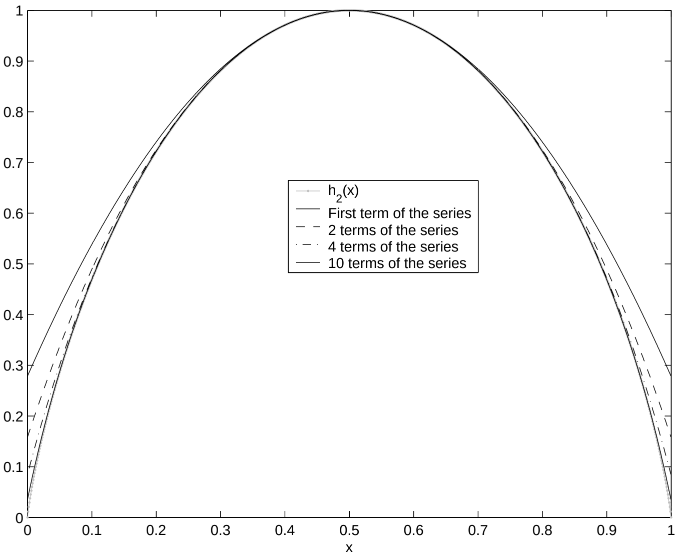

Fig. 5. Plot of the binary entropy function to base 2 and some upper bounds which are obtained by truncating its power series around $x = \frac{1}{2}$.

B. Calculation of the Multi-Dimensional Integral in (65)

Based on (99), we get

$$
\begin{aligned}
&\int_0^\infty \cdots \int_0^\infty \prod_{m=1}^{k} f_\Omega(\alpha_m) \\
&\qquad \cdot h_2\left(\frac{1}{2}\left(1 - \prod_{m=1}^{k}\left(\frac{1-e^{-\alpha_m}}{1+e^{-\alpha_m}}\right)\right)\right) d\alpha_1 \ldots d\alpha_k \\
&= 1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \Bigg\{ \frac{1}{p(2p-1)} \\
&\qquad \cdot \int_0^\infty \cdots \int_0^\infty \prod_{m=1}^{k} f_\Omega(\alpha_m) \\
&\qquad\qquad \cdot \prod_{m=1}^{k} \left(\frac{1-e^{-\alpha_m}}{1+e^{-\alpha_m}}\right)^{2p} d\alpha_1 \ldots d\alpha_k \Bigg\} \\
&= 1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \Bigg\{ \frac{1}{p(2p-1)} \\
&\qquad \cdot \int_0^\infty \cdots \int_0^\infty \prod_{m=1}^{k} \left(f_\Omega(\alpha_m)\, \tanh^{2p}\left(\frac{\alpha_m}{2}\right)\right) \\
&\qquad\qquad d\alpha_1 \ldots d\alpha_k \Bigg\} \\
&= 1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{1}{p(2p-1)} \left(\int_0^\infty f_\Omega(\alpha) \tanh^{2p}\left(\frac{\alpha}{2}\right) d\alpha \right)^{k}.
\end{aligned}
$$

This transforms the original k-dimensional integral to an infinite sum of one-dimensional integrals. Since we are interested in obtaining a tight upper bound on the k-dimensional integral above, and all the terms of the last infinite series are positive, then any truncation of the last infinite series forms an upper bound on the multi-dimensional integral given in (65). Based on the discussion in Appendix II-A, we compute the first 10 terms of this series which (see the plot in Fig. 5) give a very tight upper bound on the k-dimensional integral (for all k).

APPENDIX III

A. On the Tightness of the Lower Bound in Theorem 4.1

We show here that the lower bound on the parity-check density in Theorem 4.1 is uniformly tighter than the one in [14, Theorem 2.1] (except for the BSC and BEC where the two bounds coincide). In order to show this , we first prove the following lemma:

Lemma III.1: For any MBIOS channel, $g_1 \geq (1-2w)^2$ where w and g_1 are introduced in (2) and (57), respectively.

