On Universal Properties of Capacity-Approaching LDPC Code Ensembles

Igal Sason, *Member*

Abstract

This paper is focused on the derivation of some universal properties of capacity-approaching low-density parity-check (LDPC) code ensembles whose transmission takes place over memoryless binary-input output-symmetric (MBIOS) channels. Properties of the degree distributions, graphical complexity and the number of fundamental cycles in the bipartite graphs are considered via the derivation of information-theoretic bounds. These bounds are expressed in terms of the target block/ bit error probability and the gap (in rate) to capacity. Most of the bounds are general for any decoding algorithm, and some others are proved under belief propagation (BP) decoding. Proving these bounds under a certain decoding algorithm, validates them automatically also under any sub-optimal decoding algorithm. A proper modification of these bounds makes them universal for the set of all MBIOS channels which exhibit a given capacity. Bounds on the degree distributions and graphical complexity apply to finite-length LDPC codes and to the asymptotic case of an infinite block length. The bounds are compared with capacity-approaching LDPC code ensembles under BP decoding, and they are shown to be informative and are easy to calculate. Finally, some interesting open problems are considered.

Index Terms

Belief propagation (BP), bipartite graphs, complexity, cycles, density evolution (DE), linear programming (LP) bounds, low-density parity-check (LDPC) codes, maximum-likelihood (ML) decoding, memoryless binary-input output-symmetric (MBIOS) channels, sphere-packing bounds, stability.

I. Introduction

Low-density parity-check (LDPC) codes form a class of powerful error-correcting codes which are efficiently encoded and decoded with low-complexity algorithms. These linear block codes, originally introduced by Gallager in the early sixties [14], are characterized by sparse parity-check matrices which facilitate their low-complexity decoding with iterative message-passing algorithms. In spite of the seminal work of Gallager, LDPC codes were ignored for a long time. Following the breakthrough in coding theory, made by the introduction of turbo codes [5] and the rediscovery of LDPC codes [25] in the mid 1990s, it was realized that an efficient design of these codes enables to closely approach the channel capacity while maintaining reasonable decoding complexity. This breakthrough attracted many coding-theorists during the last decade (see, e.g., [9], [37], [55]).

The asymptotic analysis of LDPC code ensembles under iterative message-passing decoding algorithms relies on the *density evolution* (DE) approach which was developed by Richardson and Urbanke (see [34], [35], [37]). This technique is commonly used for optimizing the degree distributions of capacity-approaching LDPC code ensembles where the target is to maximize the achievable rate for a given channel model or to maximize the threshold for a given code rate subject to some constraints on the degree distributions [2]. Some approximate techniques which optimize the degree distributions of LDPC code ensembles under further practical constraints are of interest (e.g., an optimization for obtaining a good tradeoff between the asymptotic gap to capacity and the decoding complexity [3]). For the binary erasure channel (BEC), the DE approach is much simplified since it leads to a one-dimensional analysis. As a result of this significant simplification, some explicit expressions for capacity-achieving sequences of LDPC code ensembles have been derived for the BEC (see, e.g., [24], [29], [37] and [48]). For general memoryless binary-input output-symmetric (MBIOS) channels, as of yet there are no closed-form expressions for capacity-achieving LDPC code ensembles under iterative decoding, and the DE technique serves as a numerical tool for the design of capacity-approaching LDPC code ensembles in the limit where their block length tends to infinity. Although maximum-likelihood (ML) decoding is prohibitively complex, capacity-achieving sequences of LDPC

This research work was supported by the Israel Science Foundation (grant no. 1070/07).

Igal Sason is with the Department of Electrical Engineering at the Technion – Israel Institute of Technology, Haifa 32000, Israel (e-mail: sason@ee.technion.ac.il).

code ensembles have been constructed under ML decoding for any MBIOS channel where the analysis relies on upper bounds on the decoding error probability which are based on the distance spectra of these ensembles (see [18], [19], [39], and [40, Theorem 2.2]).

Consider right-regular LDPC codes (i.e., LDPC codes where the degree of the parity-check nodes is fixed to a certain value $a_{\rm R}$), and assume that their transmission takes place over a binary symmetric channel (BSC). In his thesis, Gallager derived an upper bound on the maximal achievable rate of these codes where it is required to obtain vanishing block error probability as we let the block length tend to infinity (see [14, Theorem 3.3]). This information-theoretic bound holds under ML decoding or any sub-optimal decoding algorithm. This bound shows that right-regular LDPC codes cannot achieve the channel capacity on a BSC, even under ML decoding. Based on this bound, the inherent gap between the achievable rate and the channel capacity is well approximated by an expression which decreases to zero exponentially fast in $a_{\rm R}$. Burshtein *et al.* have generalized Gallager's bound for general LDPC code ensembles whose transmission takes place over an MBIOS channel [7]. An improved upper bound on the achievable rates of LDPC code ensembles was obtained by Wiechman and Sason [53], followed by a generalization of this bound to the case where the transmission takes place over a set of parallel MBIOS channels [41]. This work partially relies on the analysis in [53] (see Section II for relevant background).

Khandekar and McEliece suggested to measure the encoding and decoding complexity of codes defined on graphs in terms of the achievable gap (in rate) to capacity, and they also had some conjectures regarding the behavior of the complexity as the gap to capacity vanishes [21]. Following their approach, the tradeoff between the performance and complexity is analyzed in the literature for LDPC code ensembles and some other variants of codes defined on graphs (see, e.g., [18], [19], [31], [32], [40], [41], [42], [53] and references therein).

In this paper, we consider some properties of capacity-approaching LDPC code ensembles whose transmission takes place over MBIOS channels. One question which is addressed in this paper is the following:

Question 1: How do the degree distributions of capacity-approaching LDPC code ensembles behave as a function of the achievable gap (in rate) to capacity ?

The behavior of the degree distributions of capacity-approaching LDPC code ensembles is addressed in this work via the derivation of some information-theoretic bounds. Some of them hold under ML decoding or any sub-optimal decoding algorithm, and some other bounds are proved under belief propagation (BP) decoding where we refer to the sum-product decoding algorithm (see [22] and [37, Chapter 2]). For the characterization of the degree distributions for capacity-approaching LDPC code ensembles, a special consideration is given to the fraction of degree–2 variable nodes (L_2) and the fraction of edges connected to these nodes (λ_2). This focus was partially motivated by the influence of λ_2 on the satisfiability of the stability condition; this condition is necessary for achieving vanishing bit error probability under iterative message-passing decoding when we let the block length tend to infinity [34]. Also, some previously reported information-combining bounds on the performance of LDPC code ensembles under iterative decoding are sensitive to this quantity (see, e.g., [49]). This motivates a study of the behavior of L_2 and λ_2 for capacity-approaching LDPC code ensembles, where the bounds on these quantities are expressed in terms of the gap between the channel capacity and the achievable rates of these code ensembles under BP decoding. We also demonstrate the tightness of these bounds for the BEC by considering the right-regular sequence of capacity-achieving LDPC code ensembles proposed by Shokrollahi [48].

General upper bounds on the degree distributions of capacity-approaching LDPC code ensembles are derived in this paper for the case where the transmission takes place over an MBIOS channel. The bounds are expressed in terms of the gap (in rate) to capacity with a target bit (or block) error probability. These linear programming (LP) upper bounds on the degree distributions of LDPC code ensembles are general with respect to the decoding algorithm, and they also hold for ensembles of finite-length codes or for the asymptotic case of an infinite block length. We note that two LP problems are formulated in [1] for optimizing the degree distributions of finite-length LDPC code ensembles whose transmission takes place over a BEC, and also a convex optimization problem is formulated in [3] for optimizing the degree distributions of LDPC code ensembles with the goal of obtaining a good tradeoff between performance and decoding complexity. It is noted that the LP-based optimizations in [1] and [3] hold under BP decoding, whereas the LP bounds which are derived in this paper are information-theoretic bounds which hold under ML decoding or any sub-optimal decoding algorithm. Although the degree distributions of the parity-check nodes are often set to be regular (or almost regular), and the irregularity often refers to the degree distributions of the variable nodes, this is not necessarily the case for capacity-approaching ensembles. For example, [32, Section VI] introduces some capacity-achieving sequences of accumulate-repeat-accumulate code ensembles for the BEC, which also possess a bounded complexity per information bit under BP decoding; they are designed in a way where the degree distributions of the LDPC code ensembles after a proper graph reduction (as explained in [32, Section II]) are self-matched and are both irregular. The irregularity of the parity-check degree distributions in the design of LDPC codes appears to be useful in various cases under BP decoding, e.g., the optimization of finite-length LDPC code ensembles whose transmission takes place over the BEC [1], the heavy-tail Poisson distribution introduced in [24] and [48] which gives rise to capacity-achieving degree distributions for the BEC, the design of bilayer LDPC code ensembles for a degraded relay AWGN channel [4], and the design of LDPC code ensembles for unequal error protection [43].

It is well known that linear block codes which are represented by cycle-free bipartite (Tanner) graphs have poor performance even under ML decoding [12]. The bipartite graphs of capacity-approaching LDPC codes should have cycles. Hence, another question which is addressed in this paper, as a continuation to a previous study in [12] and [40] (see also [37, Problems 4.52 and 4.53]), is the following:

Question 2: How does the average cardinality of the fundamental system of cycles of bipartite graphs behave as a function of the achievable gap to capacity of the underlying LDPC code ensembles ?

The fundamental tradeoff between the graphical complexity and performance of codes defined on graphs is of interest, especially for codes of finite-length. In this paper, we address the following question:

Question 3: Consider the representation of a finite-length binary linear block code by an arbitrary bipartite graph. How simple can such a graphical representation be as a function of the channel model, target block error probability, and code rate (which is below capacity) ?

We note that the graphical complexity referred to in this paper measures the total number of edges used for the representation of finite-length codes by bipartite graphs. By referring to the total number of edges, the graphical complexity is strongly related to the decoding complexity per iteration. This differs from the graphical complexity in [3], [18], [31] and [32] which measures the number of edges per information bit in the asymptotic case where we let the block length tend to infinity. Although it may appear at first glance that the aforementioned distinction is just a matter of normalization, this is not the case: the reason is that given the target block error probability and the required gap to capacity for achieving this target with any finite-length block code, one needs first to calculate the minimal block length which potentially allows to fulfill these requirements. It is done in this work via the calculation of classical and recent sphere-packing bounds (see [44], [45], [51] and [54]).

A universal design of LDPC code ensembles which enables these codes to operate reliably over a multitude of channels is of great theoretical and practical interest. We refer the reader to recent studies on universal LDPC codes (see, e.g., [13], [30], [38] and [47]). A simple modification of the bounds derived in this paper makes them universal in the sense that they hold for the set of MBIOS channels which exhibit a given channel capacity. The universality of the bounds derived in this paper stems also from the fact that they do not depend on the full characterization of the LDPC code ensembles, but only on the gap between the channel capacity and the design rates of these ensembles, and they also depend on the target bit/ block error (or erasure) probability. The bounds derived in this work are expressed in closed form and are easily calculated.

This paper is structured as follows: Section II provides some preliminary material and notation, Section III introduces the new information-theoretic bounds of this paper, Section IV then provides their proofs followed by some discussions, and Section V formulates some algorithms related to the bounds derived in this paper, it discusses their implications, and provides numerical results. Finally, Section VI summarizes this work, and it provides some interesting open problems which are related to this research.

II. Preliminaries

We introduce in this section some preliminary material and notation which serve for the analysis in this paper.

A. LDPC Code Ensembles

LDPC codes are linear block codes which are characterized by sparse parity-check matrices. A parity-check matrix is represented by a bipartite graph where the variable and parity-check nodes are on the left and right sides of this graph, respectively. An edge connects a variable node with a parity-check node in this graph if the corresponding parity-check equation involves the code symbol which is represented by this variable node (it is illustrated in Fig. 1). The requirement for a sparse parity-check matrix is equivalent to the requirement that the number of edges in the corresponding bipartite graph scales linearly with the block length.

We move to consider ensembles of binary LDPC codes. Following standard notation, let λ_i and ρ_i denote the fraction of edges attached, respectively, to variable and parity-check nodes of degree i. Let Λ_i and Γ_i denote, respectively, the fraction of variable and parity-check nodes of degree i. The LDPC code ensemble is characterized by a triple (n, λ, ρ), where n designates the block length of the codes, and $\lambda(x) \triangleq \sum_i \lambda_i x^{i-1}$ and $\rho(x) \triangleq \sum_i \rho_i x^{i-1}$ represent, respectively, the left and right degree distributions from the edge perspective. Equivalently, this ensemble is also characterized by the triple (n, Λ, Γ) where $\Lambda(x) \triangleq \sum_i \Lambda_i x^i$ and $\Gamma(x) \triangleq \sum_i \Gamma_i x^i$ represent, respectively, the left and right degree distributions from the node perspective. We denote by LDPC(n, λ, ρ) (or LDPC(n, Λ, Γ)) the ensemble whose bipartite graphs are constructed according to the corresponding pairs of degree distributions. The connections between the edges $\mathcal{E}$ emanating from the variable nodes to the parity-check nodes are constructed by first numbering the connectors on the left and on the right sides of the graph. The number of connectors is the same on both sides of the graph, and it is equal to $|\mathcal{E}| = n \sum_i i\Lambda_i = m \sum_i i\Gamma_i$ where n and m designate the number of variable nodes and parity-check nodes, respectively. Finally, the edges which connect the variable nodes with the parity-check nodes of the bipartite graph are determined by using a permutation $\pi : \{1, \ldots, |\mathcal{E}|\} \rightarrow \{1, \ldots, |\mathcal{E}|\}$ which is chosen uniformly at random, and associates connector number i on the left side of this graph with the connector whose number is $\pi(i)$ on the right. The degree distributions with respect to the nodes and edges of a bipartite graph are related via the following equations:

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

$$\lambda(x) = \frac{\Lambda'(x)}{\Lambda'(1)}\,, \qquad \rho(x) = \frac{\Gamma'(x)}{\Gamma'(1)}\,. \tag{2}$$

For an LDPC code ensemble, whose codes are represented by parity-check matrices of dimension $m \times n$, the *design rate* is defined as $R_{\rm d} \triangleq 1 - \frac{m}{n}$. This forms a lower bound on the rate of any code from this ensemble, and the rate is equal to the design rate if the particular parity-check matrix representing this code is full rank (i.e., there are no redundant parity-check equations in this matrix). The design rate is expressed in terms of the degree distributions in the following two forms:

$$R_{\rm d} = 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx} = 1 - \frac{\Lambda'(1)}{\Gamma'(1)}\,. \tag{3}$$

Note that

$$a_{\rm L} = \Lambda'(1) = \frac{1}{\int_0^1 \lambda(x)dx} \tag{4}$$

$$a_{\rm R} = \Gamma'(1) = \frac{1}{\int_0^1 \rho(x)dx} \tag{5}$$

designate the average left and right degrees (i.e., the average degrees of the variable and parity-check nodes, respectively).

B. Functionals Related to Memoryless Binary-Input Output-Symmetric Channels

Consider an MBIOS channel whose channel input and channel output are designated by X and Y, respectively, and let $p_{Y|X}(\cdot|\cdot)$ be its transition probability. The associated log-likelihood ratio (LLR) $l(y)$ when the channel output is $Y = y$ is given by

$$l(y) = \ln\left(\frac{p_{Y|X}(y|0)}{p_{Y|X}(y|1)}\right).$$

The LLR associated with the random variable Y is defined as $L = l(Y)$. Let a designate the conditional *pdf* of the random variable L given that the channel input is $X = 0$ (to be referred as the L-density function). This density function satisfies the symmetry property $a(l) = e^l\, a(-l)$ for every $l \in \mathbb{R}$ [35].

This paper relies on the following two functionals (various other functionals are presented in [37, Section 4.1]).

Lemma 1: **[Capacity functional]** Consider an MBIOS channel whose symmetric L-density function is denoted by a. Then the capacity of this channel in units of bits per channel use, $C = C(a)$, is given by

$$C = \int_{-\infty}^{\infty} a(l)\big(1 - \log_2(1 + e^{-l})\big)\, \mathrm{d}l. \tag{6}$$

An equivalent form of the capacity is given by

$$C = \int_{0}^{\infty} a(l)(1 + e^{-l}) \left(1 - h_2\Big(\frac{1}{1+e^l}\Big)\right) \mathrm{d}l. \tag{7}$$

This lemma is proved in [37, page 193].

Definition 1: **[The Bhattacharyya functional]** The Bhattacharyya constant which is associated with the symmetric L-density function a is given by

$$\mathcal{B}(a) \triangleq \int_{-\infty}^{\infty} a(l) e^{-\frac{l}{2}}\, \mathrm{d}l. \tag{8}$$

The analysis in this paper relies partially on the *stability condition*. This condition applies to the asymptotic case where we let the block length tend to infinity, and it forms a necessary condition for successful decoding in the sense that it requires that the fixed point of zero error rate be stable. Consider an LDPC code ensemble with a given pair of degree distributions (λ, ρ) whose transmission takes place over an MBIOS channel, characterized by its L-density function a. Then, the stability condition under BP decoding gets the form (see [37, Theorem 4.125])

$$\mathcal{B}(a)\lambda_2\rho'(1) < 1. \tag{9}$$

The reader is referred to [37, Section 4.9] for a proof.

C. Lower Bound on the Conditional Entropy for Binary Linear Block Codes Transmitted over MBIOS Channels

We start this section by outlining in Section II-C.1 the derivation of a lower bound on the conditional entropy of the transmitted codeword given the received sequence at the output of an MBIOS channel. Section II-C.1 relies on [53, Section IV] and its appendices where it is assumed that the code is represented by a full-rank parity-check matrix (the same assumption is also made in [37, Section 4.11]). Section II-C.2 revisits the derivation in Section II-C.1 in order to extend the bound for the case where the binary linear block code is represented by a parity-check matrix which is not necessarily full-rank; this extension was hinted briefly in [53, Section V] (along the lines of the section on numerical results), and we take this occasion to give a rigorous proof which serves as a crucial preparatory step towards the analysis in the continuation to this paper.

1) The analysis for a full-rank parity-check matrix: We assume in the following that the transmission of a binary linear block code takes place over an MBIOS channel. Let $\mathcal{C}$ be a binary linear block code of length n and rate R, and let $\mathbf{X}$ and $\mathbf{Y}$ be the transmitted codeword and received sequence, respectively. Assume that the codewords of $\mathcal{C}$ have no bits which are set a-priori to zero. We assume that the code $\mathcal{C}$ is represented by a parity-check matrix H which is full rank. In the following, C designates the capacity of the communication channel in units of bits per channel use.

- Define an equivalent channel whose output is the LLR of the original channel.
- The LLR is represented by a pair which includes its sign and absolute value.
- For the characterization of the equivalent channel, let the function a designate the L-density function.
- We randomly generate an i.i.d. sequence $\{L_i\}_{i=1}^n$ with respect to the L-density function a, and define

$$\Omega_i \triangleq |L_i|, \quad \Theta_i \triangleq \begin{cases} 0 & \text{if } L_i > 0 \\ 1 & \text{if } L_i < 0 \\ 0 \text{ or } 1 \text{ equally likely} & \text{if } L_i = 0 \end{cases}.$$

  Note that $\{\Theta_i\}$ is a sequence which represents the signs of the LLR (conditioned on $\mathbf{X} = \mathbf{0}$).
- The output of the equivalent channel is $\widetilde{\mathbf{Y}} = (\widetilde{Y}_1, \ldots, \widetilde{Y}_n)$ where

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

  and $\Phi_i = \Theta_i + X_i$ (modulo-2 addition). This channel is memoryless.

- The output of this channel at time i is $\widetilde{Y}_i \in \{0,1\} \times \mathbb{R}_+$. Note that Φ_i is a binary random variable which is affected by the channel input X_i, and Ω_i is a non-negative random variable which is not affected by X_i.
- Due to the symmetry of the communication channel, the *pdf* of the absolute value of the LLR satisfies

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

The conditional entropy of the transmitted codeword given the received sequence at the output of the MBIOS channel satisfies

$$\begin{aligned} H(\mathbf{X}|\mathbf{Y}) &= H(\mathbf{X}|\widetilde{\mathbf{Y}}) \\ &= H(\mathbf{X}) + H(\widetilde{\mathbf{Y}}|\mathbf{X}) - H(\widetilde{\mathbf{Y}}) \\ &= nR + nH(\widetilde{Y}_1|X_1) - H(\widetilde{\mathbf{Y}}) \\ &= nR + n[H(\widetilde{Y}_1) - I(X_1;\widetilde{Y}_1)] - H(\widetilde{\mathbf{Y}}) \end{aligned} \tag{10}$$

and

$$I(X_1;\widetilde{Y}_1) = I(X_1;Y_1) \le C \tag{11}$$

$$\begin{aligned} H(\widetilde{Y}_1) &= H(\Phi_1, \Omega_1) \\ &= H(\Omega_1) + H(\Phi_1|\Omega_1) \\ &= H(\Omega_1) + 1. \end{aligned} \tag{12}$$

The last transition in (12) 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 $H(\Omega_1)$ is not expressed explicitly as it will cancel out.

The entropy of the vector $\widetilde{\mathbf{Y}}$ satisfies

$$\begin{aligned} H(\widetilde{\mathbf{Y}}) &= H\big(\Phi_1, \Omega_1, \ldots, \Phi_n, \Omega_n\big) \\ &= H(\Omega_1, \ldots, \Omega_n) + H\big(\Phi_1, \ldots, \Phi_n \mid \Omega_1, \ldots, \Omega_n\big) \\ &= nH(\Omega_1) + H\big(\Phi_1, \ldots, \Phi_n \mid \Omega_1, \ldots, \Omega_n\big). \end{aligned} \tag{13}$$

- Define the syndrome vector $\mathbf{S} \triangleq (\Phi_1, \ldots, \Phi_n)H^T$. Since H is assumed to be a full-rank parity-check matrix of $\mathcal{C}$ then $\mathbf{S} \in \{0,1\}^{n(1-R)}$, i.e., the syndrome $\mathbf{S}$ is composed of $n(1-R)$ binary components.
- Let M be the index of the vector $(\Phi_1, \ldots, \Phi_n)$ in the coset which corresponds to the syndrome $\mathbf{S}$.
- $H(M) = nR$ since all the codewords are transmitted with equal probability, and we get

$$\begin{aligned} & H\big(\Phi_1, \ldots, \Phi_n \,|\, \Omega_1, \ldots, \Omega_n\big) \\ & = H(\mathbf{S}, M \,|\, \Omega_1, \ldots, \Omega_n\big) \\ & \le H(M) + H\big(\mathbf{S} \,|\, \Omega_1, \ldots, \Omega_n\big) \\ & \le nR + \sum_{j=1}^{n(1-R)} H\big(S_j \,|\, \Omega_1, \ldots, \Omega_n\big)\,. \end{aligned} \tag{14}$$

- Since $\mathbf{X}H^T = \mathbf{0}$ for every codeword $\mathbf{X} \in \mathcal{C}$, and also $\Phi_i = X_i + \Theta_i$ for all i, then $\mathbf{S} = (\Theta_1, \ldots, \Theta_n)H^T$ is independent of the transmitted codeword.

Combining (10)–(14) gives

$$H(\mathbf{X}|\mathbf{Y}) \ge n(1-C) - \sum_{j=1}^{n(1-R)} H(S_j \big| \Omega_1, \ldots, \Omega_n) \tag{15}$$

where

- $S_j = 1$ if and only if $\Theta_i = 1$ for an odd number of indices i in the j-th parity-check equation.

- Due to the symmetry of the channel

$$P(\alpha_i) \triangleq \text{Prob}(\Theta_i = 1 \big| \Omega_i = \alpha_i) \\ = \frac{a(-\alpha_i)}{a(\alpha_i) + a(-\alpha_i)} = \frac{1}{1 + e^{\alpha_i}}.$$

In order to calculate the conditional entropy of a single component of the syndrome, the following lemma is used:

Lemma 2: If the j-th component of the syndrome $\mathbf{S}$ involves k variables whose indices are $\{i_1, \ldots, i_k\}$ then

$$\text{Prob}(S_j = 1 \big| \Omega_{i_1} = \alpha_1, \ldots, \Omega_{i_k} = \alpha_k) \\ = \frac{1}{2}\Big[1 - \prod_{m=1}^{k} \big(1 - 2P(\alpha_m)\big)\Big]$$

where

$$1 - 2P(\alpha) = \tanh\left(\frac{\alpha}{2}\right).$$

The proof of this lemma follows from [14, Lemma 4.1].

- For a parity-check node of degree k, the conditional entropy $H(S_j \big| \Omega_1, \ldots, \Omega_n)$ is equal to the k-dimensional integral

$$\int_0^\infty \ldots \int_0^\infty h_2\left(\frac{1}{2}\Big[1 - \prod_{m=1}^{k} \tanh\big(\frac{\alpha_m}{2}\big)\Big]\right) \prod_{m=1}^{k} f_\Omega(\alpha_m)\, d\alpha_1 \ldots d\alpha_k$$

  where f_Ω is the *pdf* of the absolute value of the LLR, and h_2 is the binary entropy function to the base 2.