Proof: From (2), (57) and (59)

$$\begin{aligned} g_1 &= \int_0^\infty a(l)\,\left(1+e^{-l}\right)\,\tanh^2\left(\frac{l}{2}\right)\,dl \\ &= \int_0^\infty f_\Omega(l)\,\tanh^2\left(\frac{l}{2}\right)\,dl \\ &\geq \left(\int_0^\infty f_\Omega(l)\,\tanh\left(\frac{l}{2}\right)\,dl\right)^2 \\ &= \left(\int_0^\infty a(l)\,\left(1+e^{-l}\right)\cdot\left(\frac{1-e^{-l}}{1+e^{-l}}\right)\,dl\right)^2 \\ &= \left(\int_{0^+}^\infty a(l)\,dl - \int_{0^+}^\infty e^{-l}\,a(l)\,dl\right)^2 \\ &= \left(\int_{0^+}^\infty a(l)\,dl - \int_{0^+}^\infty a(-l)\,dl\right)^2 \\ &= \left(\int_{0^+}^\infty \left[a(l)+a(-l)\right]\,dl - 2\int_{0^+}^\infty a(-l)\,dl\right)^2 \\ &= \left(\int_{-\infty}^\infty a(l)\,dl - \int_{0^-}^{0^+} a(l)\,dl - 2\int_{0^+}^\infty a(-l)\,dl\right)^2 \\ &= \left(1-2\left(\int_{-\infty}^{0^-} a(l)\,dl + \frac{1}{2}\int_{0^-}^{0^+} a(l)\,dl\right)\right)^2 \\ &= (1-2w)^2 \end{aligned}$$

where the single inequality above follows from Jensen's inequality. Note also that $g_1 \leq 1$ with equality if and only if the channel is noiseless. ■

The proof of the claim stated at the beginning of this appendix follows directly by noticing that the lower bound on the parity-check density, as given in (70)–(72), is equal to

$$\frac{K_1 + K_2 \ln\left(\frac{1}{\varepsilon}\right)}{1-\varepsilon} = \frac{1-C}{(1-\varepsilon)C}\,\frac{\ln\left(\frac{1}{2\ln 2}\,\frac{1-C}{\varepsilon C}\right)}{\ln\left(\frac{1}{g_1}\right)}$$

where we refer here to all MBIOS channels except the BEC. On the other hand, the lower bound on the parity-check density which is given in [14, Theorem 2.1] gets the form

$$\frac{1-C}{(1-\varepsilon)C}\,\frac{\ln\left(\frac{1}{2\ln 2}\,\frac{1-C}{\varepsilon C}\right)}{\ln\left(\frac{1}{(1-2w)^2}\right)}\,.$$

If the lower bound is not trivial (i.e., the common numerator in both bounds is positive), then the improvement in the tightness of the former bound over the latter bound follows from Lemma III.1. We note that for the particular case of a BSC, the above two bounds coincide (as for a BSC whose crossover probability is p, it is easy to verify from (2) and (57) that $w = p$ and $g_1 = (1-2p)^2$, respectively, hence $g_1 = (1-2w)^2$).

B. *Proof for the Claim in Remark 4.4*

In order to prove the claim in Remark 4.4 (see p. 28), it is required to show that

$$\begin{aligned} &\frac{1-C}{1-\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\left\{\frac{1}{p(2p-1)}\sum_k \Gamma_k\, g_p^{\,k}\right\}} \\ &\geq \frac{2w}{1-\sum_k \Gamma_k\,(1-2w)^k} \end{aligned} \tag{101}$$

where w is introduced in (2). The reason for showing this in light of the claim in Remark 4.4 is that the RHS of the last inequality follows from considerations related to a BEC, essentially in the same way that the second term of the maximization on the RHS of (55) is derived. By showing this, we prove that the maximization of the two expressions in the LHS and RHS of (101) doesn't affect the bound in Corollary 4.1.

Following the steps which lead to (75), we get that for any integer $k \geq 2$

$$\frac{1}{2\ln 2}\sum_{p=1}^{\infty}\frac{g_p^{\,k}}{p(2p-1)} \geq C^k.$$

Applying this to (101) and denoting $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$, we get that a sufficient condition for (101) to hold is

$$\frac{1-C}{1-\Gamma(C)} \geq \frac{2w}{1-\Gamma(1-2w)}. \tag{102}$$

From the erasure decomposition lemma, we get that an MBIOS channel is physically degraded as compared to a BEC with an erasure probability $p = 2w$. By the information processing inequality, it follows that $C \leq 1-2w$. Therefore, in order to prove (102), it is enough to show that the function

$$f(x) = \frac{1-x}{1-\Gamma(x)}$$

is monotonically decreasing for $x \in (0,1)$. We prove this property by showing that the derivative of the function f is non-positive on the interval $(0,1)$. As the denominator of the derivative is positive, we may equivalently show that

$$\Gamma'(x)\,(1-x) - \big(1-\Gamma(x)\big) \leq 0.$$

Dividing both sides of the inequality by $1-x$ which is in the interval $\in (0,1)$ and noting that $\Gamma(1) = \sum_k \Gamma_k = 1$, we get that it is enough to show that for $x \in (0,1)$

$$\Gamma'(x) - \frac{\Gamma(1)-\Gamma(x)}{1-x} \leq 0. \tag{103}$$

Since the function Γ is a polynomial and therefore analytic, by the mean-value theorem we get that for some $\tilde{x} \in (x,1)$

$$\frac{\Gamma(1)-\Gamma(x)}{1-x} = \Gamma'(\tilde{x}).$$

Since $\Gamma'(x) = \sum_k k\Gamma_k x^{k-1}$ is monotonically increasing for $x \geq 0$, then (103) follows for all $x \in (0,1)$. This in turn proves (101).

C. *Proof of Eq.* (82)

In order to prove (82), we first multiply the two sides of (80) by R, and denote $R = (1-\varepsilon)C$. This gives that the lower bound on the bit error probability in (80) is non-positive if and only if

$$(1-C)B - \varepsilon C(1-B) \leq 0. \tag{104}$$

Unless the channel is noiseless, we get

$$\begin{aligned}
B &= \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)} \\
&= \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \left\{ \frac{1}{p(2p-1)} \sum_k \Gamma_k \, {g_p}^k \right\} \\
&= \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \left\{ \frac{1}{p(2p-1)} \right. \\
&\qquad \left. \cdot \sum_k \Gamma_k \left(\int_{0^+}^{\infty} a(l)(1+e^{-l}) \tanh^{2p}\left(\frac{l}{2}\right) dl \right)^k \right\} \\
&< \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \left\{ \frac{1}{p(2p-1)} \sum_k \Gamma_k \left(\int_{0^+}^{\infty} a(l)(1+e^{-l})\, dl \right)^k \right\} \\
&= \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \left\{ \frac{1}{p(2p-1)} \sum_k \Gamma_k \left(\int_{\mathbb{R}-\{0\}} a(l)\, dl \right)^k \right\} \\
&= \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \left\{ \frac{1}{p(2p-1)} \sum_k \Gamma_k \Big(1 - \Pr(\text{LLR} = 0) \Big)^k \right\} \\
&\leq \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{1}{p(2p-1)} = 1.
\end{aligned}$$

Since $B < 1$, the LHS of (104) is monotonically decreasing in ε. We therefore deduce that the inequality (104) holds for $\varepsilon \geq \varepsilon_0$, where ε_0 is the solution of

$$(1-C)B - \varepsilon_0 C(1-B) = 0.$$

It can be readily seen that the solution of the last equation is given by ε_0 defined in (82).

D. *Proof of Eq.* (83)

We will show both that there exists a unique ε_0 that satisfies (83), and that the RHS of (81) is non-positive if and only if $\varepsilon \geq \varepsilon_0$ where ε_0 is that unique solution. As in Appendix III-C, we begin by multiplying the two sides of (81) by R and denoting $R = (1-\varepsilon)C$. It follows that the bound in the RHS of (81) is trivial (non-positive) if and only if

$$-\varepsilon C + \frac{1-(1-\varepsilon)C}{2\ln 2} \sum_{p=1}^{\infty} \frac{g_p^{\frac{(2-(1-\varepsilon)C)t}{1-(1-\varepsilon)C}}}{p(2p-1)} \leq 0 \tag{105}$$

where g_p is introduced in (57). We now show that the LHS of the last inequality is monotonically decreasing in ε. Let us denote

$$\begin{aligned}
f(\varepsilon) &\triangleq -\varepsilon C + \frac{1-(1-\varepsilon)C}{2\ln 2} \sum_{p=1}^{\infty} \frac{g_p^{\frac{(2-(1-\varepsilon)C)t}{1-(1-\varepsilon)C}}}{p(2p-1)} \\
\alpha_p &\triangleq \frac{1}{2\ln 2} \, \frac{1}{p(2p-1)}, \quad p \in \mathbb{N}.
\end{aligned}$$