- Using the following Taylor series expansion of h_2:

$$h_2(x) = 1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{(1-2x)^{2p}}{p(2p-1)}, \quad 0 \leq x \leq 1 \tag{16}$$

  then, for a parity-check node of degree k, the above k-dimensional integral is transformed to the following infinite sum of one-dimensional integrals (see [53, Appendix II]):

$$H(S_j \big| \Omega_1, \ldots, \Omega_n) \\ = 1 - \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \left\{ \frac{1}{p(2p-1)} \cdot \left(\int_0^\infty a(l)(1+e^{-l}) \tanh^{2p}\Big(\frac{l}{2}\Big)\, \mathrm{d}l \right)^k \right\}. \tag{17}$$

For an arbitrary full-rank parity-check matrix of a binary linear block code $\mathcal{C}$, let Γ_k designate the fraction of the parity-checks involving k variables, and let $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$. The combination of (15) and (17) leads to the following lower bound on the conditional entropy of the transmitted codeword given the received sequence at the channel output:

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq R - C + \frac{1-R}{2\ln 2} \sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)} \tag{18}$$

where

$$g_p \triangleq \int_0^\infty a(l)(1+e^{-l}) \tanh^{2p}\left(\frac{l}{2}\right) dl, \quad p \in \mathbb{N}. \tag{19}$$

The above lower bound on the conditional entropy holds for any representation of the code by a full-rank parity-check matrix. The symmetry condition for MBIOS channels states that $a(l) = e^l a(-l)$ for all $l \in \mathbb{R}$, and therefore (19) gives that

$$g_p = \mathbb{E}\left[\tanh^{2p}\left(\frac{L}{2}\right)\right], \quad p \in \mathbb{N} \tag{20}$$

where $\mathbb{E}$ designates the statistical expectation with respect to the L-density function a, and L is a random variable which stands for the LLR at the output of the channel given that the input bit is zero. Eq. (20) implies that the non-negative sequence $\{g_p\}_{p\geq 1}$ is monotonically non-increasing and it only depends on the communication channel (but not on the code). Note also that, from (20), $0 \leq g_p < 1$ for all $p \in \mathbb{N}$ (unless the channel is perfect, which then implies that $g_p = 1$ for all values of p).

We note that the conditional entropy on the LHS of (18) depends only on the code and the communication channel, but its lower bound on the RHS of (18) depends also on the specific representation of the code by a bipartite graph.

The lower bound in (18) improves the bound in [7, Eq. (15)], except for the binary symmetric channel (BSC) where they both coincide. The reason is that the derivation of (18) relies on the un-quantized soft output of the channel whereas the derivation of the bound in [7, Eq. (15)] relies on a two-level quantization of this output (which therefore does not loosen the bound for a BSC).

2) An adaptation of the analysis to LDPC codes which are not necessarily represented by full-rank parity-check matrices: The derivation of the lower bound in (18) relies on the assumption that the parity-check matrix is full rank. Though it seems like a feasible requirement for specific binary linear block codes, this poses a problem when considering ensembles of LDPC codes. In the latter case, a parity-check matrix which corresponds to a randomly chosen bipartite graph with a given pair of degree distributions may not be full rank.[1] To this end, we present the following lemma:

Lemma 3: For (regular and irregular) ensembles of binary LDPC codes, the inequality in (18) stays valid for every code from the ensemble with the following modifications:

- The rate R of the code is replaced with the design rate (R_{d}) of the ensemble.
- The sequence $\{\Gamma_k\}$ denotes the degree distribution of the parity-check nodes of the ensemble (where the representation of a code by a parity-check matrix, with the given degree distribution, possibly includes some linearly dependent rows).

Proof: See Appendix I. ∎

D. Sphere-Packing Bounds

Sphere-packing bounds are commonly used for the study of the performance limitations of finite-length error-correcting codes over memoryless symmetric channels. For a tutorial on classical sphere-packing bounds, the reader is referred to [39, Chapter 5]. This paper relies on the following sphere-packing bounds (see Section V-D):

- The *SP59 bound*: The 1959 sphere-packing (SP59) bound of Shannon [44] serves for the evaluation of the performance limits of block codes whose transmission takes place over an AWGN channel. This lower bound on the decoding error probability is expressed in terms of the block length and the rate of the code; however, it does not take into account the modulation used, but only assumes that the modulated signals have equal energy. It is often used as a reference for quantifying the sub-optimality of error-correcting codes under some practical decoding algorithms (see [39, Chapter 5] and references therein). An efficient algorithm for the calculation of the SP59 bound is introduced in [54, Section IV.C].
- The *ISP bound*: This sphere-packing bound was recently derived in [54, Section III]. The ISP bound applies to all memoryless symmetric channels. For codes of finite block length, it improves the classical sphere-packing bound of Shannon, Gallager and Berlekamp [45] and the sphere-packing bound of Valembois and Fossorier [51] where this improvement is especially pronounced for short to moderate block lengths. We note that the ISP bound in [54] is not uniformly tighter than the SP59 bound for equi-energy signals transmitted over an AWGN channel.

Comparisons between the sphere-packing bounds in [44], [51] and [54, Section III] are shown in [54, Section V].

E. Cycles in Graphs

We consider in this paper the cycles in bipartite graphs which represent capacity-approaching LDPC code ensembles. To this end, we define and exemplify some notions which are relevant to the analysis in this paper.

[1] A concentration of the code rate to the design rate of LDPC code ensembles is proved asymptotically (for an infinite block length) under some conditions (see [27] and [37, Lemma 3.22]). However, we are interested in a lower bound on the conditional entropy which also holds for finite-length binary linear block codes regardless of this asymptotic concentration property.

Definition 2: **[Cycle and cycle length]** A *cycle* in an un-directed graph is a closed path. The length of a cycle is the number of edges on this closed path. The *girth* of an un-directed graph is defined as the shortest length of its cycles.

Definition 3: **[Tree]** A *tree* is a connected graph that has no cycles.

From Definition 3, a removal of any edge from a tree makes the graph disconnected. An important property of trees is that any two vertices are connected by a single path.

Every graph $\mathcal{G}$ has subgraphs that are trees. This motivates the following definition:

Definition 4: **[Spanning tree]** A *spanning tree* of a connected graph $\mathcal{G}$ is a tree which spans all the vertices of $\mathcal{G}$. Note that by repeatedly removing edges which originally create cycles in the graph, it follows that every connected graph has a spanning tree.

Definition 5: **[Number of components of a graph]** Let $\mathcal{G}$ be a possibly disconnected graph. The *number of components* of $\mathcal{G}$ is the minimal number of its connected subgraphs whose union forms the graph $\mathcal{G}$ (clearly, a connected graph has a single component).

Definition 6: **[Cycle rank]** Let $\mathcal{G}$ be an un-directed graph with $|V_{\mathcal{G}}|$ vertices, $|E_{\mathcal{G}}|$ edges and $C(\mathcal{G})$ components. The *cycle rank* of $\mathcal{G}$, denoted by $\beta(\mathcal{G})$, is defined as the maximal number of edges which can be removed from the graph without increasing its number of components (note that each component becomes a spanning tree after the removal of these edges).

From Definition 6, the cycle rank of a graph is a measure of the edge redundancy with respect to the connectedness of this graph. The cycle rank satisfies the following equality (see [16, p. 154]):

$$\beta(\mathcal{G}) = |E_{\mathcal{G}}| - |V_{\mathcal{G}}| + C(\mathcal{G}). \tag{21}$$

Definition 7: **[Full spanning forest]** Let $\mathcal{G}$ be an un-directed graph. A *full spanning forest* $\mathcal{F}$ of the graph $\mathcal{G}$ is the subgraph of $\mathcal{G}$ that results from removing the $\beta(\mathcal{G})$ edges from Definition 6. Clearly, the number of components of $\mathcal{F}$ and $\mathcal{G}$ is the same. Note that a graph may have a multiplicity of full spanning forests.

Definition 8: **[Fundamental cycle]** Let $\mathcal{F}$ be a full spanning forest of an un-directed graph $\mathcal{G}$, and let e be an edge in the relative complement of $\mathcal{F}$. The cycle of the subgraph $\mathcal{F} \cup \{e\}$ (whose existence and uniqueness is guaranteed by [16, Theorem 3.1.11]) is called a *fundamental cycle* of $\mathcal{G}$ which is associated with $\mathcal{F}$.

Remark 1: Each of the edges in the relative complement of a full spanning forest $\mathcal{F}$ gives rise to a *different* fundamental cycle of the graph $\mathcal{G}$.

Definition 9: **[Fundamental system of cycles]** The *fundamental system of cycles* of a graph $\mathcal{G}$ which is associated with a full spanning forest $\mathcal{F}$ is the set of all fundamental cycles of $\mathcal{G}$ associated with $\mathcal{F}$.

Remark 2: From Remark 1, the cardinality of the fundamental system of cycles of $\mathcal{G}$ associated with a full spanning forest of this graph is equal to the cycle rank $\beta(\mathcal{G})$.

Example 1: **[Fundamental system of cycles in a bipartite graph]** This example refers to the bipartite graph in Fig. 1. This graph is connected, but it is clearly not a tree. As an example, consider the cycle $\langle v_9, c_4, v_{10}, c_5, v_9 \rangle$ whose length is 4. Since the number of vertices in this graph is 15 and the number of its edges is 30, then from (21), the cycle rank of this connected graph is $30 - 15 + 1 = 16$.

In order to get a spanning tree of the graph in Fig. 1, we remove repeatedly 16 edges which create cycles while preserving the connectivity of the graph.

The parity-check matrix $\widetilde{H} = [\widetilde{h}_{i,j}]$ in Fig. 2, with 16 bolded zero entries which correspond to the removed edges from the original graph in Fig. 1, represents a spanning tree of this graph. To exemplify its connectivity, note that the variable nodes v_5 and v_6 are connected by the path $\langle v_6, c_2, v_3, c_1, v_1, c_4, v_5 \rangle$ which is of length 6. This path can be observed directly from the parity-check matrix $\widetilde{H} = [\widetilde{h}_{i,j}]$ by alternate horizontal and vertical moves through the ones of $\widetilde{H}$; explicitly, this path is determined by a horizontal move from $\widetilde{h}_{2,6}$ to $\widetilde{h}_{2,3}$, a vertical move to $\widetilde{h}_{1,3}$, a horizontal move to $\widetilde{h}_{1,1}$, a vertical move to $\widetilde{h}_{4,1}$ and finally a horizontal move to $\widetilde{h}_{4,5}$. In a similar way, it can be verified that every two vertices in the bipartite graph of $\widetilde{H}$ are connected, and it spans all the 15 vertices of the graph in Fig. 1 (since there is no row or column in $\widetilde{H}$ which is a zero vector). Hence, this graph is indeed a spanning tree of the bipartite graph in Fig. 1. This spanning tree enables to obtain a set of 16 fundamental cycles by returning back a single bolded zero in Fig. 2 (among its 16 bolded zeros) to 1. For example, by setting $\widetilde{h}_{1,6} = 1$ (which is equivalent to returning the edge which connects v_6 with c_1), we get the fundamental cycle $\langle v_3, c_2, v_6, c_1, v_3 \rangle$.

$$H := \begin{pmatrix} 1&1&1&1&0&1&1&\mathbf{0}&0&0\\ \mathbf{0}&\mathbf{0}&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{0}&\mathbf{0}\\ 0&1&0&1&0&1&0&\mathbf{1}&1&1\\ 1&0&1&0&1&0&0&\mathbf{1}&1&1\\ 1&1&0&0&1&0&1&\mathbf{0}&1&1 \end{pmatrix}$$

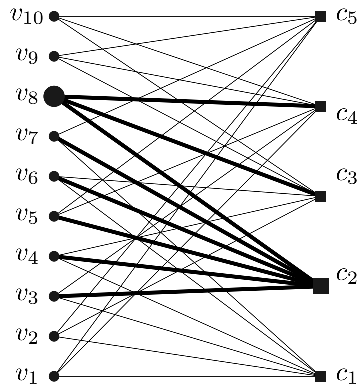

Fig. 1. A parity-check matrix H and the corresponding bipartite graph. For illustrating this relationship, column 8 and row 2 of H are bolded; the corresponding variable and parity-check nodes, and the attached edges are also bolded (this figure appears in [36]).

$$\widetilde{H} = \begin{pmatrix} 1&1&1&\mathbf{0}&0&\mathbf{0}&\mathbf{0}&0&0&0\\ 0&0&1&1&\mathbf{0}&1&1&1&0&0\\ 0&\mathbf{0}&0&\mathbf{0}&0&\mathbf{0}&0&\mathbf{0}&1&\mathbf{0}\\ 1&0&\mathbf{0}&0&1&0&0&\mathbf{0}&\mathbf{0}&\mathbf{0}\\ 1&\mathbf{0}&0&0&\mathbf{0}&0&\mathbf{0}&0&1&1 \end{pmatrix}$$

Fig. 2. A parity-check matrix which corresponds to a spanning tree of the bipartite graph in Fig. 1. As compared to the parity-check matrix H in Fig. 1, the new parity-check matrix $\widetilde{H}$ is obtained by changing the values of the bolded 16 entries from 1 to 0.

F. Notation

We consider in this paper sequences of capacity-approaching LDPC code ensembles, and refer to the case where the fractional gap (in rate) to capacity (ε) vanishes. Accordingly, based on standard notation [56], we define

- $f(\varepsilon) = O\big(g(\varepsilon)\big)$ means that there are positive constants c and δ, such that $0 \leq f(\varepsilon) \leq c\, g(\varepsilon)$ for all $0 \leq \varepsilon \leq \delta$.
- $f(\varepsilon) = \Omega\big(g(\varepsilon)\big)$ means that there are positive constants c and δ, such that $0 \leq c\, g(\varepsilon) \leq f(\varepsilon)$ for all $0 \leq \varepsilon \leq \delta$.

Note that the values of c and δ must be fixed, and should not depend on ε.

Throughout the paper

$$h_2(x) \triangleq -x \log_2(x) - (1-x) \log_2(1-x), \quad 0 \leq x \leq 1$$

denotes the binary entropy function to the base 2, and $h_2^{-1} : [0,1] \to \big[0, \frac{1}{2}\big]$ is the inverse of the restriction of h_2 to $\big[0, \frac{1}{2}\big]$. We also denote the block error probability and the bit error probability of a code by P_{B} and P_{b}, respectively (for the BEC, the error probability is replaced with an erasure probability). Note that P_{b} refers to the bit error probability of the information bits.

This paper is focused on the analysis for MBIOS channels. For basic definitions and examples of MBIOS channels, the reader is referred to [37, Section 4.1] (which uses a slightly different abbreviation: BMS channels).

For further notation used throughout this paper, Section II-A provides the setting and notation for the degree distributions and the design rate of LDPC code ensembles, Section II-B provides the notation for the capacity and Bhattacharyya functionals, Section II-C presents the notation for the lower bound on the conditional entropy (see (18)–(20)), and Section II-E provides the terminology and notation used here in the context of cycles in bipartite graphs.

III. NEW INFORMATION-THEORETIC BOUNDS

This section introduces information-theoretic bounds which are related to the degree distributions, graphical complexity, and the number of fundamental systems of cycles in the bipartite graphs of LDPC code ensembles.

Theorem 1: **[On the average degree of the parity-check nodes]** Let $\mathcal{C}$ be a binary linear block code of block length n whose transmission takes place over an MBIOS channel. Let $\mathcal{G}$ be a bipartite graph which corresponds to a full-rank parity-check matrix of $\mathcal{C}$. Let C designate the capacity of the channel, in bits per channel use, and a be the L-density function of this channel. Assume that the code rate is (at least) a fraction $1-\varepsilon$ of the channel capacity (where $0<\varepsilon<1$), and the code achieves a block error probability $P_{\rm B}$ or a bit error probability $P_{\rm b}$ under some decoding algorithm. Then, the average right degree of the bipartite graph (i.e., the average degree of the parity-check nodes in $\mathcal{G}$) satisfies

$$a_{\rm R} \geq \frac{2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-\delta}{1-(1-\varepsilon)C}\right)}\right)}{\ln\left(\frac{1}{g_1}\right)} \tag{22}$$

where g_1 is given in (19) (and it depends only on the channel), and

$$\delta \triangleq \begin{cases} P_{\rm B} + \frac{h_2(P_{\rm B})}{n} & \text{for a block error probability } P_{\rm B} \\ h_2(P_{\rm b}) & \text{for a bit error probability } P_{\rm b} \end{cases}. \tag{23}$$

Furthermore, among all the MBIOS channels which exhibit a given capacity C and for which a target block error probability $(P_{\rm B})$ or a bit error probability $(P_{\rm b})$ is obtained under some decoding algorithm, a universal lower bound on $a_{\rm R}$ holds by replacing g_1 on the RHS of (22) with C.

For the BEC, the following tightened version of (22) holds:

$$a_{\rm R} \geq \frac{\ln\left(1+\frac{p-P_{\rm b}}{(1-p)\varepsilon+P_{\rm b}}\right)}{\ln\left(\frac{1}{1-p}\right)} \tag{24}$$

where p is the erasure probability of the channel, and $P_{\rm b}$ is the bit erasure probability at the decoder.

Remark 3: **[The relation of Theorem 1 to the bound in [53]]** In the particular case where $P_{\rm b}$ vanishes, the bound in (22) forms a tightened version of the bound given in [53, Eq. (77)]. This point is clarified in Discussion 1 which succeeds the proof of Theorem 1 (see page 17). In the limit where the gap (in rate) to capacity vanishes (and with vanishing $P_{\rm b}$), the lower bounds on the average right degree in (22) and [53, Eq. (77)] both grow like the logarithm of the inverse of this gap, and they therefore possess the same asymptotic behavior where

$$a_{\rm R} \triangleq a_{\rm R}(\varepsilon) = \Omega\left(\ln\frac{1}{\varepsilon}\right). \tag{25}$$

However, in spite of the similarity in the asymptotic behavior of the two lower bounds as $\varepsilon \rightarrow 0$, they may differ significantly even for rather small values of ε (see Example 3 on p. 27).

Theorem 1 also provides a universal lower bound on the average right degree for the set of all MBIOS channels with a given capacity C. This theorem states the conditions where the bound in (22) gets its extreme values among all MBIOS channels which exhibit a given capacity.

Remark 4: **[Adaptation of Theorem 1 to LDPC code ensembles]** As is clarified in Discussion 2 (see page 17), Theorem 1 can be adapted to hold for an arbitrary ensemble of (n,λ,ρ) LDPC codes. In this case, the requirement of a full-rank parity-check matrix of a particular code $\mathcal{C}$ from this ensemble is relaxed by requiring that the design rate of the LDPC code ensemble is equal to a fraction $1-\varepsilon$ of the channel capacity. In this case, $P_{\rm b}$ and $P_{\rm B}$ stand for the average bit and block error (or erasure) probabilities of the ensemble under some decoding algorithm.

Remark 5: **[The graphical complexity of finite-length LDPC codes]** In Section V-D, we apply Theorem 1 and sphere-packing bounds on the decoding error probability (see [44], [45], [51], [54]) to obtain information-theoretic lower bounds on the graphical complexity of finite-length LDPC codes. These bounds are expressed as a function of the target block error probability and the gap between the design rate of the code and the channel capacity. We note that in this context, the graphical complexity measures the number of edges used for the representation of finite-length codes by bipartite graphs. By referring to the total number of edges, the graphical complexity is strongly related to the decoding complexity per iteration. The bounds are compared with capacity-approaching LDPC code ensembles under BP decoding, and they are shown to be informative (see Section V-D).

Based on Remark 4 and the background which is provided in Section II-E, the following result is derived:

Corollary 1: **[On the asymptotic average cardinality of the fundamental system of cycles of LDPC code ensembles]** Let $\big\{(n,\lambda,\rho)\big\}$ be a sequence of LDPC code ensembles whose transmission takes place over an MBIOS channel. Let the design rate of these ensembles be a fraction $1-\varepsilon$ of the channel capacity C, and assume that the average bit error/ erasure probability of this sequence vanishes under some decoding algorithm as we let the block length (n) tend to infinity. Consider the average cardinality of the fundamental system of cycles in bipartite graphs from the LDPC code ensemble (n, λ, ρ) where the graphs are chosen uniformly at random (from Remark 2, the cardinality of the fundamental system of cycles in a graph $\mathcal{G}$ is equal to its cycle rank $\beta(\mathcal{G})$). Then, the following asymptotic lower bound holds:

$$\liminf_{n\to\infty} \frac{\mathbb{E}_{\text{LDPC}(n,\lambda,\rho)}\big[\beta(\mathcal{G})\big]}{n} \geq \frac{(1-C)\ \ln\left(g_1\ \left[1-2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)\right]^{-2}\right)}{\ln\left(\frac{1}{g_1}\right)} - 1 \tag{26}$$

where g_1 is introduced in (19). For a BEC whose erasure probability is p, a tightened bound gets the form:

$$\liminf_{n\to\infty} \frac{\mathbb{E}_{\text{LDPC}(n,\lambda,\rho)}\big[\beta(\mathcal{G})\big]}{n} \geq \frac{p\ \ln\left(1-p+\frac{p}{\varepsilon}\right)}{\ln\left(\frac{1}{1-p}\right)} - 1. \tag{27}$$

Remark 6: Corollary 1 provides two results which are of the type $\Omega\left(\ln\frac{1}{\varepsilon}\right)$.

Theorem 2: **[On the degree distributions of capacity-approaching LDPC code ensembles]** Let (n, λ, ρ) (or (n, Λ, Γ)) be an ensemble of LDPC codes whose transmission takes place over an MBIOS channel. Assume that the design rate of the ensemble is equal to a fraction $1-\varepsilon$ of the channel capacity C, and let P_{b} designate the average bit error (or erasure) probability of the ensemble under ML decoding or any sub-optimal decoding algorithm. Then, the following properties hold for an arbitrary finite (and fixed) degree i

$$\Lambda_i(\varepsilon) = O(1) \tag{28}$$

$$\Gamma_i(\varepsilon) = O\big(\varepsilon C + h_2(P_{\text{b}})\big) \tag{29}$$

$$\lambda_i(\varepsilon) = O\left(\frac{1}{\ln\frac{1}{\varepsilon C + h_2(P_{\text{b}})}}\right) \tag{30}$$

$$\rho_i(\varepsilon) = O\left(\frac{\varepsilon C + h_2(P_{\text{b}})}{\ln\frac{1}{\varepsilon C + h_2(P_{\text{b}})}}\right). \tag{31}$$

For the case where the transmission takes place over the BEC, the bounds above are tightened by replacing $h_2(P_{\text{b}})$ with P_{b}.

Remark 7: **[On the connection between Theorems 1 and 2]** Theorem 2 implies that for every capacity-approaching LDPC code ensemble whose bit error probability vanishes and also for an arbitrary finite degree i in their bipartite graphs, the fraction of edges attached to variable nodes or parity-check nodes of degree i tends to zero as the gap to capacity (ε) vanishes. This conclusion is consistent with Theorem 1 which states that the average left and right degrees of the bipartite graphs scale at least like $\ln\frac{1}{\varepsilon}$; hence, these average degrees necessarily become unbounded as the gap to capacity vanishes.

Corollary 2: Under the assumptions of Theorem 2, if the asymptotic bit error/ erasure probability vanishes then the following properties hold for an arbitrary finite degree i

$$\Lambda_i = O(1)\,, \qquad \Gamma_i = O(\varepsilon)\,,$$
$$\lambda_i = O\left(\frac{1}{\ln\frac{1}{\varepsilon}}\right), \quad \rho_i = O\left(\frac{\varepsilon}{\ln\frac{1}{\varepsilon}}\right).$$

Remark 8: **[Linear programming upper bounds on the degree distributions of LDPC code ensembles]** Theorem 2 and Corollary 2 provide asymptotic results for the degree distributions of LDPC code ensembles in the limit where the gap to capacity vanishes (i.e., $\varepsilon \to 0$). Section V-C provides linear programming (LP) upper bounds on the degree distributions which are expressed in terms of the target average bit error probability, and the (possibly non-zero) gap between the channel capacity and the design rate of the ensemble for achieving this target. Similarly to Theorem 2 and Corollary 2, the LP bounds in Section V-C hold under ML decoding, and are therefore general in terms of the decoding algorithm. We note that these LP bounds apply to finite-length LDPC code ensembles and to the asymptotic case of an infinite block length. Analytical solutions for these LP bounds are provided in Section V-C, and these bounds are also compared with some capacity-achieving sequences of LDPC code ensembles for the BEC under BP decoding. Additional LP bounds are derived to hold for the set of all the MBIOS channels which exhibit a given capacity, and that also achieve a target bit error probability. These universal LP bounds are compared with the LP bounds which refer to specific MBIOS channels (see Section V-C).

We turn now our attention to sequences of LDPC code ensembles which asymptotically achieve vanishing bit error probability under BP decoding. The following theorem gives upper bounds on the fraction of degree-2 variable nodes (Λ_2) and the fraction of edges attached to these nodes (λ_2) for an arbitrary sequence of LDPC code ensembles whose transmission takes place over an MBIOS channel. It relies on information-theoretic arguments and the stability condition. We note that λ_2 is involved in the stability condition (see (9)). Moreover, some previously reported information-combining bounds on the performance of LDPC code ensembles under BP decoding are sensitive to the value of λ_2 (see, e.g., [49]).