By dividing the derivative of f w.r.t. ε by C, we get

$$\begin{aligned}\frac{f'(\varepsilon)}{C} &= \frac{1}{C}\Bigg(-C + C\sum_{p=1}^{\infty}\alpha_p\, g_p^{\frac{(2-(1-\varepsilon)C)t}{1-(1-\varepsilon)C}} \\ &\qquad + \big(1-(1-\varepsilon)C\big)\sum_{p=1}^{\infty}\alpha_p\, g_p^{\frac{(2-(1-\varepsilon)C)t}{1-(1-\varepsilon)C}}\,\ln(g_p) \\ &\qquad\qquad \cdot\left(-\frac{tC}{(1-(1-\varepsilon)C)^2}\right)\Bigg) \\ &= \sum_{p=1}^{\infty}\left\{\alpha_p\,\left(1-\ln\Big(g_p^{\frac{t}{1-(1-\varepsilon)C}}\Big)\right) g_p^{\frac{(2-(1-\varepsilon)C)t}{1-(1-\varepsilon)C}}\right\} - 1\,. \end{aligned} \qquad (106)$$

From the symmetry property of the pdf a then (57) yields that

$$\begin{aligned} g_p &= \int_0^{\infty} a(l)(1+e^{-l})\tanh^{2p}\left(\frac{l}{2}\right)\,dl \\ &= \int_{-\infty}^{\infty} a(l)\tanh^{2p}\left(\frac{l}{2}\right)\,dl \\ &\le \int_{-\infty}^{\infty} a(l)dl = 1 \end{aligned}$$

and, since $g_p \le 1$, it follows that $g_p^{\frac{(2-(1-\varepsilon)C)t}{1-(1-\varepsilon)C}} \le g_p^{\frac{t}{1-(1-\varepsilon)C}}$. Therefore, (106) gives

$$\frac{f'(\varepsilon)}{C} \le \sum_{p=1}^{\infty}\left\{\alpha_p\,\left(1-\ln\big(g_p^{\frac{t}{1-(1-\varepsilon)C}}\big)\right) g_p^{\frac{t}{1-(1-\varepsilon)C}}\right\} - 1.$$

For $p\in\mathbb{N}$, let us denote $g_p^{\frac{t}{1-(1-\varepsilon)C}} \triangleq 1-\delta_p$ where $0<\delta_p<1$, then the last inequality gives

$$\begin{aligned}\frac{f'(\varepsilon)}{C} &\le \sum_{p=1}^{\infty}\Big\{\alpha_p\,\big(1-\ln(1-\delta_p)\big)(1-\delta_p)\Big\} - 1 \\ &\le \sum_{p=1}^{\infty}\alpha_p - 1 = 0\end{aligned}$$

where the second transition follows from the inequality

$$\ln(1-x) > -\frac{x}{1-x}, \quad x\in(0,1).$$

This concludes the proof of the monotonicity of the LHS of (105). Observing that

$$f(0) = (1-C)\sum_{p=1}^{\infty}\alpha_p\, g_p^{\frac{(2-C)t}{1-C}} > 0$$

and

$$\begin{aligned} f(1) &= -C + \sum_{p=1}^{\infty}\alpha_p\, g_p^{2t} \\ &\le -C + \sum_{p=1}^{\infty}\alpha_p\, g_p \\ &= -C + C = 0 \end{aligned}$$

where the first inequality follows since $g_p \le 1$ and $t \ge 1$ (asymptotically, $t = 1$ if and only if the code is cycle-free). The second equality follows from the last three equalities leading to (75). From the continuity of the function f w.r.t. ε, we conclude that the monotonicity property of f, as shown above, ensures a unique solution for (83). From (105), it also follows from the monotonicity and continuity properties of the function f in terms of $\varepsilon\in(0,1)$ that the RHS of (81) is non-positive if and only if $\varepsilon \ge \varepsilon_0$ where ε_0 is the unique solution of (83).