Theorem 3: **[On the fraction of degree-2 variable nodes and the fraction of edges attached to these nodes for LDPC code ensembles]** Let $\left\{\left(n_m, \lambda(x), \rho(x)\right)\right\}_{m\geq 1}$ be a sequence of LDPC code ensembles whose transmission takes place over an MBIOS channel. Assume that this sequence asymptotically achieves a fraction $1-\varepsilon$ of the channel capacity under BP decoding with vanishing bit error probability. Then, the fraction of degree-2 variable nodes satisfies

$$\Lambda_2 < \frac{1-C}{2\,\mathcal{B}(a)}\left(1+\frac{\varepsilon C}{1-C}\right) \cdot \left[1+\frac{\ln\left(\frac{1}{g_1}\right)}{\ln\left(\frac{g_1}{\left[1-2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)\right]^2}\right)}\right] \quad (32)$$

and the fraction of edges attached to these nodes satisfies

$$\lambda_2 < \frac{\ln\left(\frac{1}{g_1}\right)}{\mathcal{B}(a)\,\ln\left(\frac{g_1}{\left[1-2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)\right]^2}\right)} \quad (33)$$

where the Bhattacharyya constant $\mathcal{B}(a)$ and the parameter g_1 are introduced in (8) and (19), respectively. Consider the set of all the MBIOS channels with a given capacity C and a Bhattacharyya constant $\mathcal{B}(a)$, for which the bit error probability vanishes under BP decoding. Then, universal upper bounds on Λ_2 and λ_2 hold for this set of channels by replacing g_1 on the RHS of (32) and (33), respectively, with C.

For a BEC with an erasure probability p, the following tightened bounds hold:

$$\Lambda_2 < \frac{1}{2}\left(1+\frac{\varepsilon(1-p)}{p}\right)\left[1+\frac{\ln\left(\frac{1}{1-p}\right)}{\ln\left(1-p+\frac{p}{\varepsilon}\right)}\right] \quad (34)$$

and

$$\lambda_2 < \frac{\ln\left(\frac{1}{1-p}\right)}{p\,\ln\left(1-p+\frac{p}{\varepsilon}\right)}. \quad (35)$$

Corollary 3: Under the assumptions of Theorem 3, in the limit where the gap to capacity vanishes under BP decoding (i.e., $\varepsilon \to 0$), the fraction of degree-2 variable nodes satisfies

$$\Lambda_2 \leq \frac{1-C}{2\,\mathcal{B}(a)} \quad (36)$$

where this upper bound is necessarily not larger than $\frac{1}{2}$. Note that this forms a universal upper bound on the fraction of degree-2 variable nodes for all MBIOS channels with a given capacity C and a Bhattacharyya constant $\mathcal{B}(a)$ for which the bit error probability vanishes under BP decoding, and for which the gap to capacity vanishes.

In the continuation to this paper, sufficient conditions for the tightness of (36) are considered (see Lemma 7 on page 24).

Remark 9: Note that for capacity-achieving sequences of LDPC code ensembles whose transmission takes place over the BEC, the bound in (36) is particularized to $\frac{1}{2}$ *regardless of the erasure probability of this channel*. This is indeed the case for some sequences of LDPC code ensembles which achieve the capacity of the BEC under BP decoding (see, e.g., [24], [29], [48]).

Corollary 4: **[A looser and simpler version of the upper bound on λ_2]** The bound (33) implies that

$$\lambda_2 < \frac{1}{\left[c_1 + c_2 \ln\left(\frac{1}{\varepsilon}\right)\right]^+} \tag{37}$$

for some constants c_1 and c_2 which only depend on the MBIOS channel, and where $[x]^+ \triangleq \max(x, 1)$; the coefficient c_2 of the logarithm in (37) is given by

$$c_2 = \frac{\mathcal{B}(a)}{\ln\left(\frac{1}{g_1}\right)} \tag{38}$$

and it is strictly positive.

In the following proposition, it is shown that for the BEC, the bounds in (35) and (37) are tight under BP decoding.

Proposition 1: **[On the tightness of the upper bound on λ_2 for capacity-achieving sequences of LDPC code ensembles over the BEC]** The bounds in (35) and (37) are tight for the capacity-achieving sequence of right-regular LDPC code ensembles over the BEC in [48]. For this sequence, $\lambda_2 \triangleq \lambda_2(\varepsilon)$ vanishes as $\varepsilon \to 0$ similarly to the upper bound in (37) with the same coefficient c_2 in (38).

IV. Proofs and Discussions

A. Proof of Theorem 1

Let $\mathbf{X}$ be a random codeword from the binary linear block code $\mathcal{C}$, and let $\mathbf{Y}$ designate the output of the communication channel when $\mathbf{X}$ is transmitted. Based on the assumption that the code $\mathcal{C}$ is represented by a full-rank parity-check matrix and $\mathcal{G}$ is the corresponding bipartite graph which represents this code, then inequality (18) holds. Since $f(t) = x^t$ is convex for any $x \geq 0$ then Jensen's inequality gives

$$\Gamma(x) = \sum_i \Gamma_i x^i \geq x^{\sum_i i \Gamma_i} = x^{a_{\text{R}}}, \quad x \geq 0\,.$$

Substituting the inequality above in (18) implies that

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq R - C + \frac{1-R}{2\ln 2} \sum_{k=1}^{\infty} \frac{g_k^{a_{\text{R}}}}{k(2k-1)}\,. \tag{39}$$

Lemma 4:

$$g_k \geq (g_1)^k\,, \quad \forall k \in \mathbb{N}. \tag{40}$$

Proof: For $k \geq 1$, Jensen's inequality and (20) give

$$\begin{aligned} g_k &= \mathbb{E}\left[\tanh^{2k}\left(\frac{L}{2}\right)\right] \\ &\geq \left(\mathbb{E}\left[\tanh^{2}\left(\frac{L}{2}\right)\right]\right)^k \\ &= (g_1)^k\,. \end{aligned}$$

■

The substitution of (40) in (39) gives

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq R - C + \frac{1-R}{2\ln 2} \sum_{k=1}^{\infty} \frac{(g_1^{a_{\mathrm{R}}})^k}{k(2k-1)}\,. \tag{41}$$

The substitution $x = \frac{1-\sqrt{u}}{2}$ in (16) gives

$$\frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{u^k}{k(2k-1)} = 1 - h_2\left(\frac{1-\sqrt{u}}{2}\right), \ \forall\, u \in [0,1]. \tag{42}$$

Since $0 \leq \tanh^2(x) < 1$ for all $x \in \mathbb{R}$, we get from (20) that $0 \leq g_1 \leq 1$ (this property holds for the entire sequence $\{g_k\}_{k=1}^{\infty}$). Substituting (42) into (41) gives the following lower bound on the conditional entropy:

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq 1 - C - (1-R)\, h_2\left(\frac{1-g_1^{a_{\mathrm{R}}/2}}{2}\right). \tag{43}$$

On the other hand, Fano's inequality provides the upper bound

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \leq \begin{cases} R\,P_{\mathrm{B}} + \frac{h_2(P_{\mathrm{B}})}{n} \\ R\,h_2(P_{\mathrm{b}}) \end{cases} \tag{44}$$

where, for the bound which is expressed in terms of the bit error probability P_{b}, one can assume without any loss of generality that the first nR bits of the code are its information bits, and their knowledge is sufficient for determining the codeword.

In order to make the statement also valid for code ensembles (to be clarified in Discussion 2), we rely on the inequality $R \leq 1$, and loosen the bound in (44) to get

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \leq \delta \tag{45}$$

where δ is introduced in (23). Combining (43) and (45) gives

$$\delta \geq 1 - C - (1-R)\, h_2\left(\frac{1-g_1^{a_{\mathrm{R}}/2}}{2}\right). \tag{46}$$

Since the RHS of (46) is monotonically increasing in R, then following our assumption that $R \geq (1-\varepsilon)C$, the bound is loosened by replacing R with $(1-\varepsilon)C$. This gives the inequality

$$h_2\left(\frac{1-g_1^{a_{\mathrm{R}}/2}}{2}\right) \geq \frac{1-C-\delta}{1-(1-\varepsilon)C}\,.$$

Since the binary entropy function h_2 is monotonically increasing on $[0, \frac{1}{2}]$ then

$$g_1^{\frac{a_{\mathrm{R}}}{2}} \leq 1 - 2h_2^{-1}\left(\frac{1-C-\delta}{1-(1-\varepsilon)C}\right)$$

which gives the lower bound on a_{R} in (22).

Let us now consider the particular case where the transmission is over the BEC. Note that for a BEC with erasure probability p, $g_k = 1-p$ for all $k \in \mathbb{N}$ (in this case we have $L \in \{0, +\infty\}$ with probabilities p and $1-p$, respectively, and the equality $\tanh(+\infty) = 1$ is exploited in (20)). Therefore (39) is particularized to

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq R - C + \frac{(1-R)(1-p)^{a_{\mathrm{R}}}}{2\ln 2} \sum_{k=1}^{\infty} \frac{1}{k(2k-1)}\,.$$

Substituting $u = 1$ in (42) gives the equality

$$\frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{1}{k(2k-1)} = 1 \tag{47}$$

and

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq R - C + (1-R)(1-p)^{a_{\text{R}}}\,. \tag{48}$$

Note that the RHS of (48) is monotonic increasing as a function of the rate R. Following the assumption that $R \geq (1-\varepsilon)C$ where $C = 1-p$ is the capacity of the BEC, we get

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq -\varepsilon(1-p) + \big(1-(1-\varepsilon)(1-p)\big)(1-p)^{a_{\text{R}}}\,. \tag{49}$$

Similarly to (44) and (45), we get for the BEC

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \leq P_{\text{b}} \tag{50}$$

where the decoder finds X_i with probability $1-P_{\text{b}}$; otherwise, the bit X_i is not determined by the decoder, and its conditional entropy (given the sequence $\mathbf{Y}$) is upper bounded by 1 bit. Combining (49) with (50) gives

$$P_{\text{b}} \geq -\varepsilon(1-p) + \big(1-(1-\varepsilon)(1-p)\big)(1-p)^{a_{\text{R}}}\,. \tag{51}$$

Finally, the lower bound on the average right degree in (24) follows from (51) by simple algebra. Note that in the case where $P_{\text{b}} = 0$, the resulting lower bound coincides with the result obtained in [40, p. 1619] (though it was derived there in a different way), and it gets the form

$$a_{\text{R}} \geq \frac{\ln\left(1+\frac{p}{(1-p)\varepsilon}\right)}{\ln\left(\frac{1}{1-p}\right)}. \tag{52}$$

We wish now to show that among all the MBIOS channels which exhibit a given capacity C, the lower bound on the average degree of the parity-check nodes as given in (22) attains its maximal and minimal values for a BSC and BEC, respectively.

Lemma 5: **[Extreme values of g_1 among all MBIOS channels with a given capacity]** Among all the MBIOS channels with a given capacity C, the value of g_1 satisfies

$$C \leq g_1 \leq \big(1-2h_2^{-1}(1-C)\big)^2 \tag{53}$$

and these upper and lower bounds on g_1 are attained for a BSC and BEC, respectively.

Proof: See Appendix II. ■

Remark 10: This lemma is in fact equivalent to the statement in [20, Theorem 1] with the extreme values derived in its proof (note that (20) implies that the sequence $\{g_k\}$ is equal to the sequence $\{m_{2k}\}$ in [20], from which the equivalence between Lemma 5 and [20, Theorem 1] follows directly). In Appendix II, we present an alternative proof which is more elementary.[2]

Remark 11: The ratio between the upper and lower bounds on g_1 (see Lemma 5) is equal to $\eta(C) = \frac{\left(1-2h_2^{-1}(1-C)\right)^2}{C}$. Based on (42), one can verify that η is a monotonic decreasing function of the capacity where it tends to $2\ln 2 \approx 1.386$ when $C \to 0$, and it is 1 (i.e., the upper and lower bounds coincide) for $C = 1$.

Consider the set of all MBIOS channels with a given capacity C for which a target block error probability $(P_{\text{B}}))$ or bit error probability (P_{b}) is obtained under some decoding algorithm. To complete the proof of the last statement in Theorem 1, note that among this set of channels, the lower bound in (22) is maximized or minimized by maximizing or minimizing the value of g_1, respectively. It therefore follows from Lemma 5 that a universal bound on a_{R} for the above set of channels holds by replacing g_1 on the RHS of (22) with C. The gives the following universal lower bound:

$$a_{\text{R}} \geq \frac{2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-\delta}{1-(1-\varepsilon)C}\right)}\right)}{\ln\left(\frac{1}{C}\right)}. \tag{54}$$

[2]The author was un-aware of [20] until its publication as a journal paper. The alternative proof on Lemma 5 was found independently of this work.

Discussions on Theorem 1 via its Proof

In the following we discuss Theorem 1 via its proof, and consider some of the generalizations of this theorem.

Discussion 1: **[A discussion on the bounds in Theorem 1 and [53, Eq. (77)]]** If the bit error probability vanishes, the lower bound in (22) forms a tightened version of [53, Eq. (77)]. We note that both bounds are based on (18) but the difference in their derivation follows since [53] relies on the fact that the RHS of (18) is an infinite sum of non-negative terms, and a simple lower bound is obtained in [53] by truncating this sum after its first term. In the proof of Theorem 1, on the other hand, a tightened lower bound on the average right degree (a_{R}) is derived by applying Jensen's inequality to the RHS of (18) (see (41)), and calculating exactly the resulting bound via (42). In this context, see Remark 3 on page 11. The additional dependence of the bound in (22) on P_{b} makes Theorem 1 valid for codes of finite block length, whereas the bound in [53, Eq. (77)] can be only applied to the asymptotic case of vanishing bit error (or erasure) probability by letting the block length tend to infinity.

Discussion 2: **[An adaptation of Theorem 1 for LDPC code ensembles]** The statement in Theorem 1 can be adapted for finite-length LDPC code ensembles whose transmission takes place over an MBIOS channel. First, from Section II-C.2, the lower bound on the conditional entropy (18) holds for every code from this ensemble if we relax the requirement of a full-rank parity-check matrix, and instead replace the rate R of the code by the design rate R_{d} of the ensemble. Similarly to the derivation of (43), we get

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq 1 - C - (1-R_{\text{d}})\, h_2\left(\frac{1-g_1^{a_{\text{R}}/2}}{2}\right).$$

Assume that $R_{\text{d}} \geq (1-\varepsilon)C$. Since the RHS of the above inequality is monotonic increasing with R_{d}, then for every code in this ensemble

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq 1 - C - \big(1-(1-\varepsilon)C\big)\, h_2\left(\frac{1-g_1^{a_{\text{R}}/2}}{2}\right). \tag{55}$$

Note that this lower bound on the conditional entropy is global in the sense that it does not depend on the code from the (n, λ, ρ) LDPC code ensemble; all these codes are represented by bipartite graphs whose common value of a_{R} is equal to $\left(\int_0^1 \rho(x)\, dx\right)^{-1}$. Note also that the parameter g_1 does not depend on the code. Taking the expectation over the LDPC code ensemble gives

$$\mathbb{E}\left[\frac{H(\mathbf{X}|\mathbf{Y})}{n}\right] \geq 1 - C - \big(1-(1-\varepsilon)C\big)\, h_2\left(\frac{1-g_1^{a_{\text{R}}/2}}{2}\right). \tag{56}$$

Note that $0 \leq g_1 < 1$ (unless $g_1 = 1$ when the capacity of the binary-input channel is 1 bit per channel use which implies that the channel is noiseless).

The loosening of the bound in the transition from (44) to (45) is due to the fact that an upper bound on the rate R of a code from this ensemble is required; since binary codes are considered, a trivial upper bound on the rate is 1 bit per channel use (note that the rate of an arbitrarily chosen code from this ensemble may exceed the channel capacity). Due to the concavity of the binary entropy function, Jensen's inequality gives

$$\mathbb{E}\left[\frac{H(\mathbf{X}|\mathbf{Y})}{n}\right] \leq \begin{cases} \overline{P_{\text{B}}} + \frac{h_2(\overline{P_{\text{B}}})}{n} \\ h_2(\overline{P_{\text{b}}}) \end{cases} \tag{57}$$

where $\overline{P_{\text{B}}} \triangleq \mathbb{E}\big[P_{\text{B}}\big]$ and $\overline{P_{\text{b}}} \triangleq \mathbb{E}\big[P_{\text{b}}\big]$ designate the average block and bit error probabilities, respectively, of the ensemble. Combining (56) and (57) leads to an adaptation of Theorem 1 for LDPC code ensembles with the following modifications:

- The parity-check matrices of the codes are not required to be full-rank (which otherwise would be problematic for LDPC code ensembles).
- The requirement on the rate a code is replaced by the same requirement on the design rate of the LDPC code ensemble where we refer to the average block and bit error probabilities of this ensemble.

Note that the adaptation of the statement in Theorem 1 for LDPC code ensembles whose transmission takes place over the BEC is more direct. For a BEC, since $h_2(P_{\text{b}})$ on the LHS of (46) is replaced by P_{b} on the LHS of (51), then there is no need for Jensen's inequality as in (57).

Discussion 3: **[Adaptation of Theorem 1 for punctured LDPC code ensembles]** In the following, we consider an adaptation of Theorem 1 for LDPC code ensembles with random or intentional puncturing where the transmission takes place over an MBIOS channel. To this end, the reader is referred to [41, Section V] where lower bounds are derived on the average right degree and the graphical complexity of such ensembles. The derivation of these bounds relies on a lower bound [41, Eqs. (2) and (3)] which generalizes (18) to the case of statistically independent parallel MBIOS channels. This lower bound was particularized in [41, Sections II–IV] for the two settings of randomly and intentionally punctured LDPC code ensembles which are communicated over a single MBIOS channel. The concept of the proof of Theorem 1 enables to tighten the lower bounds on the average right degree and the graphical complexity, as presented in [41, Section V], for both randomly and intentionally punctured LDPC code ensembles. More explicitly, by comparing the proof of (22) with the derivation of [53, Eq. (77)] under the assumption of vanishing bit error probability, one notices that the tightening of the bound in the former case is enabled by combining Lemma 4 with the equality in (42) (instead of the truncation of a non-negative infinite series after its first term, as was done for the derivation of the looser bound in [53]). This difference can be exploited exactly in the same way in connection with the results from [41, Section V] for improving the tightness of the lower bounds on the average right degree and the graphical complexity for punctured LDPC code ensembles.

Proof of Corollary 1

The following lemma relies on the background material in Section II-E, and it serves for proving Corollary 1.

Lemma 6: **[Cardinality of the fundamental system of cycles]** Under the assumptions of Theorem 1, the cardinality of the fundamental system of cycles of a bipartite graph $\mathcal{G}$, associated with a full spanning forest of $\mathcal{G}$, is larger than

$$n\big[(1-R)(a_{\text{R}}-1)-1\big] \tag{58}$$

where a_{R} can be replaced by the lower bounds in (22) and (24) for a general MBIOS channel and a BEC, respectively. From (25), the cardinality of the fundamental system of cycles of the bipartite graph $\mathcal{G}$ which is associated with a full spanning forest of this graph is $\Omega\left(\ln \frac{1}{\varepsilon}\right)$.

Proof: From Remark 2 (see Section II-E), the cardinality of the fundamental system of cycles of a bipartite graph $\mathcal{G}$, which is associated with a full spanning forest of $\mathcal{G}$, is equal to the cycle rank $\beta(\mathcal{G})$. From Eq. (21), $\beta(\mathcal{G}) > |E_{\mathcal{G}}| - |V_{\mathcal{G}}|$ where $|E_{\mathcal{G}}|$ and $|V_{\mathcal{G}}|$ designate the number of edges and vertices. Specializing this for a bipartite graph $\mathcal{G}$ which represents a full-rank parity-check matrix of a binary linear block code, the number of vertices satisfies $|V_{\mathcal{G}}| = n(2-R)$ (since there are n variable nodes and $n(1-R)$ parity-check nodes in the graph) and the number of edges satisfies $|E_{\mathcal{G}}| = n(1-R)a_{\text{R}}$. Combining these equalities gives the lower bound on the cardinality of the fundamental system of cycles in (58). ∎

The proof of (26) and (27) is based on Remark 4 and Lemma 6. By substituting $P_{\text{b}} = 0$ in (22), one obtains the following lower bound on the average right degree as the average bit error probability of the LDPC code ensemble vanishes:

$$a_{\text{R}} \geq \frac{2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)}\right)}{\ln\left(\frac{1}{g_1}\right)}. \tag{59}$$

Since the average bit error probability of the ensemble is assumed to vanish as the block length tends to infinity, then asymptotically with probability 1, the code rate of an arbitrary code from the considered ensemble does not exceed the channel capacity. By substituting the lower bound on a_{R} from (59) and an upper bound on R (i.e., $R \leq C$) into (58), the asymptotic result in (26) follows readily. A similar proof of the tightened bound for the BEC in (27) follows by substituting $P_{\text{b}} = 0$ in (24). This concludes the proof of Corollary 1.

B. Proof of Theorem 2

Eq. (28) is trivial (though it is demonstrated in the continuation that, for degree-2 variable nodes, this result is asymptotically tight as the gap to capacity vanishes).

We turn now to consider the degrees of the parity-check nodes. Similarly to Discussion 2 (which succeeds the proof of Theorem 1), we denote by $\mathbf{X}$ a random codeword from the LDPC code ensemble (n, λ, ρ) where the randomness is over the selected code from the ensemble and the codeword which is selected from the code. Let $\mathbf{Y}$ designate the output of the communication channel when $\mathbf{X}$ is transmitted. From (18) and its adaptation to LDPC code ensembles (see Section II-C.2)

$$\begin{aligned}&\frac{H(\mathbf{X}|\mathbf{Y})}{n}\\&\geq R_{\text{d}} - C + \frac{1-R_{\text{d}}}{2\ln 2}\sum_{k=1}^{\infty}\frac{\Gamma(g_k)}{k(2k-1)}\\&= -\varepsilon\, C + \frac{1-(1-\varepsilon)C}{2\ln 2}\sum_{i=1}^{\infty}\left\{\Gamma_i \sum_{k=1}^{\infty}\frac{g_k^i}{k(2k-1)}\right\}\end{aligned} \tag{60}$$

where the last equality follows from the equality $\Gamma(x) = \sum_i \Gamma_i x^i$ (see Section II-A) and also since, by assumption, the design rate of the LDPC code ensemble forms a fraction $1-\varepsilon$ of the channel capacity. Applying Lemma 4 to the RHS of (60), we get

$$\begin{aligned}&\frac{H(\mathbf{X}|\mathbf{Y})}{n}\\&\geq -\varepsilon\, C + \frac{1-(1-\varepsilon)C}{2\ln 2}\sum_{i=1}^{\infty}\left\{\Gamma_i \sum_{k=1}^{\infty}\frac{\left(g_1^i\right)^k}{k(2k-1)}\right\}\\&= -\varepsilon\, C + \big(1-(1-\varepsilon)C\big)\sum_{i=1}^{\infty}\left\{\left[1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)\right]\,\Gamma_i\right\}\end{aligned}$$

where the last equality follows from (42). Combining (45) with the last result gives

$$h_2(P_{\text{b}}) \geq -\varepsilon\, C + \big(1-(1-\varepsilon)C\big)\sum_{i=1}^{\infty}\left[1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)\right]\,\Gamma_i$$

and therefore

$$\sum_{i=1}^{\infty}\left\{\left[1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)\right]\,\Gamma_i\right\} \leq \frac{\varepsilon\, C + h_2(P_{\text{b}})}{1-(1-\varepsilon)C} \tag{61}$$

where P_{b} designates the average bit error probability of the ensemble under the considered decoding algorithm. Since all the terms in the sum on the LHS of (61) are non-negative, this sum is lower bounded by its i-th term, for any degree i. This provides the following upper bound on the fraction of parity-check nodes of any finite degree i:

$$\begin{aligned}\Gamma_i &\leq \frac{\varepsilon C + h_2(P_{\text{b}})}{1-(1-\varepsilon)C}\,\frac{1}{1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)}\\&\leq \big(\varepsilon C + h_2(P_{\text{b}})\big)\left[\frac{1}{1-C}\,\frac{1}{1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)}\right].\end{aligned} \tag{62}$$

This completes the proof of (29) for a general MBIOS channel. Let us now consider the particular case where the transmission is over a BEC with an erasure probability p. In this case, $g_k = 1-p$ for all $k \in \mathbb{N}$ (this equality follows directly from (19)), and the channel capacity is equal to $1-p$ bits per channel use. Therefore, (60) is particularized to

$$\begin{aligned}&\frac{H(\mathbf{X}|\mathbf{Y})}{n}\\&\geq -\varepsilon(1-p) + \frac{1-(1-\varepsilon)(1-p)}{2\ln 2}\\&\qquad\cdot\sum_{i=1}^{\infty}\left[\Gamma_i\,(1-p)^i \sum_{k=1}^{\infty}\frac{1}{k(2k-1)}\right]\\&= -\varepsilon(1-p) + \big(1-(1-\varepsilon)(1-p)\big)\sum_{i=1}^{\infty}\Gamma_i\,(1-p)^i\end{aligned} \tag{63}$$

where the above equality holds since $\sum_{k=1}^{\infty} \frac{1}{k(2k-1)} = 2\ln 2$. Applying the upper bound on the conditional entropy (50) to the LHS of (63), we get

$$P_{\text{b}} \geq -\varepsilon(1-p) + \big(p + \varepsilon\,(1-p)\big) \sum_{i=1}^{\infty} \Gamma_i\,(1-p)^i$$

where P_{b} denotes the average bit erasure probability of the ensemble, and therefore

$$\sum_{i=1}^{\infty} \Big\{\Gamma_i (1-p)^i\Big\} \leq \frac{\varepsilon\,(1-p) + P_{\text{b}}}{p + \varepsilon\,(1-p)}. \tag{64}$$

Since the sum on the LHS of (64) is of non-negative terms, then we get

$$\Gamma_i \leq \big(\varepsilon\,(1-p) + P_{\text{b}}\big) \left(\frac{1}{p\,(1-p)^i}\right) \tag{65}$$

so $h_2(P_{\text{b}})$ in (29) is replaced for the BEC with P_{b}.

We turn now to consider the pair of degree distributions from the edge perspective. The average left degree (a_{L}) of the LDPC code ensemble satisfies

$$\frac{1}{a_{\text{L}}} = \sum_{i=2}^{\infty} \frac{\lambda_i}{i} \tag{66}$$

which implies that for any degree i of the variable nodes

$$\lambda_i \leq \frac{i}{a_{\text{L}}}\,. \tag{67}$$

Since the design rate of the LDPC code ensemble is assumed to be a fraction $1-\varepsilon$ of the channel capacity, then the average right and left degrees satisfy

$$\begin{aligned} a_{\text{L}} &= \big(1-(1-\varepsilon)C\big)a_{\text{R}} \\ &\geq (1-C)a_{\text{R}}. \end{aligned} \tag{68}$$

Substituting (68) on the RHS of (67) and applying the lower bound on a_{R} in (22) gives

$$\lambda_i \leq \frac{i\ln\!\left(\frac{1}{g_1}\right)}{2(1-C)\ \ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_{\text{b}})}{1-(1-\varepsilon)C}\right)}\right)}\,. \tag{69}$$

Using the power series for the binary entropy function in (16) and truncating the sum on the RHS after the first term gives

$$1 - h_2(x) \geq \frac{(1-2x)^2}{2\,\ln 2}$$

and substituting $u = h_2(x)$ yields

$$\big(1 - 2h_2^{-1}(u)\big)^2 \leq 2\,\ln 2 \cdot (1-u), \quad \forall\, 0 \leq u \leq 1. \tag{70}$$

Combining (69) and (70) gives

$$\begin{aligned} \lambda_i &\leq \frac{i\ln\!\left(\frac{1}{g_1}\right)}{(1-C)\ \ln\left(\frac{1}{2\ln 2}\ \frac{1}{1-\frac{1-C-h_2(P_{\text{b}})}{1-(1-\varepsilon)C}}\right)} \\ &= \frac{i\ln\!\left(\frac{1}{g_1}\right)}{(1-C)\ \ln\left(\frac{1}{2\ln 2}\ \frac{1-(1-\varepsilon)C}{\varepsilon C + h_2(P_{\text{b}})}\right)} \\ &\leq \frac{i\ln\!\left(\frac{1}{g_1}\right)}{(1-C)\ \left[\ln\left(\frac{1}{\varepsilon C + h_2(P_{\text{b}})}\right) + \ln\left(\frac{1-C}{2\ln 2}\right)\right]} \end{aligned}$$

which completes the proof of (30) for general MBIOS channels. For the BEC, we substitute (68) and the lower bound on the average right degree in (24) into the RHS of (67) to get

$$\begin{aligned}\lambda_i &\leq \frac{i \ln\left(\frac{1}{1-p}\right)}{p\ \ln\left(1+\frac{p-P_{\rm b}}{\varepsilon\,(1-p)+P_{\rm b}}\right)} \\ &= \frac{i \ln\left(\frac{1}{1-p}\right)}{p\ \ln\left(\frac{\varepsilon\,(1-p)+p}{\varepsilon\,(1-p)+P_{\rm b}}\right)} \\ &\leq \frac{i \ln\left(\frac{1}{1-p}\right)}{p\ \left[\ln\left(\frac{1}{\varepsilon\,(1-p)+P_{\rm b}}\right)+\ln(p)\right]}\,. \end{aligned} \tag{71}$$

Hence, $h_2(P_{\rm b})$ in (30) is replaced by $P_{\rm b}$ when the communication channel is a BEC. Considering the right degree distribution of the ensemble, we have

$$\frac{1}{a_{\rm R}} = \sum_{i=1}^{\infty} \frac{\rho_i}{i}\,.$$

By following the same steps as in (66)–(71), one obtains an upper bound on ρ_i for any degree i of the parity-check nodes. The asymptotic behavior of the resulting upper bound on ρ_i is similar to the upper bound on λ_i as given in (71). However, as we show in the following, a tighter upper bound on the fraction of edges connected to parity-check nodes of degree i is derived from the equality

$$\rho_i = \frac{i\,\Gamma_i}{a_{\rm R}}\,. \tag{72}$$

Substituting (22) and (62) in the above equality, we get

$$\rho_i \leq \frac{\varepsilon\,C+h_2(P_{\rm b})}{1-C}\ \frac{\ln\left(\frac{1}{g_1}\right)}{2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_{\rm b})}{1-(1-\varepsilon)C}\right)}\right)}\cdot\frac{i}{1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)}\,. \tag{73}$$

Applying (70) to the denominator of the second term on the RHS of (73) gives

$$\begin{aligned}\rho_i &\leq \frac{\varepsilon\,C+h_2(P_{\rm b})}{1-C}\ \frac{\ln\left(\frac{1}{g_1}\right)}{\ln\left(\frac{1}{2\ln 2}\ \frac{1}{1-\frac{1-C-h_2(P_{\rm b})}{1-(1-\varepsilon)C}}\right)}\ \frac{i}{1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)} \\ &= \frac{\varepsilon\,C+h_2(P_{\rm b})}{1-C}\ \frac{\ln\left(\frac{1}{g_1}\right)}{\ln\left(\frac{1}{2\ln 2}\ \frac{1-(1-\varepsilon)C}{\varepsilon C+h_2(P_{\rm b})}\right)}\ \frac{i}{1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)} \\ &\leq \frac{\ln\left(\frac{1}{g_1}\right)}{1-C}\ \frac{\varepsilon\,C+h_2(P_{\rm b})}{\ln\left(\frac{1}{\varepsilon\,C+h_2(P_{\rm b})}\right)+\ln\left(\frac{1-C}{2\ln 2}\right)}\ \frac{i}{1-h_2\left(\frac{1-g_1^{i/2}}{2}\right)}\,.\end{aligned}$$

This proves (31) regarding the fraction of edges connected to parity-check nodes of an arbitrary finite degree i. For a BEC, a substitution of (24) and (65) in (72) gives

$$\rho_i \leq \frac{i\big[\varepsilon\,(1-p)+P_{\rm b}\big]}{p(1-p)^i}\ \frac{\ln\left(\frac{1}{1-p}\right)}{\ln\left(1+\frac{p-P_{\rm b}}{\varepsilon\,(1-p)+P_{\rm b}}\right)}.$$

Followed by some straightforward algebra, this proves (31) for the BEC when $h_2(P_{\rm b})$ is replaced with $P_{\rm b}$.

Remark 12: **[Note on Theorem 2 and Corollary 2]** Consider the capacity-achieving sequence of right-regular LDPC code ensemble as introduced in [48]. The gap to capacity (ε) can be made arbitrarily small for this sequence (even under BP decoding), although $\rho_i = 1$ for some integer i. At first glance, it looks contradictory to Corollary 2 (see p. 12) which states that ρ_i is upper bounded by an expression which scales like $\frac{\varepsilon}{\ln \frac{1}{\varepsilon}}$ for any finite degree i, and it therefore should tend to zero as the gap to capacity vanishes. However, the right degree of this sequence scales like $\ln \frac{1}{\varepsilon}$ (see [48] and [40, Theorem 2.3]), hence the index i for which $\rho_i = 1$ becomes unbounded as $\varepsilon \to 0$. Note that Corollary 2 applies on the other hand to finite and bounded degrees i in the limit where the gap to capacity vanishes. Moreover, as we let $\varepsilon \to 0$ for this capacity-achieving and right-regular sequence, then ρ_i is identically zero for all finite and bounded degrees i.

Remark 13: **[On the degree distribution of the parity-check nodes for the set of MBIOS channels with a given capacity]** Consider the set of all MBIOS channels of a given capacity C, and consider a required bit error probability p_{b}. By combining the inequality constraint (61) with the extreme values of g_1 in Lemma 5 (see (53)), we obtain the following universal inequality constraint which should hold for this set of channels:

$$\sum_{i=1}^{\infty} \left\{ \left[1 - h_2 \left(\frac{1 - C^{\frac{i}{2}}}{2} \right) \right] \, \Gamma_i \right\} \leq \frac{\varepsilon \, C + h_2(P_{\text{b}})}{1 - (1-\varepsilon) C} \, . \tag{74}$$

We refer later to this inequality when we consider linear programming bounds for the degree distributions of capacity-approaching LDPC code ensembles (see Section V).

C. Proof of Theorem 3

Consider bipartite graphs which correspond to an LDPC code ensemble with pair of degree distributions (λ, ρ). The average degrees of the variable nodes and the parity-check nodes of these graphs are given in (4) and (5), respectively. Hence, the fraction of degree-2 variable nodes is given by

$$\Lambda_2 = \frac{\lambda_2 \, a_{\text{L}}}{2} = \frac{\lambda_2}{2 \int_0^1 \lambda(x) \mathrm{d}x} \tag{75}$$

and the design rate of this ensemble is given by (3). Using (3), we rewrite $\int_0^1 \lambda(x) \mathrm{d}x$ at the denominator of (75) as

$$\int_0^1 \lambda(x) \mathrm{d}x = \frac{1}{1 - R_{\text{d}}} \int_0^1 \rho(x) \mathrm{d}x \, . \tag{76}$$

By assumption, the considered sequence of ensembles achieves vanishing bit error probability under BP decoding, and hence the stability condition in (9) is satisfied. Combining (9), (75) and (76) leads to the following upper bound on Λ_2:

$$\Lambda_2 < \frac{1 - R_{\text{d}}}{2 \, \mathcal{B}(a) \; \rho'(1) \int_0^1 \rho(x) \mathrm{d}x} \, . \tag{77}$$

From the convexity of $f(t) = x^t$ for $x > 0$, Jensen's inequality gives

$$\begin{aligned}
&\int_0^1 \rho(x) \mathrm{d}x \\
&= \int_0^1 \sum_i \rho_i x^{i-1} \mathrm{d}x \\
&\geq \int_0^1 x^{\sum_i \rho_i \, (i-1)} \mathrm{d}x \\
&= \int_0^1 x^{\rho'(1)} \mathrm{d}x \\
&= \frac{1}{\rho'(1) + 1}
\end{aligned}$$

which implies that

$$\rho'(1) \geq \frac{1}{\int_0^1 \rho(x) \mathrm{d}x} - 1 = a_{\text{R}} - 1 \, . \tag{78}$$

Substituting (78) in (77) and since $R_{\text{d}} = (1-\varepsilon)C$ then

$$\begin{aligned} \Lambda_2 &< \frac{1-R_{\text{d}}}{2\,\mathcal{B}(a)}\left(1+\frac{1}{\rho'(1)}\right) \\ &\leq \frac{1-R_{\text{d}}}{2\,\mathcal{B}(a)}\left(1+\frac{1}{a_{\text{R}}-1}\right) \\ &= \frac{1-C}{2\,\mathcal{B}(a)}\left(1+\frac{\varepsilon C}{1-C}\right)\left(1+\frac{1}{a_{\text{R}}-1}\right). \end{aligned} \tag{79}$$

Since the RHS of (79) is monotonically decreasing with the average right degree (a_{R}), this bound still holds when a_{R} is replaced by a lower bound. For all $m \in \mathbb{N}$, let $P_{\text{b},m}$ designate the average bit error probability of the LDPC code ensemble $\bigl(n_m, \lambda(x), \rho(x)\bigr)$ under BP decoding. Applying Theorem 1 where $P_{\text{b},m}$ vanishes as $m \to \infty$ gives

$$a_{\text{R}} \geq \frac{2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)}\right)}{\ln\left(\frac{1}{g_1}\right)}. \tag{80}$$

The upper bound in (32) follows by substituting (80) in (79).

We now turn to derive the upper bound on the fraction of edges which are connected to degree-2 variable nodes. Since the considered sequence of LDPC code ensembles achieves vanishing bit error probability under BP decoding, then the stability condition (9) implies that

$$\lambda_2 = \lambda'(0) < \frac{1}{\rho'(1)\,\mathcal{B}(a)}$$

where $\mathcal{B}(a)$ is given in (8). Combining this with (78) gives

$$\lambda_2 < \frac{1}{(a_{\text{R}}-1)\,\mathcal{B}(a)} \tag{81}$$

where a_{R} designates the common average right degree of the sequence of ensembles. The upper bounds on λ_2 in (33) and (35) are obtained by substituting (80) and (52) (these are the lower bounds on a_{R} derived in Theorem 1 for vanishing bit error/ erasure probability), respectively, in (81).

Consider the set of all MBIOS channels with a given capacity C and a Bhattacharyya constant $\mathcal{B}(a)$, for which the bit error probability of the BP decoder vanishes for the considered sequence of LDPC code ensembles. Universal upper bound on Λ_2 and λ_2 follow directly by combining the bounds in (32) and (33), respectively, with Lemma 5 (note that the upper bound on the RHS of (32) is a monotonic decreasing function of g_1; this bound therefore attains its maximal value at the minimal value of g_1, i.e., when $g_1 = C$). Therefore, the universal upper bounds on Λ_2 and λ_2 hold for all the channels from the above set by substituting $g_1 = C$ on the RHS of (32) and (33), respectively.

For a transmission over the BEC, the improved upper bound on the degree-2 variable nodes follows by substituting the lower bound in (24) (where the bit erasure probability P_{b} vanishes) into (79). Note that for a BEC with erasure probability p, $1-C = \mathcal{B}(a) = p$ and $\frac{1-C}{2\mathcal{B}(a)} = \frac{1}{2}$. Similarly, the upper bound on the fraction of edges which are attached to degree-2 variable nodes follows by substituting (24) and $\mathcal{B}(a) = p$ into (81).

Discussion 4: **[On the tightness of the upper bound** (36) **on the fraction of degree-2 variable nodes for capacity-achieving LDPC code ensembles over MBIOS channels]** In the following, the tightness of the bound in (36) is considered:

Lemma 7: **[On the asymptotic fraction of degree 2 variable nodes for capacity-achieving sequences of LDPC code ensembles]** Let (n_m, λ_m, ρ_m) be a sequence of LDPC code ensembles whose transmission takes place over an MBIOS channel of capacity C (in bits per channel use). Assume that this sequence is capacity-achieving under BP decoding, and also that the flatness condition is asymptotically satisfied for this sequence (i.e., the stability condition in (9) is satisfied asymptotically with equality). Let us also assume that the limit of the ratio between the standard deviation and the expectation of the right degree distribution in the LDPC code ensemble (n_m, λ_m, ρ_m)

is finite as $m \to \infty$, and denote this limit by K. Then, the asymptotic fraction of degree-2 variable nodes in this sequence is equal to

$$\lim_{m\to\infty} \Lambda_2^{(m)} = \frac{1-C}{2(1+K^2)\,\mathcal{B}(a)} \tag{82}$$

where $\mathcal{B}(a)$ is introduced in (8).

Proof: See Appendix III. ■

As a particular case of Lemma 7, if $K = 0$ (this happens, e.g., when the right degree is fixed), then the asymptotic fraction of degree-2 variable nodes in (82) coincides with the upper bound in (36).

Remark 14: We note that the property proved in Lemma 7 for the non-vanishing asymptotic fraction of degree-2 *variable nodes* of capacity-achieving sequences of LDPC code ensembles is reminiscent of another information-theoretic property which was proved by Shokrollahi with respect to the non-vanishing fraction of degree-2 *output nodes* for capacity-achieving sequences of Raptor codes whose transmission takes place over an MBIOS channel (see [11, Theorem 11 and Proposition 12]).

Proof of Corollary 3: The upper bound (36) on the fraction of degree-2 variable nodes for capacity-achieving LDPC code ensembles follows directly by letting the gap to capacity ε tend to zero in (32). We wish to show that the upper bound in (36) is necessarily not larger than $\frac{1}{2}$ for all MBIOS channels, and it is equal to $\frac{1}{2}$ for a BEC regardless of the erasure probability of this channel. To this end, we prove the following lemma:

Lemma 8: For every MBIOS channel, the sum of its capacity and its Bhattacharyya constant is at least 1. The minimal value of this sum is attained for a BEC, irrespectively of the erasure probability of this channel, and is equal to 1.

Proof: See Appendix IV. ■

Combining Lemma 8 and the RHS of (36) implies that the fraction of degree-2 variable nodes for an arbitrary capacity-achieving sequence of LDPC code ensembles under BP decoding is upper bounded by $\frac{1}{2}$. Note that this maximal value is attained for a BEC (see also Remark 9 on page 14). This completes the proof of Corollary 3.

In the following, we compare two upper bounds on the fraction of edges connected to degree-2 variable nodes. One of these bounds is given in Theorem 3, and the other bound follows along the lines of the proof of Theorem 2.

Discussion 5: **[Comparison between two upper bounds on λ_2: ML versus iterative decoding]** In the proof of Theorem 2, we derive an upper bound on the fraction of edges connected to variable nodes of degree i for ensembles of LDPC codes which achieve a bit error (or erasure) probability P_{b} under an arbitrary decoding algorithm (see (69) and the tightened version (71) of this bound for the BEC). Referring to degree-2 variable nodes and letting P_{b} vanish, (69) gives

$$\lambda_2 \leq \frac{\ln\left(\frac{1}{g_1}\right)}{(1-C)\ \ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)}\right)} \tag{83}$$

where $R = (1-\varepsilon)C$. It is interesting to see that there is some similarity between the two upper bounds on λ_2 as given in (33) and (83). In the following, we compare between the two bounds on λ_2 by calculating the ratio between the bound in (33) which relies on the stability condition, and the bound in (83) which follows along the lines of the proof of Theorem 2. This gives

$$\begin{aligned}
&\frac{\ln\left(\frac{1}{g_1}\right)}{\mathcal{B}(a)\,\ln\left(\frac{g_1}{\left[1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)\right]^2}\right)} \\
&\cdot\frac{\left(1-C\right)\,\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)}\right)}{\ln\left(\frac{1}{g_1}\right)} \\
&= \frac{1-C}{\mathcal{B}(a)}\;\frac{\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)}\right)}{\ln\left(\frac{g_1}{\left[1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)\right]^2}\right)} \\
&= \frac{1-C}{\mathcal{B}(a)}\;\frac{\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)}\right)}{\ln(g_1)+2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C}{1-R}\right)}\right)}\,. \qquad (84)
\end{aligned}$$

Hence, as the gap to capacity vanishes (i.e., $\varepsilon \to 0$), the expression in (84) for the ratio between the two bounds on λ_2 tends to $\frac{1-C}{2\mathcal{B}(a)}$. By Lemma 8, $\mathcal{B}(a)+C-1 \geq 0$, which implies that $\frac{1-C}{2\mathcal{B}(a)} \leq \frac{1}{2}$. Hence, the upper bound on λ_2 in (33) improves the bound in (83) by at least a factor of 2 (where the former bound is given in Theorem 3, and the latter bound follows along the lines of the proof of Theorem 2). We note that the basis of the comparison between these two upper bounds on λ_2 is the assumption of vanishing bit error probability under BP decoding, though the bound in (83) also holds with the weaker requirement of vanishing bit error probability under ML decoding.

Proof of Corollary 4: See Appendix V.

Proof of Proposition 1: See Appendix VI.

V. Implications of the Information-Theoretic Bounds and Numerical Results

We provide here some implications of the information-theoretic bounds and numerical results which refer to the following issues:

- Examination of the tightness of the bounds provided in Section III by comparing these bounds to the asymptotic performance of some LDPC code ensembles under BP decoding (referring here to the sum-product decoding). In order to make this comparison more conclusive, we compare the new bounds with previously reported bounds (see Section V-A) in order to exemplify their practicality.
- Information-theoretic lower bound on the cardinality of the fundamental system of cycles of LDPC code ensembles, expressed in terms of the achievable gap to capacity (see Section V-B).
- Linear programming (LP) bounds on the degree distributions of capacity-approaching LDPC code ensembles. The bounds refer to the case where the communication takes place over an MBIOS channel, as well as universal bounds which are valid for the set of all MBIOS channels which exhibit a given capacity C. These bounds are valid under ML decoding (and hence, they are also valid under any sub-optimal decoding algorithm). These LP bounds are solved analytically, and are also compared with the degree distributions of capacity-approaching LDPC code ensembles under BP decoding (see Section V-C).
- Lower bounds on the graphical complexity of binary linear block codes which are represented by an arbitrary bipartite graph and whose transmission takes place over an MBIOS channel. The graphical complexity is measured by the total number of edges in the graph, and the bound provides a quantitative measure of the minimal number of edges required for this graphical representation as a function of the target block error probability and the gap (in rate) to capacity. This bound refers to codes of finite-length, and is valid under ML decoding (or any sub-optimal decoding). It can be also applied to LDPC code ensembles, and then it provides a lower bound on the decoding complexity per iteration of a BP decoder. Comparison of the information-theoretic lower bound on the graphical complexity in terms of the achievable gap to capacity with a target block error

probability with some efficient finite-length LDPC codes which are provided in the literature enables to evaluate the maximal potential gain that can be attained by future design of such finite-length codes in terms of the tradeoff between performance and graphical complexity (see Section V-D).

A. Numerical Results for the Asymptotic Analysis under BP Decoding

The following sub-section relies on the theoretic results provided in Section III, and it exemplifies the use of these results in the context of capacity-approaching sequences of LDPC code ensembles whose transmission takes place over an MBIOS channel, and whose bit error probability vanishes under BP decoding. As representatives of MBIOS channels, the considered communication channels are the binary erasure channel (BEC), binary symmetric channel (BSC) and the binary-input AWGN channel (BIAWGNC) (as presented in [37, Example 4.1]).

Example 2: **[BEC]** Consider a sequence of LDPC code ensembles (n, λ, ρ) where the block length (n) tends to infinity and the pair of degree distributions is given by

$$\lambda(x) = 0.409x + 0.202x^2 + 0.0768x^3 + 0.1971x^6 + 0.1151x^7$$
$$\rho(x) = x^5.$$

The design rate of this ensemble is $R = 0.5004$, and the threshold under BP decoding is (see [37, Theorem 3.59])

$$p^{\text{BP}} = \inf_{x \in (0,1]} \frac{x}{\lambda\big(1 - \rho(1-x)\big)} = 0.4810$$

so the minimum capacity of a BEC over which it is possible to transmit with vanishing P_{b} under BP decoding is $C = 1 - p^{\text{BP}} = 0.5190$ bits per channel use, and the multiplicative gap to capacity is $\varepsilon = 1 - \frac{R}{C} = 0.0358$. The lower bound on the average right degree in (24) with vanishing bit erasure probability (i.e., $P_{\text{b}} = 0$) gives that the average right degree should be at least 5.0189. By imposing a prior assumption that the LDPC code ensemble has a fixed right degree (as is the case with the above LDPC code ensemble), then it follows that this right degree cannot be below 6. Hence, the lower bound is attained in this case with equality. An upper bound on the fraction of edges which are connected to degree-2 variable nodes (λ_2) is calculated from (81) with $\mathcal{B}(a) = p^{\text{BP}} = 0.4810$, and the above lower bound on a_{R} (for LDPC code ensembles of a fixed right degree) which is equal to 6; this gives from (81) that $\lambda_2 \leq 0.4158$ as compared to the exact value which is equal to 0.409. The exact value of the fraction of degree-2 variable nodes is

$$\Lambda_2 = \frac{\lambda_2 \, a_{\text{L}}}{2} = \frac{\lambda_2 \, (1-R) \, a_{\text{R}}}{2} = 0.6130$$

as compared to the upper bound in (79), combined with the tight lower bound $a_{\text{R}} \geq 6$, which gives $\Lambda_2 \leq 0.6232$. We note that without the prior assumption about the fixed right degree, the universal bounds give $a_{\text{R}} \geq 5.0189$ and $\lambda_2 < 0.5173$ so these bounds are clearly loosened.

Example 3: **[Comparison of the lower bound on the average right degree from Theorem 1 and Discussion 2 with the bound in [53]]** In the following, we exemplify the practical use of the lower bound on the average right degree of LDPC code ensembles, as given in Theorem 1 and its adaptation to LDPC code ensembles in Discussion 2, and compare it with the previously reported bound in [53, Section IV]. Consider the case where the communications takes place over a BIAWGNC. The LDPC code ensembles in each sequence are specified by the following pairs of degree distributions, followed by their corresponding design rates and thresholds under BP decoding:

Ensemble 1:

$$\lambda(x) = x, \quad \rho(x) = x^{19}, \quad R_{\text{d}} = 0.9000.$$
$$\sigma_{\text{BP}} = 0.4156590.$$

Ensemble 2:

$$\lambda(x) = 0.4012x + 0.5981x^2 + 0.0007x^{29}, \quad \rho(x) = x^{24}$$
$$R_{\text{d}} = 0.9000, \quad \sigma_{\text{BP}} = 0.4741840.$$

These code ensembles are taken from the data base in [2]. From [37, Example 4.38] which expresses the capacity of the BIAWGNC in terms of the standard deviation σ of the Gaussian noise, the minimum capacity of a BIAWGNC

TABLE I

BOUNDS VS. EXACT VALUES OF λ_2 AND a_{R} FOR TWO SEQUENCES OF LDPC CODE ENSEMBLES OF DESIGN RATE $\frac{1}{2}$ TRANSMITTED OVER THE BIAWGNC. THE SEQUENCES ARE GIVEN IN [8, TABLE II] AND ACHIEVE VANISHING BIT ERROR PROBABILITY UNDER THE BELIEF PROPAGATION (BP) DECODING ALGORITHM WITH THE INDICATED GAPS TO CAPACITY.