Acknowledgment

I. Sason wishes to acknowledge Rüdiger Urbanke for stimulating discussions during the preparation of the work in [14] which motivated the research in this paper, and for his comments on an earlier version of this paper. We wish to acknowledge the anonymous reviewers for their detailed comments which greatly improved the lucidity of the presentation. The work of I. Sason was supported by the Taub and Shalom Foundations.

References

[1] D. Burshtein, M. Krivelevich, S. Litsyn and G. Miller, "Upper bounds on the rate of LDPC codes," *IEEE Trans. on Information Theory*, vol. 48, no. 9, pp. 2437–2449, September 2002.

[2] A. Bennatan and D. Burshtein, "On the application of LDPC codes to arbitrary discrete-memoryless channels," *IEEE Trans. on Information Theory*, vol. 50, no. 3, pp. 417–438, March 2004.

[3] R. G. Gallager, *Low-density parity-check codes*, Cambridge, MA, USA, MIT Press, 1963.

[4] H. Jin and R. J. McEliece, "Typical pairs decoding on the AWGN channel," *Proceedings 2000 International Symposium on Information Theory and Its Applications*, pp. 180–183, Honolulu, Hawaii, U.S.A., November 5–8, 2000.

[5] A. Khandekar and R. J. McEliece, "On the complexity of reliable communications on the erasure channel," *Proceedings 2001 International Symposium on Information Theory*, p. 1, Washington, DC, USA, June 24–29, 2001.

[6] C. Measson, A. Montanari, T. Richardson and R. Urbanke, "Life above threshold: From list decoding to area theorem and MSE," *2004 IEEE Information Theory Workshop*, San Antonio, TX, USA, October 24–29, 2004.

[7] C. Measson, A. Montanari and R. Urbanke, "Maxwell construction: the hidden bridge between iterative and maximum a posteriori decoding," accepted to *IEEE Trans. on Information Theory*. [Online]. Available: http://www.arxiv.org/abs/cs.IT/0506083.

[8] A. Montanari, "Tight bounds for LDPC codes and LDGM codes under MAP decoding," *IEEE Trans. on Information Theory*, vol. 51, no. 9, pp. 3221–3246, September 2005.

[9] A. Montanari, *personal communications*, May 2005.

[10] H. D. Pfister, I. Sason, and R. Urbanke, "Capacity-achieving ensembles for the binary erasure channel with bounded complexity," *IEEE Trans. on Information Theory*, vol. 51, no. 7, pp. 2352–2379, July 2005.

[11] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 599–618, February 2001.

[12] T. Richardson, A. Shokrollahi and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 619–637, February 2001.

[13] T. Richardson and R. Urbanke, *Modern Coding Theory*, to be published, Cambridge Press. [Online]. Available: http://lthcwww.epfl.ch/mct/index.php.

[14] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1611–1635, July 2003.

[15] I. Sason and G. Wiechman, "On achievable rates and complexity of LDPC codes over parallel channels: Bounds and applications," to appear in the *IEEE Trans. on Information Theory*, February 2007.

[16] E. Sharon, A. Ashikmin, and S. Litsyn, "EXIT functions for binary input memoryless symmetric channels," *IEEE Trans. on Communications*, vol. 54, no. 7, pp. 1207–1214, July 2006.

[17] A. Shokrollahi, "New sequences of time erasure codes approaching channel capacity," in *Proceedings of the 13th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, Lectures Notes in Computer Science 1719, Springer Verlag, pp. 65–76, 1999.

[18] A. Trachtenberg, T. Etzion and A. Vardy, "Which codes have cycle-free Tanner graphs ?," *IEEE Trans. on Information Theory*, vol. 45, no. 6, pp. 2173–2181, September 1999.

[19] R. Urbanke, *Optimization of degree distributions for ensembles of LDPC codes*. [Online]. Available: http://lthcwww.epfl.ch/research/ldpcopt/index.php.