LDPC ensemble	Gap to capacity (ε)	a_{R}	Lower bound on a_{R} (Theorem 1)	λ_2	Upper bound on λ_2 (Theorem 3)
1	$3.72 \cdot 10^{-3}$	10.938	9.249	0.170	0.205
2	$2.22 \cdot 10^{-3}$	12.000	10.129	0.153	0.185

over which it is possible to communicate with vanishing bit error probability under BP decoding is $C = 0.9685$ and 0.9323 bits per channel use for Ensembles 1 and 2, respectively. The corresponding gap (in rate) to capacity $\varepsilon = 1 - \frac{R_{\mathrm{d}}}{C}$ is equal to $\varepsilon = 7.07 \cdot 10^{-2}$ and $3.46 \cdot 10^{-2}$, respectively. Therefore, for the first ensemble which is a (2,20) regular LDPC code ensemble, the new lower bound on the average right degree which follows from Discussion 2 is equal to 9.949 whereas the lower bound from [53, Section IV] (i.e., the un-numbered equation before [53, Eq. (77)]) is equal to 2.392. For the second ensemble whose fixed right degree is equal to 25, the new lower bound on the average right degree is 16.269 whereas the lower bound from [53] is 14.788. This shows that the improvement obtained in Theorem 1 followed by Discussion 2 is of practical use.

We note that the gap which still exists between the lower bounds on the average right degrees and the actual values of a_{R} for the above two ensembles is partially attributed to the fact that this information-theoretic lower bound holds even under ML decoding, although we apply this bound here under the sub-optimal BP decoding algorithm. The gaps to capacity under ML decoding are smaller than those calculated under BP decoding, and smaller values of ε provide improved lower bounds on a_{R}.

Example 4: **[BIAWGNC]** Table I considers two sequences of LDPC code ensembles of design rate $\frac{1}{2}$ which are taken from [8, Table II]. The transmission of these ensembles is assumed to take place over the BIAWGNC. The pair of degree distributions of the ensembles in each sequence is fixed and the block length of these ensembles tends to infinity. The LDPC code ensembles in each sequence are specified by the following pairs of degree distributions:
Ensemble 1:

$$\begin{aligned}\lambda(x) &= 0.170031x + 0.160460x^2 + 0.112837x^5 \\ &\quad +0.047489x^6 + 0.011481x^9 + 0.091537x^{10} \\ &\quad +0.152978x^{25} + 0.036131x^{26} + 0.217056x^{99} \\ \rho(x) &= \frac{1}{16}\,x^9 + \frac{15}{16}\,x^{10}.\end{aligned}$$

Ensemble 2:

$$\begin{aligned}\lambda(x) &= 0.153425x + 0.147526x^2 + 0.041539x^5 \\ &\quad +0.147551x^6 + 0.047938x^{17} + 0.119555x^{18} \\ &\quad +0.036379x^{54} + 0.126714x^{55} + 0.179373x^{199} \\ \rho(x) &= x^{11}.\end{aligned}$$

The asymptotic thresholds of the considered LDPC code ensembles under BP decoding are calculated with the DE technique, and these calculations provide the thresholds $\sigma_{\mathrm{BP}} = 0.97592$ and 0.97704, respectively. The minimum capacity of a BIAWGNC which enables to communicate Ensembles 1 and 2 with vanishing bit error probability under BP decoding is therefore $C = 0.5019$ and 0.5011 bits per channel use, respectively (it is calculated via the power series expansion of the capacity of a BIAWGNC as given in [37, page 194]). This leads to the indicated gaps (in rate) to capacity as given in Table I. The value of λ_2 for each sequence of LDPC code ensembles (where we let the block length tend to infinity) is compared with the upper bound in Theorem 3 which corresponds to BP decoding. Note that for calculating the bound in Theorem 3, the Bhattacharyya constant in (8) is given by $\mathcal{B}(a) = \exp\bigl(-\frac{RE_{\mathrm{b}}}{N_0}\bigr)$ for the BIAWGNC where $\frac{E_{\mathrm{b}}}{N_0}$ designates the energy per information bit over the one-sided noise spectral density, and we substitute here the threshold value of $\frac{E_{\mathrm{b}}}{N_0}$ under BP decoding. The average right

TABLE II

COMPARISON OF THEORETICAL BOUNDS AND ACTUAL VALUES OF λ_2 AND a_{R} FOR TWO SEQUENCES OF LDPC CODE ENSEMBLES TRANSMITTED OVER THE BSC. THE SEQUENCES ARE TAKEN FROM [2] AND ACHIEVE VANISHING BIT ERROR PROBABILITY UNDER THE BELIEF PROPAGATION (BP) DECODING ALGORITHM WITH THE INDICATED GAPS TO CAPACITY.

LDPC ensemble	Gap to capacity (ε)	a_{R}	Lower bound on a_{R} (Theorem 1)	λ_2	Upper bound on λ_2 (Theorem 3)
1	$1.85 \cdot 10^{-2}$	5.172	4.301	0.291	0.371
2	$6.18 \cdot 10^{-3}$	11.000	9.670	0.160	0.185

degree of each sequence is also compared with the lower bound in Theorem 1. These comparisons exemplify that for the examined LDPC code ensembles, both of the theoretical bounds are informative.

Example 5: **[BSC]** Table II considers two sequences of LDPC code ensembles, taken from [2], where the pair of degree distributions of the ensembles in each sequence is fixed and the block length of these ensembles tends to infinity. The transmission of these ensembles is assumed to take place over the BSC. The LDPC code ensembles in each sequence are specified by the following pairs of degree distributions and design rates:

Ensemble 1:

$$\begin{aligned}\lambda(x) &= 0.291157x + 0.189174x^2 + 0.0408389x^4 \\ &\quad +0.0873393x^5 + 0.00742718x^6 + 0.112581x^7 \\ &\quad +0.0925954x^{15} + 0.0186572x^{20} + 0.124064x^{32} \\ &\quad +0.016002x^{39} + 0.0201644x^{44} \\ \rho(x) &= 0.8x^4 + 0.2x^5 \\ R &= 0.250\end{aligned}$$

Ensemble 2:

$$\begin{aligned}\lambda(x) &= 0.160424x + 0.160541x^2 + 0.0610339x^5 \\ &\quad +0.153434x^6 + 0.0369041x^{12} + 0.020068x^{15} \\ &\quad +0.0054856x^{16} + 0.128127x^{19} + 0.0233812x^{24} \\ &\quad +0.05285542x^{34} + 0.0574104x^{67} + 0.0898442x^{68} \\ &\quad +0.0504923x^{85} \\ \rho(x) &= x^{10} \\ R &= 0.500.\end{aligned}$$

The thresholds of the above LDPC code ensembles under BP decoding are equal to $p_{\mathrm{BSC}} = 0.2120$ and 0.1090, respectively. Hence, for Ensembles 1 and 2, the minimum capacity of a BSC which enables to communicate with vanishing bit error probability under BP decoding is $C = 0.2547$ and 0.5031 bits per channel use. Since of the design rates of these two ensembles are $R_{\mathrm{d}} = 0.250$ and 0.500, respectively, then the gaps to capacity are given in Table II. The value of λ_2 for each sequence is compared with the upper bound given in Theorem 3. Note that for calculating the bound in Theorem 3, the Bhattacharyya constant $\mathcal{B}(a)$ introduced in (8) satisfies $\mathcal{B}(a) = \sqrt{4p(1-p)}$ for a BSC whose crossover probability is equal to p, and we substitute here the threshold value of p under BP decoding. Also, for the calculation of this bound for such a BSC, Eq. (97) gives that $g_1 = (1-2p)^2$. The average right degree of each sequence is also compared with the lower bound in Theorem 1. These comparisons show that for the considered sequences of LDPC code ensembles, both of the theoretical bounds are fairly tight; the upper bound on λ_2 is within a factor of 1.3 from the actual value for the two sequences of LDPC code ensembles while the lower bound on the average right degree is not lower than 83% of the corresponding actual values. The LDPC code ensembles referred to in Table II were obtained in [2] by the DE technique with the goal of minimizing the gap to capacity under a constraint on the maximal degree.

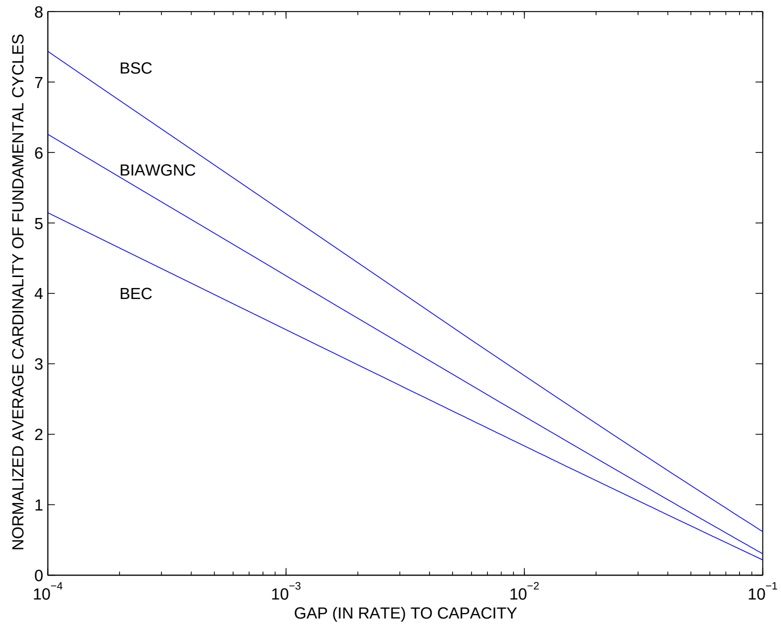

Fig. 3. Plot of the asymptotic lower bounds in Corollary 1 (see Eqs. (26) and (27)) for memoryless binary-input output-symmetric (MBIOS) channels. These lower bounds correspond to the average cardinality of the fundamental system of cycles for bipartite graphs representing codes from an arbitrary LDPC code ensemble; the above quantity is normalized with respect to the block length of the ensemble, and the asymptotic result refers to the case where we consider a sequence of LDPC code ensembles whose block lengths tend to infinity. The bounds are plotted versus the achievable gap (in rate) between the channel capacity and the design rate of the LDPC code ensembles. This figure shows the bounds for the binary symmetric channel (BSC), binary-input AWGN channel (BIAWGNC) and the binary erasure channel (BEC) where it is assumed that the design rate of the LDPC code ensembles is equal to one-half bit per channel use.

B. On the Fundamental System of Cycles for Capacity-Approaching Sequences of LDPC Code Ensembles

Corollary 1 considers an arbitrary sequence of LDPC code ensembles, specified by a pair of degree distributions, whose transmission takes place over an MBIOS channel. This corollary refers to the asymptotic case where we let the block length of the ensembles in this sequence tend to infinity and the bit error (or erasure) probability vanishes; the design rate of these ensembles is assumed to be a fraction $1-\varepsilon$ of the channel capacity (for an arbitrary $\varepsilon \in (0,1)$). In Corollary 1, Eq. (26) applies to a general MBIOS channel and a tightened version of this bound is given in (27) for the BEC. Based on these results, the asymptotic average cardinality of the fundamental system of cycles for bipartite graphs representing codes from LDPC code ensembles as above, where this average cardinality is normalized with respect to the block length, grows at least like $\ln \frac{1}{\varepsilon}$. We consider here the BSC, BEC, and BIAWGNC as three representatives of the class of MBIOS channels, and assume that the design rate of the LDPC code ensembles is fixed to one-half bit per channel use. It is shown in Fig. 3 that for a given gap (ε) to the channel capacity and for a fixed design rate, the extreme values of this lower bounds correspond to the BSC and BEC (which attain the maximal and minimal values, respectively). This observation is consistent with the last part of the statement in Corollary 1.

C. Linear Programming Bounds for the Degree Distributions of LDPC Code Ensembles

This sub-section provides LP bounds on the degree distributions of LDPC code ensembles. These bounds, which are based on Sections III and IV, are formulated in terms of the target bit error probability and the gap (in rate) to capacity required to achieve this target. The following LP bounds refer to the node and the edge perspectives of the pair of degree distributions, and they provide upper bounds on the fraction of edges or nodes up to degree k where k is a parameter. Similarly to Theorem 2, the LP bounds which are introduced in this section hold under ML decoding, and are therefore general in terms of the decoding algorithm. These LP bounds apply to finite-length LDPC code ensembles as well as to the asymptotic case of an infinite block length. Analytical solutions for these LP bounds are provided in Section V-C, and these bounds are also compared with some capacity-achieving sequences of LDPC code ensembles for the BEC under BP decoding. The following LP bounds are separated into four categories:

- **LP1**: 'LP1' forms an LP upper bound on the degree distribution of the parity-check nodes for LDPC code ensembles whose transmission takes place over an MBIOS channel. Its first version gives an upper bound on the fraction of parity-check nodes up to degree k (where $k \geq 1$ is an integer) as a function of the achievable rate (and its gap to the channel capacity) with a given bit error probability P_{b}. By combining (61) with the trivial constraints for an arbitrary degree distribution, the following optimization problem follows:

$$\begin{array}{l} \text{maximize} \quad \sum_{i=1}^{k} \Gamma_i, \quad k = 1, 2, \ldots \\ \text{subject to} \\ \begin{cases} \sum_{i=1}^{\infty} \left\{ \left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2} \right) \right] \Gamma_i \right\} \leq \frac{\varepsilon\, C + h_2(P_{\text{b}})}{1-(1-\varepsilon)C} \\ \sum_{i=1}^{\infty} \Gamma_i = 1 \\ \Gamma_i \geq 0, \quad i = 1, 2, \ldots \end{cases} \end{array}$$

  where the optimization variables are $\{\Gamma_i\}_{i \geq 1}$. From (1), the following equality holds:

$$\Gamma_i = \frac{\rho_i}{i} \left(\sum_{j=1}^{\infty} \frac{\rho_j}{j} \right)^{-1}. \tag{85}$$

  The substitution of this equality in the first constraint of the above LP bound gives the following optimization problem for the degree distribution of the parity-check nodes from the edge perspective (i.e., we get an upper bound on the fraction of edges which are connected to parity-check nodes up to degree $k \geq 1$):

$$\begin{array}{l} \text{maximize} \quad \sum_{i=1}^{k} \rho_i, \quad k = 1, 2, \ldots \\ \text{subject to} \\ \begin{cases} \sum_{i=1}^{\infty} \left\{ \left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2} \right) \right] \frac{\rho_i}{i} \right\} \leq \frac{\varepsilon\, C + h_2(P_{\text{b}})}{1-(1-\varepsilon)C} \sum_{i=1}^{\infty} \frac{\rho_i}{i} \\ \sum_{i=1}^{\infty} \rho_i = 1 \\ \rho_i \geq 0, \quad i = 1, 2, \ldots \end{cases} \end{array}$$

  where the optimization variables are $\{\rho_i\}_{i \geq 1}$. These two LP bounds on the parity-check degree distribution (from the node and edge perspectives) rely both on Theorems 1 and 2, and are therefore valid under ML decoding (hence, they also hold under any other decoding algorithm). These bounds hold for finite-length codes and also for the asymptotic case of an infinite block length.
  An analytical solution of the LP1 bound is given in Appendix VIII. This bound is tightened in Appendix VIII for the BEC, followed by its analytical solution.
- **LP2**: 'LP2' provides a universal LP upper bound on the degree distribution of the parity-check nodes for LDPC code ensembles as a function of the required achievable rate (and its gap to the channel capacity) with

a required bit error probability P_{b}. This bound follows from (74) and (85), and it gets the form:

$$\begin{array}{l}
\text{maximize} \quad \sum_{i=1}^{k} \rho_i, \quad k = 1, 2, \ldots \\
\text{subject to} \\
\begin{cases}
\sum_{i=1}^{\infty} \left\{ \left[1 - h_2\left(\frac{1-C^{\frac{i}{2}}}{2}\right) \right] \frac{\rho_i}{i} \right\} \leq \frac{\varepsilon\, C + h_2(P_{\text{b}})}{1-(1-\varepsilon)C} \sum_{i=1}^{\infty} \frac{\rho_i}{i} \\
\sum_{i=1}^{\infty} \rho_i = 1 \\
\rho_i \geq 0, \quad i = 1, 2, \ldots
\end{cases}
\end{array}$$

where the optimization variables are $\{\rho_i\}_{i \geq 1}$, and the bound holds under the same conditions as of the previous item. However, as opposed to the LP1 bound, the LP2 bound is universal since it holds for all MBIOS channels which exhibit a given capacity C. Note that the LP2 bound is similar to the LP1 bound, except of replacing the parameter g_1 in the LP1 bound with the channel capacity C. This follows directly by comparing (61) and (74). Note that the transition from (61) to (74) follows from Lemma 5 which implies that among all MBIOS channels with a given capacity C, the channel which attains the minimal value of g_1 is the BEC, and the minimal value of g_1 is equal to C.

The analytical solution of the LP2 bound follows directly from the analysis in Appendix VIII for the LP1 bound, by replacing g_1 in the LP1 bound with the channel capacity C in the LP2 bound.

- **LP3**: 'LP3' provides an LP upper bound on the degree distribution of the variable nodes (from the edge perspective) for LDPC code ensembles whose transmission takes place over an MBIOS channel. This bound provides an upper bound on the fraction of edges which are connected to variable nodes up to degree k for a parameter $k \geq 2$, and it is expressed in terms of the required achievable rate (and its gap to capacity) with a given bit error probability P_{b}. From (3) and (22), this LP bound gets the form

$$\begin{array}{l}
\text{maximize} \quad \sum_{i=2}^{k} \lambda_i, \quad k = 2, 3, \ldots \\
\text{subject to} \\
\begin{cases}
\sum_{i=2}^{\infty} \frac{\lambda_i}{i} \leq \dfrac{\ln\left(\frac{1}{g_1}\right)}{2(1-C)\left(1+\frac{\varepsilon C}{1-C}\right) \ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_{\text{b}})}{1-(1-\varepsilon)C}\right)} \right)} \\
\sum_{i=2}^{\infty} \lambda_i = 1 \\
\lambda_i \geq 0, \quad i = 2, 3, \ldots
\end{cases}
\end{array}$$

where the optimization variables are $\{\lambda_i\}_{i \geq 2}$. Since the bound relies on Theorem 1, then it is therefore valid under ML decoding (or any other decoding algorithm). It holds for finite block-length as well as in the asymptotic case where we let the block length tend to infinity. We note that the focus on the degree distribution of the variable nodes from the edge perspective is due to Theorem 2 and Remark 9 (see p. 14).

- **LP4**: 'LP4' provides a universal LP upper bound on the degree distribution of the variable nodes for LDPC

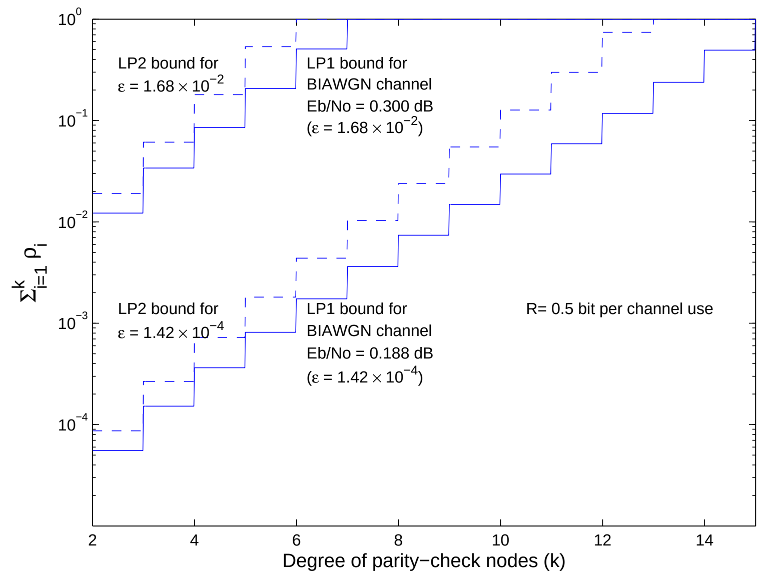

Fig. 4. LP1 versus LP2 upper bounds on the degree distributions of the parity-check nodes, from the edge perspective, for LDPC code ensembles whose design rate is $R = \frac{1}{2}$. The stair functions show upper bounds on the fraction of the edges which are connected to parity-check nodes whose degrees are at most k for an integer $k \geq 2$. The bounds are valid under ML decoding or any sub-optimal decoding algorithm. All these curves refer to a target bit error probability of $P_b = 10^{-10}$. The two LP1 bounds (solid lines) refer to binary-input AWGN (BIAWGN) channels for which $\frac{E_b}{N_0}$ = 0.300 and 0.188 dB, so the corresponding channel capacities are $C = 0.5086$ and 0.5001 bits per channel use, respectively; the corresponding gaps (in rate) to capacity are therefore equal to $\varepsilon = 1.68 \cdot 10^{-2}$ and $1.42 \cdot 10^{-4}$, respectively. The two universal LP2 bounds (dashed lines) correspond to all the MBIOS channels which exhibit a given capacity, whose value coincides in each case with the capacity of the considered BIAWGN channel.

code ensembles (from the edge perspective). It is based on (3) and (54) which give the following problem:

$$\begin{array}{|l|}\hline \text{maximize} \quad \sum_{i=2}^{k} \lambda_i, \quad k = 2, 3, \ldots \\ \text{subject to} \\ \begin{cases} \sum_{i=2}^{\infty} \frac{\lambda_i}{i} \leq \dfrac{\ln\left(\frac{1}{C}\right)}{2(1-C)\left(1+\frac{\varepsilon C}{1-C}\right) \ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_b)}{1-(1-\varepsilon)C}\right)}\right)} \\ \\ \sum_{i=2}^{\infty} \lambda_i = 1 \\ \\ \lambda_i \geq 0, \quad i = 2, 3, \ldots \end{cases} \\ \hline \end{array}$$

where the optimization variables are $\{\lambda_i\}_{i \geq 2}$. This bound holds for all MBIOS channels with a given capacity C.

The universal (LP2) bound is compared in Figure 4 to the LP1 bound for the BIAWGN channel with the same capacity. It is shown in this figure that the difference between these two bounds is not large. Note that the universal bound is attained for the BEC with the same capacity as of the BIAWGN channel.

Remark 15: **[A discussion on the constraints given in the LP1 and LP2 bounds and the un-necessity of adding the constraint in Theorem 1]** We prove in Appendix VII that adding the constraint which is imposed by the lower bound on the average right degree (i.e., the lower bound on $a_R = \sum_{i=1}^{\infty} i\Gamma_i$) does not affect the LP1 and LP2 bounds introduced here. This simplifies the formulation of the LP bounds serves for the derivation of closed-form analytical solutions of these bounds later in this section.

Remark 16: **[The LP1 and LP2 bounds and their connection with the asymptotic behavior as given in Theorem 2]** As shown via the upper bounds in Fig. 4, the fraction of edges which are connected to parity-check nodes of low degree is small, especially when the achievable gap to capacity vanishes. This is consistent with the theoretical result in Theorem 2 and Corollary 2 which states that the fraction of parity-check nodes of any finite degree scales at most like ε and the fraction of edges connected to parity-check nodes of any finite degree scales at most like $\frac{\varepsilon}{\ln \frac{1}{\varepsilon}}$ where ε designates the gap in rate to capacity, so both quantities tend to zero as the gap to capacity vanishes.

For solving the LP1 and LP2 bounds which are introduced in this section we originally used [15], a package for specifying and solving convex optimization problems [6]. It enables to solve these problems on a standard PC in a fraction of a second. However, it is still nice to get an analytic solution of these LP bounds.

Analytical solutions for the LP1 and LP2 bounds: The LP1 problem can be expressed in the following equivalent form:

$$\begin{aligned}
&\text{maximize} \quad \sum_{i=1}^{k} \rho_i, \quad k = 1, 2, \ldots \\
&\text{subject to} \\
&\begin{cases}
\sum_{i=1}^{\infty} d_i \rho_i \leq 0 \\
d_i \triangleq \frac{1}{i} \left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2} \right) - \frac{\varepsilon C + h_2(P_b)}{1-(1-\varepsilon)C} \right] \\
\sum_{i=1}^{\infty} \rho_i = 1 \\
\rho_i \geq 0, \quad i = 1, 2, \ldots
\end{cases}
\end{aligned}$$

An analytical solution for the LP1 bound is obtained in Appendix VIII (via the use of strong Lagrange duality).

In the following, the final solution of the LP1 bound is presented. To this end, note that for indices i large enough, $d_i < 0$ and also $\lim_{i\to\infty} d_i = 0$. Let $d^* \triangleq \min_{i\geq 1} d_i$ be the minimal value of this sequence, and let $i = l$ be the corresponding index of d_i which achieves this minimal value of the sequence $\{d_i\}$. Clearly, $d^* < 0$. The resulting closed-form solution for the LP1 bound gets the following form (see Appendix VIII):

- For values of k below the lower bound on the average right degree in (22), it is equal to $-\frac{d^*}{d_k - d^*}$.
- For values of k larger or equal to the lower bound on the average right degree in (22), it is equal to 1.

A similar solution is obtained for the LP2 bound where the only difference is that g_1 in the definition of the sequence $\{d_i\}$ is replaced by the channel capacity C. These analytical solutions match the numerical solutions obtained via [15].

Example 6: **[A comparison of the LP1 bound and capacity-achieving LDPC code ensembles over the BEC]** In the following, we compare the LP1 bound for the BEC and the degree distributions of two capacity-achieving sequences of LDPC code ensembles under iterative message-passing decoding.

The first capacity-achieving sequence for the BEC refers to the heavy-tail Poisson distribution, and it was introduced in [24, Section IV], [48] (see also [37, Problem 3.20]). The second capacity-achieving sequence refers to the right-regular LDPC code ensembles [48], based also on the analysis in the proof of Proposition 1 (see Section IV).

This first capacity-achieving sequence is obtained via the pair of degree distributions

$$\hat{\lambda}_\alpha(x) = -\frac{1}{\alpha} \cdot \ln(1-x) = \frac{1}{\alpha} \sum_{i=1}^{\infty} \frac{x^i}{i}$$

$$\rho_\alpha(x) = e^{\alpha(x-1)} = e^{-\alpha} \sum_{i=0}^{\infty} \frac{\alpha^i x^i}{i!}$$

which satisfies the equality $\hat{\lambda}_\alpha(1 - \rho_\alpha(1-x)) = x$ for all $\alpha > 0$. Starting with the heavy-tail Poisson distribution as above and proceeding along the lines in [37, Section 3.15], the following two steps are performed for the construction of capacity-approaching LDPC code ensembles for the BEC:

- The degree distribution $\hat{\lambda}_\alpha(x)$ is truncated so that it consists of the first N terms of its Taylor series expansion (up to and including the term x^{N-1}).
- The truncated power series $\hat{\lambda}_\alpha^{(N)}(x)$ is normalized so that it is equal to 1 at $x = 1$. The left degree distribution (from the edge perspective) is then equal to $\lambda_\alpha^{(N)}(x) = \frac{\hat{\lambda}_\alpha^{(N)}(x)}{\hat{\lambda}_\alpha^{(N)}(1)}$. The right degree distribution, $\rho_\alpha(x)$, is not modified.

This procedure provides the following degree distributions:

$$\begin{aligned} \lambda_i &= \frac{1}{H(N-1)\,(i-1)}, \quad i = 2, 3, \ldots N \\ \rho_i &= \frac{e^{-\alpha}\alpha^{i-1}}{(i-1)!}, \qquad i = 1, 2, \ldots \end{aligned} \tag{86}$$

where $H(k) \triangleq \sum_{i=1}^{k} \frac{1}{i}$ for $k \geq 1$ is a truncated harmonic sum. From (3), straightforward calculus shows that the design rate of the corresponding LDPC code ensemble is equal to

$$\begin{aligned} R_{\text{d}}(\alpha, N) &= 1 - \frac{\int_0^1 \rho_\alpha(x)\,\text{d}x}{\int_0^1 \lambda_\alpha^{(N)}(x)\,\text{d}x} \\ &= 1 - \frac{N\,H(N-1)\,(1-e^{-\alpha})}{(N-1)\alpha}. \end{aligned} \tag{87}$$

We need to determine the parameters α and N so that the design rate in (87) forms (at least) a fraction $1-\varepsilon$ of the capacity of the BEC. Let p designate the erasure probability of the channel, and let $r = (1-\varepsilon)(1-p)$ be the lower bound on the required design rate. We need to choose α and N to satisfy the inequality $R_{\text{d}}(\alpha, N) \geq r$ with vanishing bit erasure probability under BP decoding. Similarly to the calculations in [37, Example 3.88], the satisfiability of the inequality

$$\frac{\hat{\lambda}_\alpha^{(N)}(1)}{1-\hat{\lambda}_\alpha^{(N)}(1)} \left(\frac{\int_0^1 \rho_\alpha(x)\,\text{d}x}{\int_0^1 \hat{\lambda}_\alpha^{(N)}(x)\,\text{d}x} - 1 \right) \leq \varepsilon$$

implies this requirement, and straightforward algebra gives the inequality

$$\frac{\frac{H(N-1)}{\alpha}}{1-\frac{H(N-1)}{\alpha}} \left(\frac{N(1-e^{-\alpha})}{N-1} - 1 \right) \leq \varepsilon. \tag{88}$$

By choosing α to satisfy the equality $\frac{H(N-1)}{\alpha} = 1 - r$ and replacing $1 - e^{-\alpha}$ by 1, we get from (88) the following stronger requirement:

$$\frac{1-r}{r}\,\frac{1}{N-1} \leq \varepsilon \tag{89}$$

which then provides a proper choice for N. To conclude, the parameters α and N are chosen to be

$$\alpha = \frac{H(N-1)}{1-r}, \quad N = \left\lceil \frac{1-r}{\varepsilon r} \right\rceil + 1. \tag{90}$$

In the following, we calculate the heavy-tail Poisson distribution in (86) with the choice of parameters in (90). The resulting degree distribution of the parity-check nodes (from the edge perspective) is compared with the LP1 bound for the BEC where the analytical solution of this bound is given in Appendix VIII.

Comparisons between the heavy-tail Poisson distribution and the LP1 bound are shown in Figure 5. We note that the LP1 bound is an upper bound on the parity-check degree distribution which is valid under ML decoding (and hence, it is general for any decoding algorithm), whereas the heavy-tail Poisson distribution is designed to achieve a certain gap to capacity under BP decoding. We also show in this figure the fixed degree of the parity-check nodes for the right-regular LDPC code ensemble; this calculation is done via (112), (113), (118) where the right degree is equal to $a_{\text{R}} = \lceil \frac{1}{\alpha} \rceil + 1$. Although the latter case corresponds to a step function, the degree where this function switches from zero to one provides an indication to the reasonable tightness of the LP1 upper bound with respect to the value of the parity-check degree k where this upper bound is close to 1.

The following analysis compares between the behavior of the upper bound on ρ_i as given in Corollary 2 with the behavior of the heavy-tail Poisson distribution in the limit where the gap to capacity vanishes under BP decoding:

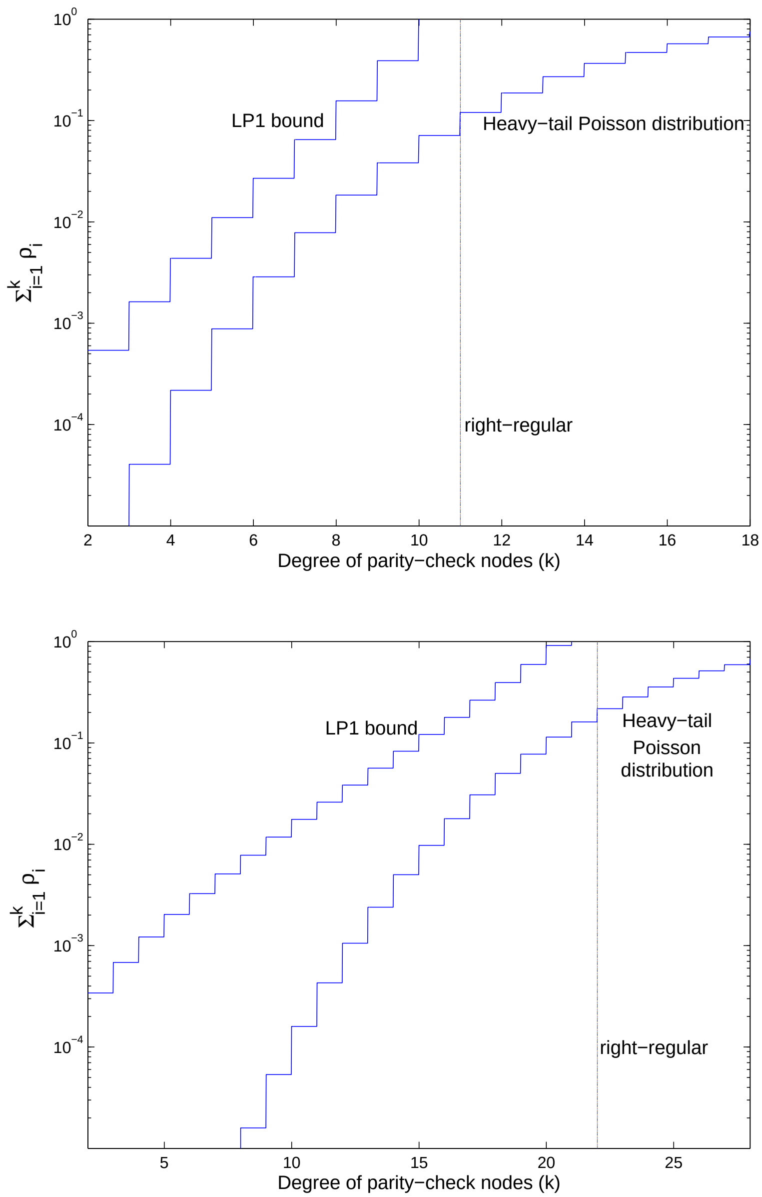

Fig. 5. A comparison between the LP1 bound, the heavy-tail Poisson degree distribution in (86) and (90), and the parity-check degree distribution of the right-regular LDPC ensemble (it is calculated via (112), (113), (118) where the right degree is equal to $a_{\rm R} = \lceil \frac{1}{\alpha} \rceil + 1$). This comparison refers to a BEC whose capacity is one-half (upper plot) and three-quarters (lower plot) bits per channel use, and the setting where 99.9% of the channel capacity is achieved under BP decoding with vanishing bit erasure probability. The stair functions correspond to the fraction of edges which are attached to parity-check nodes whose degrees are at most k for a positive integer k.

Note that the truncated harmonic sum $H(k)$ scales like the logarithm of k (more precisely, $H(k) \approx \ln(k) + \gamma$ for $k \gg 1$ where $\gamma \approx 0.5772$ is Euler's constant), and the value of N as given in (90) becomes un-bounded as the gap to capacity vanishes (since it is inversely proportional to ε). Hence, for small values of the gap to capacity (i.e., when $\varepsilon \ll 1$), we get from (90)

$$\alpha \approx \frac{\ln \frac{1-r}{\varepsilon r}}{1-r}, \quad N \approx \frac{1-r}{\varepsilon r} + 1$$

and therefore (86) yields that the fraction of edges which are attached to parity-check nodes of a given degree i scales like $\varepsilon^{\frac{1}{1-r}} \left(\ln \frac{1}{\varepsilon}\right)^{i-1}$ for $i \geq 1$. The upper bound on ρ_i as given in Corollary 2 scales like $\frac{\varepsilon}{\ln \frac{1}{\varepsilon}}$, where this bound is even valid under ML decoding. For a comparison between this general upper bound and the behavior of the Poisson distribution when the gap to capacity vanishes, we note that for any rate $r < 1$, a positive integer i and

$\varepsilon \ll 1$, the inequality $\varepsilon^{\frac{1}{1-r}} \left(\ln \frac{1}{\varepsilon}\right)^{i-1} \ll \frac{\varepsilon}{\ln \frac{1}{\varepsilon}}$ holds, as expected from a comparison of a degree distribution with a general upper bound. Moreover, it follows from the asymptotic analysis that for small design rates (i.e., $r \ll 1$), the Poisson distribution gets closer to the LP1 bound in the limit where $\varepsilon \to 0$ (as exemplified in Fig. 5 by comparing the upper and lower plots which correspond to a capacity of $\frac{1}{2}$ and $\frac{3}{4}$ bits per channel use, respectively).

Analytical solutions for the LP3 and LP4 bounds: Consider an LP problem of the form

$$\begin{array}{l} \text{maximize} \;\; \sum_{i=2}^{k} \lambda_i, \quad k = 2, 3, \ldots \\ \text{subject to} \\ \begin{cases} \sum_{i=2}^{\infty} \frac{\lambda_i}{i} \leq \alpha \\ \sum_{i=2}^{\infty} \lambda_i = 1 \\ \lambda_i \geq 0, \quad i = 2, 3, \ldots \end{cases} \end{array}$$

If $k\alpha \leq 1$ then the optimal solution is obtained by setting $\lambda_k = k\alpha$, $\lambda_j = 1 - k\alpha$ for some $j \to \infty$ where all the other λ_i's are set to zero. This gives a solution which is equal to $\sum_{i=1}^{k} \lambda_i = \lambda_k = k\alpha$. If $k\alpha > 1$ then the optimal solution is obtained by setting $\lambda_k = 1$ and all the other λ_i's to be zero. Hence, the solution of this LP problem is given by $\min\{k\alpha, 1\}$ which implies that the closed-form solutions of the LP3 and LP4 bounds are given by

$$\min \left\{ 1, \frac{k \ln\left(\frac{1}{g_1}\right)}{2(1-C)\left(1 + \frac{\varepsilon C}{1-C}\right) \ln\left(\frac{1}{1 - 2h_2^{-1}\left(\frac{1-C-h_2(P_{\mathrm{b}})}{1-(1-\varepsilon)C}\right)}\right)} \right\} \tag{91}$$

and

$$\min \left\{ 1, \frac{k \ln\left(\frac{1}{C}\right)}{2(1-C)\left(1 + \frac{\varepsilon C}{1-C}\right) \ln\left(\frac{1}{1 - 2h_2^{-1}\left(\frac{1-C-h_2(P_{\mathrm{b}})}{1-(1-\varepsilon)C}\right)}\right)} \right\} \tag{92}$$

respectively.

Based on the observations in Theorems 2 and 3, the fraction of edges connected to variable nodes of small degree is expected to be significantly larger than the fraction of edges which are connected to parity-check nodes of the same degree. This is shown in the following example:

*Example 7 (***LP3 bound):** Consider LDPC code ensembles whose design rate is one-half bit per channel use, and whose transmission takes place over a BIAWGN channel. Lets assume that we wish to find upper bounds on the fraction of edges up to degree k (for a parameter $k \geq 2$) for the setting of a bit error probability of (at most) $P_{\mathrm{b}} = 10^{-10}$ under ML decoding (or any sub-optimal decoding algorithm) at $\frac{E_{\mathrm{b}}}{N_0} = 0.188$ dB. This implies a gap to capacity which is equal to $\varepsilon = 1.42 \cdot 10^{-4}$. From (91), we obtain the following inequalities (also verified numerically via [15]):

$$\begin{aligned} &\lambda_2 \leq 0.2683 \\ &\lambda_2 + \lambda_3 \leq 0.4025 \\ &\lambda_2 + \lambda_3 + \lambda_4 \leq 0.5367 \\ &\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 \leq 0.6709 \\ &\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 \leq 0.8051 \\ &\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 \leq 0.9392 \\ &\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 \leq 1.0000. \end{aligned}$$

A comparison of these numerical results with those presented in Fig. 4 for the same value of $\frac{E_{\mathrm{b}}}{N_0}$ shows a big difference between the two upper bounds on the sequences $\{\lambda_i\}$ and $\{\rho_i\}$. This difference is well expected in light of the bounds in Corollary 2 where for every finite degree i, the upper bounds on λ_i and ρ_i scale like $\frac{1}{\log \frac{1}{\varepsilon}}$ and $\frac{\varepsilon}{\log \frac{1}{\varepsilon}}$, respectively. We note that this difference is not an artifact of the bounding technique, as is demonstrated in Proposition 1 for the BEC.

D. Bounds on the Graphical Complexity of Finite-Length Codes

In various applications, there is a need to design a communication system which fulfills several requirements on the available bandwidth with acceptable delay for transmitting and processing the data while maintaining a certain fidelity criterion in reconstructing the data. In this setting, one wishes to design a code which satisfies the delay constraint (i.e., the block length is limited) while adhering to the required performance over the given channel. By fixing the communication channel model and code rate (which is related to the bandwidth expansion caused by the error-correcting code), sphere-packing bounds are transformed into lower bounds on the minimal block length required to achieve a target block error probability at a certain gap to capacity using an arbitrary block code and decoding algorithm. This issue is studied in [54, Section V].

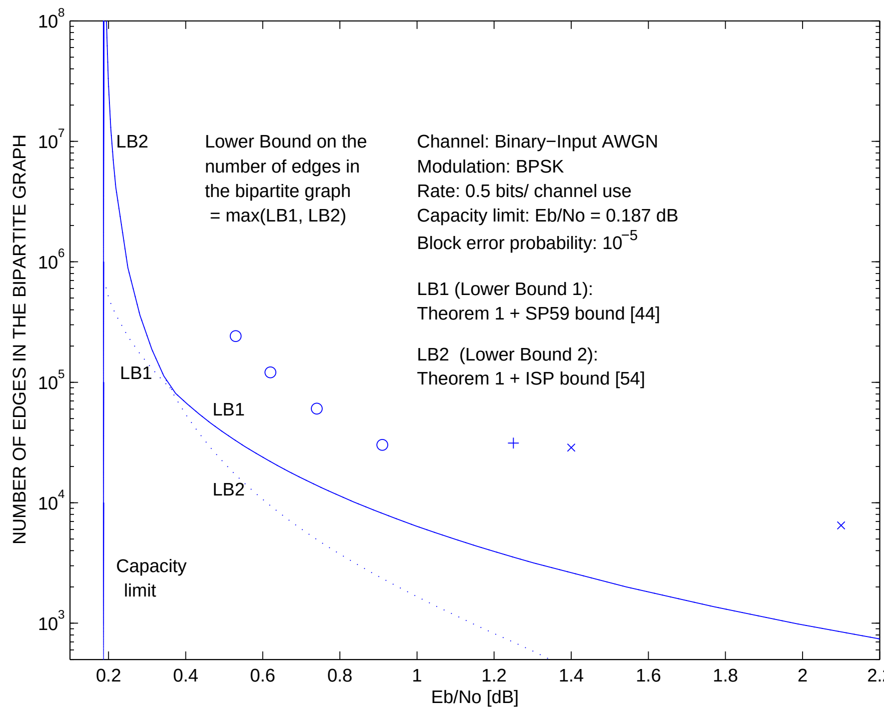

Fig. 6. A comparison between the graphical complexity of various efficient LDPC code ensembles and an information-theoretic lower bound. The graphical complexity is measured by the number of edges which are used to represent the codes (or code ensembles) by bipartite graphs in order to achieve a fixed target block error probability over a given communication channel. It is assumed that the code is BPSK modulated and transmitted over a binary-input AWGN channel. This figure refers to a target block error probability of $P_{\text{B}} = 10^{-5}$, and a design rate of one-half bit per channel use. The information-theoretic lower bound is valid under maximum-likelihood (ML) decoding (and, hence, it also holds under any sub-optimal decoding algorithm). For the comparison of the lower bound with various LDPC code ensembles, we refer to both ML and belief-propagation (BP) decoding algorithms. The circled points refer to ML decoding, and they are based on the tangential-sphere upper bound which is applied to the (6,12) regular LDPC code ensembles of Gallager for block lengths of 5040, 10080, 20160 and 40320 bits (these points rely on [50, Table II]). The other three points in this figure refer to LDPC code ensembles which are decoded by a BP decoder. The point marked by '$+$' refers to a non-punctured protograph LDPC code ensemble of block length 7360 bits and of rate one-half (see [10, Fig. 9]). The other two points which are marked by '$\times$' refer to irregular quasi-cyclic LDPC code ensembles (see [23, Figs. 10 and 11]). The two information-theoretic lower bounds on the graphical complexity ('LB1' and 'LB2') rely, respectively, on the sphere-packing bound of Shannon [44] and the recently introduced sphere-packing bound in [54]. Both of these bounds also rely on Theorem 1 which serves as a lower bound on the average right degree. The information-theoretic lower bound that is shown in this figure is obtained by taking the maximum of the LB1 and LB2 bounds.

In the following, we refer to the graphical complexity of an arbitrary bipartite graph which represents a binary linear block code. The graphical complexity has an operational meaning for an iterative message-passing decoder since the number of edges is equal to the number of right-to-left and left-to-right messages which are delivered in each iteration. As opposed to [18], [31] and [32], we refer here to the graphical complexity of *finite-length codes*. In order to evaluate an information-theoretic lower bound on the graphical complexity which is expressed in terms of the target block error probability and the corresponding achievable gap to capacity, we rely here on the following algorithm:

- *Step 1*: Sphere-packing bounds are used to calculate a lower bound on the minimal required block length in terms of the achievable rate with a target block error probability and its gap to capacity. For a memoryless symmetric channel, the lower bound on the minimal block length is calculated via the ISP bound (for finite-length codes, this recent sphere-packing bound suggests a significant improvement over the bounds in [45] and [51], see Section II-D and [54, Section III]). In addition, this lower bound is also compared with the 1959 sphere-packing (SP59) bound of Shannon (see Section II-D and [44]) for a binary-input AWGN channel where the transmitted signals are assumed to have equal energy.
- *Step 2*: A lower bound on the average right degree is calculated via Theorem 1 for an arbitrary bipartite graph which is used to represent a binary linear block code. Note that for an LDPC code whose parity-check matrix is not necessarily full-rank, one can apply this lower bound by replacing the code rate with the design rate (see Discussion 2 in Section IV). The calculation of this lower bound for a target block error probability P_{B} also stays valid if the block length n is replaced in (23) with a lower bound n' (as calculated in the previous step).
- *Step 3:* The total number of edges of a bipartite graph is a measure of its graphical complexity. For a bipartite graph which refers to a design rate of R_{d}, the total number of edges is equal to $|\mathcal{E}| = (1 - R_{\mathrm{d}})na_{\mathrm{R}}$. Replacing n and a_{R} by the lower bounds calculated in Steps 1 and 2, respectively, gives a lower bound on the number of edges.

The resulting lower bound on the total number of edges is general for every representation of a binary linear block code by a parity-check matrix and its respective bipartite graph. This bound depends on the code rate (or design rate), the communication channel, the achievable gap to capacity, and the target block error probability. This lower bound holds for an arbitrary representation of the code by a bipartite graph.

According to the above description of the three steps used to calculate the information-theoretic lower bound on the graphical complexity, we calculate here two lower bounds on the graphical complexity:

- **LB1**: A lower bound which combines a lower bound on the block length calculated via the SP59 bound [44], and a lower bound on the average right degree which is calculated via Theorem 1 for a target block error probability P_{B} and a given code rate (or design rate).
- **LB2**: A lower bound which combines a lower bound on the block length calculated via the ISP bound [54, Section III], and the same lower bound on the average right degree.

We note that Steps 2 and 3 in the above algorithm are common for the calculation of the LB1 and LB2 bounds, and the only difference in the calculation of these two bounds is in Step 1 where the SP59 and ISP bounds are used for the LB1 and LB2 bounds, respectively. The resulting lower bound (LB) on the graphical complexity is the maximal value of the LB1 and LB2 bounds, i.e., LB $= \max(\text{LB1}, \text{LB2})$. We note that the resulting lower bound on the graphical complexity holds under ML decoding or any sub-optimal decoding algorithm.

The above algorithm is applied in Figure 6 to obtain a lower bound on the graphical complexity of an arbitrary binary linear block code of rate one-half and with a target block error probability of $P_{\mathrm{B}} = 10^{-5}$. It is assumed that the code is BPSK modulated, and the transmission takes place over a binary-input AWGN channel. The unbounded complexity in the limit where the gap to capacity vanishes is due to the infinite block length which is required to obtain reliable communications at rates which are arbitrarily close to capacity. We note that the *bounded* graphical complexity for the BEC, as demonstrated in [18], [31] and [32], is obtained by addressing the graphical complexity per information bit, and by also allowing more complicated Tanner graphs which include state nodes (e.g., punctured bits) in addition to the variable and parity-check nodes which are used for a representation of these codes by bipartite graphs.

As shown in Figure 6, the bound LB2 is advantageous over LB1 for low values of $\frac{E_{\mathrm{b}}}{N_0}$ which are close to the capacity limit; this phenomenon is even more pronounced for higher code rates (above one-half bit per channel use). This observation is partially due to the fact that the ISP bound depends on the particular type of modulation used, in contrast to the SP59 bound which only assumes that the modulated signals have equal energy but does not consider the particular modulation used.

The lower bound on the graphical complexity is compared here with some efficient LDPC codes (or code ensembles) as reported in the literature. To this end, we refer to computer simulations under BP decoding, and also to upper bounds on the block error probability under ML decoding. Although the number of edges is relevant for the decoding complexity per iteration under BP decoding, some comparisons with ML decoding provide a better assessment of the tightness of this information-theoretic lower bound. The circled points in Figure 6 are based on the tangential-sphere upper bound[3] which is applied to the (6,12) regular LDPC code ensembles of Gallager for block lengths of 5040, 10080, 20160 and 40320 bits whose block error probability is upper bounded by 10^{-5} (see [50, Table II]). The other three points which are shown in Figure 6 refer to LDPC code ensembles which are decoded by a BP decoder. The point marked by '$+$' refers to a non-punctured protograph LDPC code ensemble of block length of 7360 bits and a design rate of one-half (see [10, Fig. 9]). The other two points which are marked by '$\times$' refer to irregular quasi-cyclic LDPC code ensembles (see [23, Figs. 10 and 11]) where the graphical complexity is obtained via the degree distributions which are given in [23, Examples 10 and 11]. To conclude, the information-theoretic lower bound on the graphical complexity becomes un-bounded as the gap to capacity vanishes (even under ML decoding). It also behaves in a similar way to the circled points in Figure 6 (where these points refer to the performance of a regular LDPC code ensembles under ML decoding). Moreover, the comparison of this lower bound in Figure 6 with some efficient LDPC code ensembles under BP decoding (where the corresponding points are marked by '$+$' and '$\times$') indicate the gain that can be potentially obtained by improved designs of efficient LDPC codes and iterative decoding algorithms defined on graphs.

VI. OUTLOOK

This work considers some universal properties of capacity-approaching low-density parity-check (LDPC) code ensembles whose transmission takes place over memoryless binary-input output-symmetric (MBIOS) channels. Properties of the degree distributions, graphical complexity and the fundamental cycles of the bipartite graphs are studied in this paper via the derivation of information-theoretic bounds (see Sections III and IV). The applications of these bounds are exemplified in Section V.

In the following, we gather some interesting open problems which are related to this research work:

- The analysis in this paper relies (in part) on the lower bound (18) on the conditional entropy (see [53]). Note that this bound depends on the right degree distribution (i.e., the degree distribution of the parity-check nodes), but the dependence on the left degree distribution is rather weak (according to Section II-C.2, this dependence is made only through the design rate of the LDPC code ensemble). It would be interesting to improve this bound by also having an explicit dependence on the left degree distribution. This goal can be obtained by improving the weak link in the derivation of this bound, namely, by tightening the upper bound (14) on the conditional entropy of the syndrome vector (which is expressed by the sum of the respective conditional entropies of the components of the syndrome). Note that for the BSC, the bound in (18) coincides with the bound of Gallager in [14, Section 3.8] (since the conditioning on the RHS of (14) becomes irrelevant for the BSC, due to the fact that the absolute value of the LLR is a constant for this channel). A step towards the improvement of Gallager's bound for the BSC was done by Wadayama [52] where the entropy of the syndrome vector was calculated exactly in terms of the coset weight distribution of the code (or the average coset weight distribution of the ensemble). For a general MBIOS channel, the improvement of the bound in (18) is an open problem, and it may provide an explicit dependence of the bound on the pair of degree distributions for a code which is represented by a bipartite graph.
- Unlike the information-theoretic bound in (18), the bounds presented in [28] rely on statistical physics, and therefore do not provide a bound on the conditional entropy which is valid for every binary linear block code from the considered ensembles. It would be interesting to get some theory that unifies the information-theoretic and statistical physics approaches, and provides bounds that are tight on the average and valid for each code. We note that the bounds in [28] depend on both the left and right degree distributions for LDPC code ensembles (though their computation is more complicated than the bound given in (18)).
- The asymptotic bounds in Corollary 1 address the average cardinality of the fundamental system of cycles for bipartite graphs representing LDPC code ensembles where the results are directly linked to the average right degree of these ensembles. Further study of the possible link between the statistical properties of the degree

[3] For a presentation of the tangential-sphere bound, originally introduced by Poltyrev [33], we refer the reader to [39, pp. 23–32].

distributions of capacity-approaching LDPC code ensembles and some other graphical properties related to the bipartite graphs of these ensembles is of interest.

- The graphical complexity of capacity-approaching LDPC codes is studied in this paper via an information-theoretic lower bound which relies on both Theorem 1 and sphere-packing bounds (see Section V-D). The graphical complexity is defined to be the number of edges in the bipartite graphs used to represent these codes. A recent sphere-packing bound which was introduced in [54] is shown to be helpful for the calculation of the lower bound on the graphical complexity, especially when the gap to capacity becomes small (see the algorithm for the calculation of this bound in Section V-D and the results shown in Figure 6). Further tightening of sphere-packing bounds for finite-length codes, especially for codes of short to moderate block lengths, is of interest and it has the potential of further improving the resulting lower bound on the graphical complexity. An improvement of the sphere-packing bounds introduced in [44] and [54] will also contribute to the study of the sub-optimality of iteratively decoded codes for finite block lengths.
- The derivation of universal bounds on the number of iterations of code ensembles defined on graphs, measured in terms of the achievable gap (in rate) to capacity, is of theoretical and practical interest. In a recent work [42], this issue is addressed for the BEC. It is demonstrated in [42] that the number of iterations which is required for successful message-passing decoding scales at least like the inverse of the achievable gap (in rate) to capacity, provided that the fraction of degree-2 variable nodes of these turbo-like code ensembles does not vanish (hence, the number of iterations becomes unbounded as the gap to capacity vanishes). Note that Lemma 7 (see p. 24) provides a condition which ensures that the fraction of degree-2 variable nodes stays strictly positive for capacity-achieving LDPC code ensembles. A generalization of such a lower bound on the number of iterations for an arbitrary MBIOS channel is of interest. The matching condition for generalized extrinsic information transfer (GEXIT) curves serves to conjecture in [26, Section XI] that, also for an arbitrary MBIOS channel, this number of iterations scales like the inverse of the achievable gap to capacity.
- Extension of the results in this paper to channels with memory (e.g., finite-state channels) is of interest. In this respect, the reader is referred to [17] which considers information-theoretic bounds on the achievable rates of LDPC code ensembles for a class of finite-state channels.
- Extension of the results in this work to general ensembles of multi-edge type LDPC codes (see [37, Chapter 7]) is of interest.

APPENDIX I
PROOF OF LEMMA 3

The following proof deviates from the analysis in Section II-C.1, starting from (14).

- In the transition to the last line in (14), the conditional entropy $H\big(\mathbf{S}\,|\,\Omega_1,\ldots,\Omega_n\big)$ is upper bounded by the sum of the conditional entropies of the $n(1-R)$ independent components of the syndrome $\mathbf{S}$ under the assumption that the parity-check matrix is full-rank. In the general case where this parity-check matrix is not necessarily full-rank, the rate R of the code may exceed the design rate $R_{\rm d}$ due to a possible linear dependence of the rows in this matrix. Therefore, we obtain an upper bound on the conditional entropy by summing over the $n(1-R_{\rm d})$ components of the syndrome.
- In parallel to (14), we get the inequality

$$\begin{aligned}&H\big(\Phi_1,\ldots,\Phi_n\,|\,\Omega_1,\ldots,\Omega_n\big)\\&\leq H(M)+\sum_{j=1}^{n(1-R_{\rm d})} H\big(S_j\,|\,\Omega_1,\ldots,\Omega_n\big)\,.\end{aligned}\tag{93}$$

- The entropy of the transmitted codeword $\mathbf{X}$ is equal to the entropy of the index M of the received vector in the appropriate coset, regardless of the rank of H. Hence, $H(\mathbf{X})$ in the second line of (10) can be replaced by $H(M)$, and we get

$$H(\mathbf{X}|\mathbf{Y})=H(M)+n[H(\widetilde{Y}_1)-I(X_1;\widetilde{Y}_1)]-H(\widetilde{\mathbf{Y}}).$$

- Combining (11)–(13), (93) and the last equality, we get the inequality (note that the entropy $H(M)$ cancels out)

$$H(\mathbf{X}|\mathbf{Y}) \geq n(1-C) - \sum_{j=1}^{n(1-R_{\text{d}})} H(S_j \big| \Omega_1, \ldots, \Omega_n)$$

  which is similar to (15) except that the sum on the RHS is over the $n(1-R_{\text{d}})$ (possibly linearly-dependent) components of the syndrome.

From this point, the analysis is similar to Section II-C.1 which then yields an extension of (18) with R replaced by R_{d} when the parity-check matrix is not necessarily full-rank.

APPENDIX II
PROOF OF LEMMA 5

This lemma is proved by expressing the channel capacity as a non-negative infinite series which depends on the sequence $\{g_k\}_{k\geq 1}$, and solving an optimization problem for the extreme values of g_1 subject to a constraint on the channel capacity C. To this end, we rely on the equality in (7) for the capacity of an MBIOS channel:

$$\begin{aligned} C &= \int_0^\infty a(l)(1+e^{-l}) \left[1 - h_2\left(\frac{1}{1+e^l}\right)\right] \, \mathrm{d}l \\ &\stackrel{\text{(a)}}{=} \int_0^\infty a(l)(1+e^{-l}) \, \frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{\tanh^{2k}\left(\frac{l}{2}\right)}{k(2k-1)} \, \mathrm{d}l \\ &= \frac{1}{2\ln 2} \sum_{k=1}^{\infty} \left\{ \frac{\int_0^\infty a(l)(1+e^{-l}) \tanh^{2k}\left(\frac{l}{2}\right) \, \mathrm{d}l}{k(2k-1)} \right\} \\ &\stackrel{\text{(b)}}{=} \frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{g_k}{k(2k-1)} \end{aligned} \tag{94}$$

where equality (a) follows by substituting $x = \frac{1}{1+e^l}$ in (16), and equality (b) follows from (19); this provides an expression for the channel capacity in terms of the non-negative sequence $\{g_k\}_{k=0}^{\infty}$ defined in (19). The representation of the capacity as the infinite series in (94) follows in fact from the result which is obtained via [46, Propositions 3.1–3.3] by referring to an equi-probable binary input, though the derivation here is more direct.

We start with the proof of the upper bound on g_1, as given on the RHS of (53). Since we look for the maximal value of g_1 among all MBIOS channels with a given capacity C, then we need to solve the optimization problem

$$\begin{aligned} &\text{maximize} \quad g_1 \\ &\text{subject to} \quad \frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{g_k}{k(2k-1)} = C. \end{aligned} \tag{95}$$

Based on Lemma 4, for every MBIOS channel, $g_k \geq (g_1)^k$ for all $k \in \mathbb{N}$. Therefore

$$\begin{aligned} &\frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{g_k}{k(2k-1)} \\ &\geq \frac{1}{2\ln 2} \sum_{k=1}^{\infty} \frac{(g_1)^k}{k(2k-1)} \\ &= 1 - h_2\left(\frac{1-\sqrt{g_1}}{2}\right) \end{aligned} \tag{96}$$

where the last equality is based on (42). The equality constraint in (95) and the inequality (96) yield that

$$1 - h_2\left(\frac{1-\sqrt{g_1}}{2}\right) \leq C$$

from which the RHS of (53) follows. Note that this upper bound on g_1 is attained when $g_k = (g_1)^k$ for all $k \in \mathbb{N}$. To show this equality, note that for a BSC with crossover probability p, the LLR at the channel output (L) is bimodal and it gets the values $l_1 = +\ln\left(\frac{1-p}{p}\right)$ and $l_2 = -l_1$ with probabilities $1-p$ and p, respectively. Eq. (20) then gives

$$\begin{aligned} g_k &\triangleq \mathbb{E}\left[\tanh^{2k}\left(\frac{L}{2}\right)\right] \\ &= (1-p)\tanh^{2k}\left(\frac{l_1}{2}\right) + p\tanh^{2k}\left(-\frac{l_1}{2}\right) \\ &= \tanh^{2k}\left(\frac{l_1}{2}\right) \\ &= \left(\frac{e^{l_1}-1}{e^{l_1}+1}\right)^{2k} \\ &= (1-2p)^{2k}, \quad \forall\, k \in \mathbb{N}. \end{aligned} \tag{97}$$

Hence for the BSC, $g_k = (g_1)^k$ for all $k \in \mathbb{N}$. The upper bound on g_1 on the RHS of (53) is therefore achieved for a BSC whose crossover probability is $p = h_2^{-1}(1-C)$.

The proof of the lower bound on g_1 relies on (94). Since the sequence $\{g_k\}_{k\geq 1}$ is monotonically non-increasing and non-negative (this property follows directly from (20)), then

$$\begin{aligned} C &= \frac{1}{2\ln 2}\sum_{k=1}^{\infty}\frac{g_k}{k(2k-1)} \\ &\leq \frac{g_1}{2\ln 2}\sum_{k=1}^{\infty}\frac{1}{k(2k-1)} \\ &= g_1 \end{aligned}$$

where the last equality follows from (47). This lower bound on g_1 is attained for a BEC (since for a BEC whose erasure probability is p, (20) implies that the sequence $\{g_k\}$ is constant and $g_1 = 1-p = C$).

Appendix III
Proof of Lemma 7

From the assumption in Lemma 7, the satisfiability of the flatness condition for this capacity-achieving sequence gives that

$$\lim_{m\to\infty} \mathcal{B}(a)\,\lambda_2^{(m)}\,\rho_m'(1) = 1\,. \tag{98}$$

From (75), the fraction of degree-2 variable nodes is given by

$$\Lambda_2^{(m)} = \frac{\lambda_2^{(m)}}{2\int_0^1 \lambda_m(x)\mathrm{d}x}\,, \qquad \forall m \in \mathbb{N} \tag{99}$$

and therefore

$$\begin{aligned} &\lim_{m\to\infty}\Lambda_2^{(m)} \\ &\stackrel{\text{(a)}}{=} \lim_{m\to\infty}\frac{1}{2\,\mathcal{B}(a)\,\rho_m'(1)\,\int_0^1\lambda_m(x)\mathrm{d}x} \\ &\stackrel{\text{(b)}}{=} \lim_{m\to\infty}\frac{1-R_m}{2\,\mathcal{B}(a)\,\rho_m'(1)\,\int_0^1\rho_m(x)\mathrm{d}x} \\ &\stackrel{\text{(c)}}{=} \frac{1-C}{2\,\mathcal{B}(a)}\lim_{m\to\infty}\frac{1}{\rho_m'(1)\,\int_0^1\rho_m(x)\mathrm{d}x} \end{aligned} \tag{100}$$

where (a) relies on (98) and (99), (b) follows from (3) where R_m designates the design rate of the m-th LDPC code ensemble in this sequence, and (c) follows by the assumption that the sequence is capacity-achieving. Let $a_{\mathrm{R}}^{(m)}$ designate the average right degree of the LDPC code ensemble (n_m, λ_n, ρ_m). From (5), this implies that

$a_{\rm R}^{(m)} = \left(\int_0^1 \rho_m(x)\,{\rm d}x\right)^{-1}$ and, from Theorem 1 followed by Discussion 2, the asymptotic average right degree of the considered capacity-achieving sequence tends to infinity, i.e.,

$$\lim_{m\to\infty} a_{\rm R}^{(m)} = \infty. \tag{101}$$

We evaluate now the expression in (100). To this end, let $\rho_m(x) \triangleq \sum_i \rho_i^{(m)} x^{i-1}$, and let $\Gamma_i^{(m)}$ designate the fraction of parity-check nodes of degree i for LDPC code ensemble (n_m, λ_m, ρ_m), then

$$\begin{aligned}
&\rho_m'(1) \int_0^1 \rho_m(x){\rm d}x \\
&= \frac{\rho_m'(1)}{a_{\rm R}^{(m)}} \\
&= \frac{\sum_i (i-1)\rho_i^{(m)}}{a_{\rm R}^{(m)}} \\
&= \frac{\sum_i i\rho_i^{(m)} - 1}{a_{\rm R}^{(m)}} \\
&= \frac{1}{a_{\rm R}^{(m)}} \left(\sum_i i\left(\frac{i\Gamma_i^{(m)}}{\sum_j j\Gamma_j^{(m)}}\right) - 1\right) \\
&= \frac{\sum_i i^2\Gamma_i^{(m)}}{\left(a_{\rm R}^{(m)}\right)^2} - \frac{1}{a_{\rm R}^{(m)}} \\
&= \left(\frac{\sqrt{\sum_i i^2\Gamma_i^{(m)} - \left(a_{\rm R}^{(m)}\right)^2}}{a_{\rm R}^{(m)}}\right)^2 + 1 - \frac{1}{a_{\rm R}^{(m)}}.
\end{aligned} \tag{102}$$

Consider any code from the LDPC code ensemble (n_m, λ_m, ρ_m). Note that the first term in (102) is the square of the ratio of the standard deviation and the average degree of the parity-check nodes for this code. Since we denote the asymptotic limit of this ratio by K (where we assume that it exists and is finite) and also (101) holds, then we get from (102) that

$$\lim_{m\to\infty} \rho_m'(1) \int_0^1 \rho_m(x){\rm d}x = K^2 + 1. \tag{103}$$

This completes the proof of the theorem by combining (100) with (103).

Appendix IV
Proof of Lemma 8

Let a denote the symmetric L-density *pdf* of the transition probability of an MBIOS channel (see [37, Theorem 4.26]). Let $C = C(a)$ and $B = \mathcal{B}(a)$ be the corresponding capacity and Bhattacharyya constant, respectively.

From (6), (8) and the symmetry of a

$$\begin{aligned}
& C+B-1 \\
&= \int_{-\infty}^{\infty} a(l) e^{-\frac{l}{2}} \, \mathrm{d}l - \int_{-\infty}^{\infty} a(l) \log_2(1+e^{-l}) \, \mathrm{d}l \\
&= \int_{-\infty}^{\infty} a(l) e^{-\frac{l}{2}} \, \mathrm{d}l \\
&\quad -\frac{1}{2} \int_{-\infty}^{\infty} \Big[a(l) \log_2(1+e^{-l}) + a(-l) \log_2(1+e^{l})\Big] \mathrm{d}l \\
&= \int_{-\infty}^{\infty} a(l) e^{-\frac{l}{2}} \, \mathrm{d}l \\
&\quad -\frac{1}{2} \int_{-\infty}^{\infty} a(l) \Big[\log_2(1+e^{-l}) + e^{-l} \log_2(1+e^{l})\Big] \mathrm{d}l \\
&= \int_{-\infty}^{\infty} e^{-\frac{l}{2}} \, a(l) g(l) \mathrm{d}l
\end{aligned}$$

where the function g is given by

$$g(l) = 1 - \frac{1}{2}\Big[e^{\frac{l}{2}} \log_2(1+e^{-l}) + e^{-\frac{l}{2}} \log_2(1+e^{l})\Big], \quad l \in \mathbb{R}.$$

In order to complete the proof, it suffices to show that the function g is non-negative. The substitution $x = \frac{1}{1+e^l}$ gives $g(l) = 1 - \frac{h_2(x)}{2\sqrt{x(1-x)}}$ where the interval $(-\infty, +\infty)$ for l is mapped into the interval $(0,1)$ for x. The non-negativity of g follows from the inequality $h_2(x) \leq 2\sqrt{x(1-x)}$ which is satisfied for $0 \leq x \leq 1$. The non-negativity of the function g implies that $C + B \geq 1$.

Note that for a BEC with erasure probability p, the channel capacity is $1-p$ bits per channel use, and the Bhattacharyya constant is equal to p. Hence, the equality $C+B=1$ holds for every BEC, irrespectively of the channel erasure probability.

Appendix V
Proof of Corollary 4

A truncation of the power series on the LHS of (42) after its first term gives the inequality

$$1 - h_2\left(\frac{1-\sqrt{u}}{2}\right) \geq \frac{u}{2\ln 2}, \quad 0 \leq u \leq 1.$$

Assigning $u = \big(1 - 2h_2^{-1}(x)\big)^2$ and rearranging terms gives

$$h_2^{-1}(x) \geq \frac{1}{2}\left(1 - \sqrt{2\ln 2 \ (1-x)}\right), \quad 0 \leq x \leq 1. \tag{104}$$

Assigning $0 \leq x \triangleq \frac{1-C}{1-(1-\varepsilon)C} \leq 1$ in (104) gives

$$\begin{aligned}
& h_2^{-1}\left(\frac{1-C}{1-C(1-\varepsilon)}\right) \\
&\geq \frac{1}{2}\left(1 - \sqrt{2\ln 2 \ \left(\frac{\varepsilon C}{1-(1-\varepsilon)C}\right)}\right) \\
&\geq \frac{1}{2}\left(1 - \sqrt{2\ln 2 \ \left(\frac{\varepsilon C}{1-C}\right)}\right)
\end{aligned}$$

and therefore

$$1 - 2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right) \leq \sqrt{2\ln 2 \ \left(\frac{\varepsilon C}{1-C}\right)}. \tag{105}$$

Substituting (105) in (80) provides the following lower bound on the average right degree of the ensembles:

$$a_{\rm R} \geq \frac{\ln\left(\frac{1}{2\ln 2}\,\frac{1-C}{\varepsilon C}\right)}{\ln\left(\frac{1}{g_1}\right)}. \tag{106}$$

As the average right degree of an LDPC code ensemble is not less than 2 (as otherwise, some bits are forced to be zeros and can be deleted from all codewords), then it follows from (106) that

$$\begin{aligned} a_{\rm R} - 1 \;&\geq\; \left[\frac{\ln\left(\frac{g_1}{2\ln 2}\,\frac{1-C}{\varepsilon C}\right)}{\ln\left(\frac{1}{g_1}\right)}\right]^{+} \\ &= \left[\frac{\ln\left(\frac{g_1}{2\ln 2}\,\frac{1-C}{C}\right) + \ln\left(\frac{1}{\varepsilon}\right)}{\ln\left(\frac{1}{g_1}\right)}\right]^{+}. \end{aligned} \tag{107}$$

The proof is completed by combining (81) with (107).

APPENDIX VI
PROOF OF PROPOSITION 1

When the transmission takes place over a BEC whose erasure probability is p, the constant c_2 given in (38) takes the form

$$c_2 = \frac{p}{\ln\left(\frac{1}{1-p}\right)}\;. \tag{108}$$

The starting point of this proof follows the concept in [37, Example 3.88], and its continuation relies on the analysis used for the proof of [40, Theorem 2.3]. For $0 < \alpha < 1$, let

$$\begin{aligned} \hat{\lambda}_\alpha(x) &= 1-(1-x)^\alpha = \sum_{k=1}^{\infty} (-1)^{k+1} \binom{\alpha}{k} x^k\,, \quad 0 \leq x \leq 1 \\ \rho_\alpha(x) &= x^{\frac{1}{\alpha}}. \end{aligned} \tag{109}$$

Note that all the coefficients in the power series expansion of $\hat{\lambda}_\alpha$ are positive for all $0 < \alpha < 1$. Let us now define the polynomials $\hat{\lambda}_\alpha^{(N)}$ and $\lambda_\alpha^{(N)}$ where $\hat{\lambda}_\alpha^{(N)}(x)$ is the truncated power series of $\hat{\lambda}_\alpha(x)$ around $x = 0$, consisting of all the terms up to (and including) the term x^{N-1}, and the polynomial

$$\lambda_\alpha^{(N)}(x) \triangleq \frac{\hat{\lambda}_\alpha^{(N)}(x)}{\hat{\lambda}_\alpha^{(N)}(1)} \tag{110}$$

is normalized to satisfy the equality $\lambda_\alpha^{(N)}(1) = 1$. The sequence of right-regular LDPC code ensembles in [48] is of the form $\left\{\left(n_m, \lambda_\alpha^{(N)}(x), \rho_\alpha(x)\right)\right\}_{m \geq 1}$ where $0 < \alpha < 1$ and $N \in \mathbb{N}$ are arbitrary parameters which need to be selected properly. Assume that the transmission takes place over a BEC whose erasure probability is p. Based on the proof of [40, Theorem 2.3], this sequence achieves a fraction $1-\varepsilon$ of the capacity of the BEC with vanishing bit erasure probability under BP decoding when α and N are chosen to satisfy

$$\frac{1}{N^\alpha} = 1-p \tag{111}$$

$$N = \max\left(\left\lceil \frac{1-(1-\varepsilon)(1-p)k_2(p)}{\varepsilon}\right\rceil, \left\lceil (1-p)^{-\frac{1}{p}}\right\rceil\right) \tag{112}$$

where

$$k_2(p) \triangleq (1-p)^{\frac{\pi^2}{6}}\; e^{\left(\frac{\pi^2}{6}-\gamma\right)\,p} \tag{113}$$

and γ is Euler's constant ($\gamma \approx 0.5772$). Combining (109) and (110), and using the equality

$$\sum_{k=1}^{N-1} (-1)^{k+1} \binom{\alpha}{k} = 1 - \frac{N}{\alpha}\binom{\alpha}{N}(-1)^{N+1}$$

gives

$$\lambda_{\alpha}^{(N)}(x) = \frac{\sum_{k=1}^{N-1} (-1)^{k+1} \binom{\alpha}{k} x^k}{1 - \frac{N}{\alpha} (-1)^{N+1} \binom{\alpha}{N}} .$$

Therefore, the fraction of edges adjacent to variable nodes of degree two is given by

$$\lambda_2 = \frac{\alpha}{1 - \frac{N}{\alpha} (-1)^{N+1} \binom{\alpha}{N}} . \tag{114}$$

We now obtain upper and lower bounds on λ_2. From [40, Eq. (67)] we have that

$$\frac{c(\alpha, N)}{N^{\alpha}} < \frac{N}{\alpha} (-1)^{N+1} \binom{\alpha}{N} \leq \frac{1}{N^{\alpha}} \tag{115}$$

where

$$c(\alpha, N) \triangleq (1-\alpha)^{\frac{\pi^2}{6}} \; e^{\alpha \left(\frac{\pi^2}{6} - \gamma + \frac{1}{2N} \right)} . \tag{116}$$

Substituting (115) in (114) and using (111), we get

$$\frac{\alpha}{1 - c(\alpha, N)\,(1-p)} < \lambda_2 \leq \frac{\alpha}{1 - (1-p)} = \frac{\alpha}{p} . \tag{117}$$

Under the parameter assignments in (111) and (112), the parameters N and α satisfy

$$\alpha = \frac{\ln\left(\frac{1}{1-p}\right)}{\ln N} \tag{118}$$

$$N \geq \frac{1 - (1-p)\, k_2(p)}{\varepsilon} . \tag{119}$$

Substituting (118) and (119) into the inequality on the RHS of (117) gives an upper bound on λ_2 which takes the form

$$\begin{aligned} \lambda_2 &\leq \frac{\alpha}{p} \\ &\leq \frac{\ln\left(\frac{1}{1-p}\right)}{p \, \ln\left(\frac{1-(1-p)\,k_2(p)}{\varepsilon}\right)} \\ &= \frac{1}{c_3 + c_2 \ln \frac{1}{\varepsilon}} \end{aligned} \tag{120}$$

where c_2 is the coefficient of the logarithmic growth rate in $\frac{1}{\varepsilon}$, which coincides here with (108), and

$$c_3 \triangleq \frac{p \, \ln\big(1 - (1-p)\, k_2(p)\big)}{\ln\left(\frac{1}{1-p}\right)} \tag{121}$$

is a constant which only depends on the BEC. We turn now to derive a lower bound on λ_2, and then examine it in the limit where the gap to capacity vanishes. From (112), we have that for small enough values of ε, the parameter N satisfies

$$\begin{aligned} N &= \left\lceil \frac{1 - k_2(p)\,(1-p)\,(1-\varepsilon)}{\varepsilon} \right\rceil \\ &\leq \frac{1 - k_2(p)\,(1-p)\,(1-\varepsilon)}{\varepsilon} + 1. \end{aligned} \tag{122}$$

Substituting (118) and (122) into the inequality on the LHS of (117), we get

$$\begin{aligned} \lambda_2 &> \frac{\alpha}{p} \; \frac{p}{1 - c(\alpha, N)\,(1-p)} \\ &\geq \frac{\ln\left(\frac{1}{1-p}\right)}{p \, \ln\left(\frac{1-k_2(p)(1-p)(1-\varepsilon)+\varepsilon}{\varepsilon}\right)} \cdot \frac{p}{1 - c(\alpha, N)\,(1-p)} \\ &= \frac{1}{c_3 + c_2 \ln\left(\frac{1}{\varepsilon}\right) + \widetilde{\varepsilon}(\varepsilon, p)} \; \frac{p}{1 - (1-p)\, c(\alpha, N)} \end{aligned} \tag{123}$$

where c_2 is the coefficient of the logarithm in the denominator of (37) and it coincides with (108) for the BEC, c_3 is given in (121), and

$$\widetilde{\varepsilon}(\varepsilon, p) \triangleq \frac{p \ln\left(1 + \frac{\varepsilon\left(1+k_2(p)\,(1-p)\right)}{1-k_2(p)\,(1-p)}\right)}{\ln\left(\frac{1}{1-p}\right)}$$

which therefore implies that for $0 \leq p < 1$

$$\lim_{\varepsilon \rightarrow 0} \widetilde{\varepsilon}(\varepsilon, p) = 0. \tag{124}$$

Using the lower bound on the parameter N in (119), in the limit where ε tends to zero, the parameter N tends to infinity (since $1-(1-p)k_2(p) > 0$ for all $0 < p < 1$ where k_2 in introduced in (113)). Also, from (112) and (118), we get

$$\lim_{\varepsilon \rightarrow 0} \alpha = 0$$

which, from (116), yields that

$$\lim_{\varepsilon \rightarrow 0} c(\alpha, N) = 1\,. \tag{125}$$

Substituting (124) and (125) in (123) yields that in the limit where the gap to capacity vanishes (i.e., $\varepsilon \rightarrow 0$), the upper and lower bounds on λ_2 in (120) and (123) coincide. Specifically, we have shown that

$$\lim_{\varepsilon \rightarrow 0} \lambda_2(\varepsilon) \cdot c_2 \, \ln\left(\frac{1}{\varepsilon}\right) = 1\,.$$

Therefore, as $\varepsilon \rightarrow 0$, the upper bound on $\lambda_2 = \lambda_2(\varepsilon)$ in Corollary 4 becomes tight for the sequence of right-regular LDPC code ensembles in [48] with the parameters chosen in (111) and (112). We note that the setting of the parameters N and α in (111) and (112) is identical to [40, p. 1615].

APPENDIX VII
A PROOF OF REMARK 15

We prove in the following the claim in Remark 15 which states that adding the constraint that is imposed by the lower bound on the average right degree (i.e., the lower bound on $a_{\text{R}} = \sum_{i=1}^{\infty} i\Gamma_i$) does not affect the LP1 and LP2 bounds introduced in Section V-C. More explicitly, for the LP1 bound, we prove that the constraint on $\{\Gamma_i\}_{i\geq 1}$ which is imposed by (61) implies the lower bound on the average right degree as given in (22) and (23).

Proof: Eq. (61) gives the first constraint in the LP1 bound. By substituting $x = \frac{1-g_1^{\frac{i}{2}}}{2}$ in (16), we get that the following equality holds for $i \geq 1$ (note that since $0 \leq g_1 \leq 1$ then $0 \leq x \leq 1$ as required in (16)):

$$1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2}\right) = \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{g_1^{pi}}{p(2p-1)}.$$

Plugging this equality into the LHS of (61) gives

$$\begin{aligned}
&\sum_{i=1}^{\infty} \left\{ \left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2}\right)\right] \Gamma_i \right\} \\
&\stackrel{(a)}{=} \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \sum_{i=1}^{\infty} \frac{\Gamma_i g_1^{pi}}{p(2p-1)} \\
&\stackrel{(b)}{\geq} \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{{g_1}^{p\sum_i i\Gamma_i}}{p(2p-1)} \\
&\stackrel{(c)}{=} \frac{1}{2\ln 2} \sum_{p=1}^{\infty} \frac{{g_1}^{pa_{\text{R}}}}{p(2p-1)} \\
&\stackrel{(d)}{=} 1 - h_2\left(\frac{1-g_1^{\frac{a_{\text{R}}}{2}}}{2}\right).
\end{aligned} \tag{126}$$

where equality (a) is obtained by interchanging the order of summation, equality (b) follows from Jensen's inequality, equality (c) follows from expressing the average right degree by the equality $a_{\text{R}} = \sum_i i\Gamma_i$, and equality (d) follows from (16). Combining (61) with (126) gives that

$$1 - h_2\left(\frac{1 - g_1^{\frac{a_{\text{R}}}{2}}}{2}\right) \leq \frac{\varepsilon\, C + h_2(P_{\text{b}})}{1 - (1-\varepsilon)C}$$

and then some straightforward algebra implies that

$$a_{\text{R}} \geq \frac{2\ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_{\text{b}})}{1-(1-\varepsilon)C}\right)}\right)}{\ln\left(\frac{1}{g_1}\right)}.$$

This lower bound on the average right degree coincides with the bound in (22) and (23) which then completes our proof for the LP1 bound. The same proof holds for the LP2 bound while referring to the lower bound given in (23) and (54). ■

APPENDIX VIII
ANALYTICAL SOLUTION OF THE LP1 BOUND

The LP1 bound in Section V-C can be equivalently expressed as the following minimization problem:

$$\begin{array}{l} \text{minimize} \quad -\sum_{i=1}^{k} \rho_i, \quad k = 1, 2, \ldots \\ \text{subject to} \\ \begin{cases} \sum_{i=1}^{\infty} d_i \rho_i \leq 0 \\ d_i \triangleq \frac{1}{i}\left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2}\right) - \frac{\varepsilon\, C + h_2(P_{\text{b}})}{1-(1-\varepsilon)C}\right], \ i \geq 1 \\ \sum_{i=1}^{\infty} \rho_i \leq 1 \\ \rho_i \geq 0, \quad i = 1, 2, \ldots \end{cases} \end{array}$$

where we negated the objective function and turned the maximization into a minimization, and also the equality constraint on $\sum_{i\geq 1} \rho_i$ was turned into an inequality constraint. By introducing the non-negative Lagrange multipliers μ_1 and μ_2, respectively, to the first and second inequality constraints, and also introducing the non-negative Lagrange multiplies $\{\theta_i\}$ to the non-negativity constraint on $\{\rho_i\}$, we get the Lagrangian

$$\begin{aligned} & L(\{\rho_i\}, \mu_1, \mu_2, \{\theta_i\}) \\ & = -\sum_{i=1}^{k} \rho_i + \mu_1 \sum_{i=1}^{\infty} d_i \rho_i + \mu_2 \left(\sum_{i=1}^{\infty} \rho_i - 1\right) - \sum_{i=1}^{\infty} \theta_i \rho_i \\ & = \sum_{i=1}^{k} \bigl(-1 + \mu_1 d_i + \mu_2 - \theta_i\bigr)\rho_i + \sum_{i=k+1}^{\infty} \bigl(\mu_1 d_i + \mu_2 - \theta_i\bigr)\rho_i \\ & \quad -\mu_2. \end{aligned} \tag{127}$$

By alternating again the sign of the objective function, we get the following dual LP problem:

$$\begin{array}{l} \text{minimize } \mu_2 \\ \text{subject to} \\ \begin{cases} -1 + \mu_1 d_i + \mu_2 - \theta_i = 0, \quad i = 1, 2, \ldots, k \\ \mu_1 d_i + \mu_2 - \theta_i = 0, \quad i = k+1, k+2, \ldots \\ \mu_1, \mu_2 \geq 0 \\ \theta_i \geq 0, \quad i = 1, 2, \ldots \end{cases} \end{array}$$

Strong duality holds for linear programming provided that the primal LP or its dual LP are feasible (see [6, Problem 5.23]). Hence, strong duality holds for the LP1 problem.

Note that the sequence $\{d_i\}$ (see the above primal problem) is positive if and only if $i < k_0$ where k_0 denotes the lower bound on the average right degree as is given in (22). For $k < k_0$, the sequence $\{d_i\}_{i=1}^k$ is positive and monotonic decreasing:

$$d_1 > d_2 > \ldots, > d_k > 0, \quad \forall\, k < k_0.$$

Also $d_i \leq 0$ for $i \geq k_0$, and $\lim_{i \to \infty} d_i = 0$. Let

$$d^* \triangleq \min_{i \geq 1} d_i \tag{128}$$

where the minimum of the sequence $\{d_i\}$ is attained for some index $i \geq k_0$, and $d^* \leq 0$ (note that except for the degenerate case where $g_1 = 0$, for which the channel is completely useless, the sequence $\{d_i\}$ is negative for $i > k_0$, and it tends asymptotically to zero in the limit where $i \rightarrow \infty$).

Let $k < k_0$. Due to the properties of the sequence $\{d_i\}$ and the non-negativity constraint on $\{\theta_i\}$ in the dual LP problem, the minimization of the objective function (μ_2) can be simplified. To this end, one can remove all the equality constraints from the dual LP problem except of the first equality constraint with the index $i = k$, and the second equality constraint with the index $i \geq k_0$ for which the sequence $\{d_i\}$ attains its minimal value (d^*). For these two indices of i, the Lagrange multipliers θ_i in the two equality constraints of the dual LP problem are set to zero; this setting attains the minimal value of μ_2 (for the other equality constraints that were removed from the dual LP problem, the corresponding θ_i's are strictly positive; however, these equality constraints are redundant for the minimization of μ_2 in the dual LP). Hence, for $k < k_0$, the dual LP problem is simplified to

minimize μ_2
subject to

$$\begin{cases} -1 + \mu_1 d_k + \mu_2 = 0 \\ \mu_1 d^* + \mu_2 = 0 \\ \mu_1, \mu_2 \geq 0 \end{cases}$$

whose solution is

$$\mu_1 = \frac{1}{d_k - d^*}, \quad \mu_2 = -\frac{d^*}{d_k - d^*}$$

and the optimal value of the dual LP is equal to $-\frac{d^*}{d_k - d^*}$ which is indeed bounded between 0 and 1 (since $d^* \leq 0$ and $d_k > 0$ for $k < k_0$).

For $k \geq k_0$, we get the following system of inequalities from the dual LP problem:

$$\begin{cases} -1 + \mu_1 d_i + \mu_2 \geq 0, & \text{for } i = 1, 2, \ldots, k \\ \mu_1 d_i + \mu_2 \geq 0, & \text{for } i = k+1, k+2, \ldots \end{cases}$$

Since $d_k \leq 0$, then the optimal solution of the dual LP is obtained at $\mu_1 = 0$ and $\mu_2 = 1$, which then gives an optimal value of 1 for the minimization of μ_2.

Remark 17: : Consider again the solution of the LP1 problem in the case where $k \leq k_0$. From the solution of the dual problem, it follows that it is obtained by setting ρ_i to be zero, except for two indices. To this end, let $i = l$ be the index for which the sequence $\{d_i\}$ achieves its negative minimal value (d^*), and let us choose the values of ρ_k and ρ_l to satisfy the two equalities:

$$d_k \rho_k + d_l \rho_l = 0$$

$$\rho_k + \rho_l = 1.$$

Since $d^* = d_l$ for some $l > k_0$, then for $k \leq k_0$ and the above selection of $\{\rho_i\}$

$$\sum_{i=1}^{k} \rho_i = \rho_k = -\frac{d^*}{d_k - d^*}$$

which indeed coincides with the solution of the dual problem.

A Tightened Version of the LP1 bound for the BEC and its Analytical Solution

A tightened version of the LP1 bound for the BEC is obtained via (64). By substituting the equality (85) into the LHS of (64) and using the equality $C = 1 - p$ for a BEC gives

$$\sum_{i=1}^{\infty} \frac{\rho_i C^i}{i} \leq \frac{\varepsilon C + P_{\text{b}}}{1-(1-\varepsilon)C} \sum_{i=1}^{\infty} \frac{\rho_i}{i}\,.$$

This inequality constraint forms a tightened constraint for the BEC, as compared to the first inequality constraint which was formulated in the LP1 problem for a general MBIOS channel. In order to use the analytical result derived earlier in this appendix and adapt it to this case, we formulate the tightened version of the LP1 bound for the BEC as follows:

$$\boxed{\begin{array}{l} \text{minimize} \quad -\sum_{i=1}^{k} \rho_i, \quad k = 1, 2, \ldots \\ \text{subject to} \\ \left\{ \begin{array}{l} \sum_{i=1}^{\infty} d_i \rho_i \leq 0 \\ d_i \triangleq \frac{1}{i}\left(C^i - \frac{\varepsilon C + P_{\text{b}}}{1-(1-\varepsilon)C}\right), \quad i = 1, 2, \ldots \\ \sum_{i=1}^{\infty} \rho_i \leq 1 \\ \rho_i \geq 0, \quad i = 1, 2, \ldots \end{array} \right. \end{array}}$$

Similarly to the above analysis in this appendix, the new sequence $\{d_i\}$ is non-negative if and only if $i \leq k_0$ where k_0 denotes the lower bound on the average right degree as is given in (24). For $k \leq k_0$, the sequence $\{d_i\}_{i=1}^{k}$ is non-negative and monotonic decreasing; moreover, $d_i < 0$ for $i > k_0$, and $\lim_{i\to\infty} d_i = 0$. By using the same notation of d^* in (128), we obtain that the tightened version of the LP1 bound for the BEC has the same analytical solution as of the general LP1 bound, except for the change of the sequence $\{d_i\}$ (and its corresponding minima d^*).

Acknowledgment

Discussions with Henry Pfister, Tom Richardson and Ruediger Urbanke during the 2008 Turbo Coding Symposium in Lausanne are acknowledged. A discussion with Gil Wiechman in an early stage of this work is also acknowledged. The author wishes to thank Boaz Shuval, Moshe Twitto, and Oren Zeitlin for pointing out typos in a previous draft. Thanks are due to the anonymous reviewers for their feedback which contributed to the lucidity of the presentation.

REFERENCES

[1] A. Amraoui, A. Montanari and R. Urbanke, "How to find good finite-length codes: from art towards science," *European Transactions on Telecommunications*, vol. 18, no. 5, pp. 491–508, April 2007.

[2] A. Amraoui and R. Urbanke, LdpcOpt: Software for optimizing the degree distrbutions of LDPC code ensembles. [Online]. Available: http://lthcwww.epfl.ch/research/ldpcopt/index.php.

[3] M. Ardakani, B. Smith, W. Yu and F. R. Kschischang, "Complexity-optimized low-density parity-check codes," *Proceedings of the Forty-Third Annual Allerton Conference on Communication, Control and Computing*, pp. 45–54, Urbana-Champaign, Illinois, USA, September 2005.

[4] M. H. Azmi, J. Yuan, J. Ning and H. Q. Huynh, "Improved bilayer LDPC codes using irregular check node degree distribution, " *Proceedings 2008 IEEE International Symposium on Information Theory*, pp. 141–145, Toronto, Canada, July 2008.

[5] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding," *Proceedings 1993 IEEE International Conference on Communications*, pp. 1064–1070, Geneva, Switzerland, May 1993.

[6] S. Boyd and L. Vanderberghe, *Convex Optimization*, Cambridge Press, 2004. [Online]. Available: http://www.stanford.edu/~boyd/cvxbook/.

[7] 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.

[8] S. Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Communications Letters*, vol. 5, no. 2, pp. 58–60, February 2001.

[9] D. J. Costello and G. D. Forney, "Channel coding: The road to channel capacity," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150–1177, June 2007.

[10] D. Divsalar, S. Dolinar and C. Jones, "Construction of protograph LDPC codes with linear minimum distance," *Proceedings 2006 IEEE International Symposium on Information Theory*, pp. 664–668, Seattle, Washington, USA, July 6–12, 2006.

[11] O. Etesami and A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. on Inofrmation Theory*, vol. 52, no. 5, pp. 2033–2051, May 2006.

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

[13] M. Franceschini, G. Ferrari, and R. Raheli, "Does the performance of LDPC codes depend on the channel?," *IEEE Trans. on Communications*, vol. 54, no. 12, pp. 2129-2132, December 2006.

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

[15] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, August 2008. [Online]. Available: `http://stanford.edu/~boyd/cvx`.

[16] J. Gross and J. Yellen, *Graph Theory and its Applications*, CRC Press Series on Discrete Mathematics and its Applications, 1999.

[17] P. Grover and A. K. Chaturvedi, "Upper bounds on the rate of LDPC codes for a class of finite-state Markov channels," *IEEE Trans. on Information Theory*, vol. 53, no. 2, pp. 794–804, February 2007.

[18] C. H. Hsu and A. Anastasopoulos, "Capacity-achieving codes with bounded graphical complexity on noisy channels," *Proceedings of the Forty-Third Annual Allerton Conference on Communications, Control and Computing*, pp. 1825–1834, Urbana-Champaign, IL, USA, September 2005.

[19] C. H. Hsu and A. Anastasopoulos, "Capacity-achieving LDPC codes through puncturing," *IEEE Trans. on Information Theory*, vol. 54, no. 10, pp. 4698–4706, October 2008.

[20] Y. Jiang, A. Ashikhmin, R. Koetter and A. C. Singer, "Extremal problems of information combining," *IEEE Trans. on Information Theory*, vol. 54, no. 1, pp. 51–71, January 2008.

[21] A. Khandekar and R. J. McEliece, "On the complexity of reliable communication on the erasure channel," *Proceedings 2001 IEEE International Symposium on Information Theory*, p. 1, Washington, D.C., USA, June 2001.

[22] F. R. Kschischang, B. J. Frey and H. A. Loeliger, "Factor graphs and the sum-product decoding," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 498–519, February 2001.

[23] L. Lan, L. Zeng, Y. Tai and L. Chen, "Construction of quasi-cyclic LDPC codes for the AWGN and binary erasure channels: A finite field approach," *IEEE Trans. on Information Theory*, vol. 53, no. 7, pp. 2429–2458, July 2007.

[24] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi and D. A. Spielman, "Efficient erasure-correcting codes," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 569–584, February 2001.

[25] D. J. C. MacKay and R. M. Neal, "Near Shannon limit perofrmance of low-density parity-check codes," *IEE Electronic Letters*, vol. 32, no. 18, pp. 1545–1646, August 1996.

[26] C. Measson, A. Montanari, T. Richardson and R. Urbanke, "The generalized area theorem and some of its consequences," submitted to *IEEE Trans. on Information Theory*, November 2005. [Online]. Available: `http://arxiv.org/abs/cs/0511039`.

[27] G. Miller and G. Cohen, "The rate of regular LDPC codes," *IEEE Transactions on Information Theory*, vol. 49, no. 11, pp. 2989–2992, November 2003.

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

[29] P. Oswald and A. Shokrollahi, "Capacity-achieving sequences for the erasure channel," *IEEE Trans. on Information Theory*, vol. 48, no. 12, pp. 3017–3028, December 2002.

[30] F. Peng, W. E. Ryan and R. D. Wesel, "Surrogate-channel design of universal LDPC codes," *IEEE Communications Letters*, vol. 10, no. 6, pp. 480–482, June 2006.

[31] 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.

[32] H. D. Pfister and I. Sason, "Accumulate-repeat-accumulate codes: Capacity-achieving ensembles of systematic codes for the erasure channel with bounded complexity," *IEEE Trans. on Information Theory*, vol. 53, no. 6, pp. 2088–2115, June 2007.

[33] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. on Information Theory*, vol. 40, pp. 1284–1292, July 1994.

[34] 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.

[35] 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.

[36] T. Richardson and R. Urbanke, "The renaissance of Gallager's low-density parity-check codes," *IEEE Communications Magazine*, vol. 41, no. 8, pp. 126–131, August 2003.

[37] T. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008. [Online]. Available: `http://lthcwww.epfl.ch/mct/index.php`.

[38] A. Sanaei, M. Ramezani and M. Ardakani, "On the design of universal LDPC codes," *Proceedings 2008 International Symposium on Information Theory*, pp. 802–806, Toronto, Canada, July 2008.

[39] I. Sason and S. Shamai, "Performance Analysis of Linear Codes under Maximum-Likelihood Decoding: A Tutorial," *Foundations and Trends in Communications and Information Theory*, vol. 3, no. 1–2, pp. 1–222, Now Publishers, Delft, the Netherlands, July 2006. [Online]. Available: `http://www.ee.technion.ac.il/people/sason/monograph_postprint.pdf`.

[40] 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.

[41] I. Sason and G. Wiechman, "On achievable rates and complexity of LDPC codes over parallel channels: Bounds and applications," *IEEE Trans. on Information Theory*, vol. 53, no. 2, pp. 580–598, February 2007.

[42] I. Sason and G. Wiechman, "Bounds on the number of iterations for turbo-like ensembles over the binary erasure channel," *Proceedings 2008 IEEE International Symposium on Information Theory*, pp. 1898–1902, Toronto, Ontario, Canada, July 2008.

[43] L. Sassatelli, W. Henkel and D. Declercq, "Check-irregular LDPC codes for unequal error protection under iterative decoding," *Proceedings of the 4th International Symposium on Turbo Codes and Related Topics*, Munich, Germany, April 2006.
[44] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, pp. 611–656, May 1959.
[45] C. Shannon, R. Gallager and E. Berlekamp, "Lower bounds to error probability for decoding on discrete memoryless channels," *Information and Control*, vol. 10, Part 1: pp. 65–103, and Part 2: pp. 522–552, February/May 1967.
[46] 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.
[47] J. Shi and R. D. Wesel, "A study on universal codes with finite block length," *IEEE Trans. on Information Theory*, vol. 53, no. 9, pp. 3066–3074, September 2007.
[48] A. Shokrollahi, "Capacity-achieving sequences," *IMA Volume in Mathematics and its Applications*, vol. 123, pp. 153–166, 2000.
[49] I. Sutskover, S. Shamai and J. Ziv, "Constrained information combining: Theory and applications for LDPC coded systems," *IEEE Trans. on Information Theory*, vol. 53, no. 5, pp. 1617–1643, May 2007.
[50] S. Tong, "Tangential-sphere bounds on the ensemble performance of ML decoded Gallager codes via their exact ensemble distance spectrum," *Proceedings 2008 IEEE International Symposium on Communications*, pp. 1150–1154, Beijing, China, May 2008.
[51] A. Valembois and M. Fossorier, "Sphere-packing bounds revisited for moderate block length," *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 2998–3014, December 2004.
[52] T. Wadayama, "Ensemble analysis on syndrome entropy of binary linear codes," *Proceedings 2006 IEEE International Symposium on Information Theory*, pp. 1559–1564, Seattle, Washington, USA, July 2006.
[53] G. Wiechman and I. Sason, "Parity-check density versus performance of binary linear block codes: New bounds and applications," *IEEE Trans. on Information Theory*, vol. 53, no. 2, pp. 550–579, February 2007.
[54] G. Wiechman and I. Sason, "An improved sphere-packing bound for finite-length codes over symmetric memoryless channels," *IEEE Trans. on Information Theory*, vol. 54, no. 5, pp. 1962–1990, May 2008.
[55] Special issue on *Codes on Graphs and Iterative Algorithms*, *IEEE Trans. on Information Theory*, vol. 47, no. 2, February 2001.
[56] Big-O notation. [Online]. Available: `http://www.nist.gov/dads/HTML/bigOnotation.html`.