ON UNIVERSAL LDPC CODES OVER MEMORYLESS SYMMETRIC CHANNELS

Igal Sason Boaz Shuval
Department of Electrical Engineering
Technion – Israel Institute of Technology
Haifa 32000, Israel

April 28, 2011

Abstract

A design of robust error-correcting codes that achieve reliable communication over various channels is of great theoretical and practical interest. Such codes are termed *universal*. This paper considers the universality of low-density parity-check (LDPC) code ensembles over families of memoryless binary-input output-symmetric (MBIOS) channels. Universality is considered both under belief-propagation (BP) and maximum-likelihood (ML) decoding.

For the BP decoding case, we derive a density-evolution-based analytical method for designing LDPC code ensembles that are universal over various families of MBIOS channels. We also derive a necessary condition for universality of LDPC code ensembles under BP decoding; this condition is used to provide bounds on the universally achievable fraction of capacity. These results enable us to provide conditions for reliable / unreliable communications under BP decoding that are based on the Bhattacharyya parameter of the channel. For the ML decoding case, we prove that properly selected regular LDPC code ensembles are universally capacity-achieving for the set of equi-capacity MBIOS channels and extend this result to punctured regular LDPC code ensembles.

Index Terms

Belief propagation (BP), Bhattacharyya parameter (B-parameter), density evolution (DE), linear programming (LP) bounds, low-density parity-check (LDPC) codes, maximum-likelihood (ML), memoryless binary-input output-symmetric (MBIOS) channels, stability.

I. INTRODUCTION

Low-density parity-check (LDPC) codes are linear block codes that can be represented by sparse parity-check matrices. This sparse structure enables to decode these codes by suboptimal iterative decoding algorithms. These low-complexity algorithms are remarkable in that they achieve rates close to capacity for properly designed LDPC code ensembles (see, e.g., [4], [16], [18], [24], and [25]).

The density evolution approach serves as a main tool for the asymptotic analysis of the performance of LDPC code ensembles under iterative message-passing decoding [25]. Using this approach, it is possible to numerically optimize LDPC code ensembles for specific memoryless binary-input output-symmetric (MBIOS) channels. The goal is to find degree distributions that asymptotically ensure convergence to error-free communications for a given channel model, and that are optimal in the sense of either achieving maximal rate for specific channel parameters, or exhibiting the best threshold for a specific chosen rate or other constraints on the degree distributions. Depending on the threshold parameter considered, the best threshold could be either a maximal value, as in the crossover probability of a binary symmetric channel, or a minimal value, as in the signal-to-noise ratio of a binary-input additive white Gaussian noise channel. Another consequence of density evolution is the stability condition which forms a necessary condition for an LDPC code ensemble to asymptotically achieve vanishing bit error probability

This research work was supported by the Israel Science Foundation (grant no. 1070/07). The material in this paper was presented in part at the 2010 International Symposium on Information Theory (ISIT '10), Austin, Texas, USA, June 2010. It was also presented in part at the 2010 IEEE 26th Convention of Electrical and Electronics Engineers, Eilat, Israel, November 2010. Igal Sason is the corresponding author (e-mail: sason@ee.technion.ac.il).

under iterative message-passing decoding for a given channel model. Density evolution is a powerful tool for numerical optimization of degree distributions, but it does not lend itself, in general, for the *analytical* design of degree distributions. An exception to this is the case of the binary erasure channel (BEC), where density evolution is greatly simplified to a single-dimensional recursive equation. Based on this, several explicit expressions of capacity-achieving sequences for the BEC have been derived (see, e.g., [22] and [35]). So far, no explicit expressions for capacity-achieving code ensembles under iterative decoding have been found for other MBIOS channel models. While ML decoding is impractical, explicit expressions for capacity-approaching codes for any MBIOS channel under ML decoding have been derived (see [11] and [30, Theorem 2.2]). The analysis under ML decoding relies on upper bounds on the decoding error probability based on the average weight distribution of the ensemble (see [20] and [29]).

It is of great interest, both practically and theoretically, to design a code that will operate reliably over a range of channels. Such robust codes are termed *universal*. There are many different notions of universality, and many approaches to the design of universal codes. An excellent survey on the matter can be found in [15], where the authors have focused on the problem of communicating reliably when there is channel uncertainty. The authors introduced several models of channel uncertainty, and discussed several universality strategies for these models, both in terms of encoder design and decoder design.

The subject of universal LDPC codes has been addressed in several recent studies. In this setting, the goal is to design an LDPC code ensemble that will perform well in terms of error probability over a family of channels, using a standard decoder for LDPC codes, such as a belief propagation decoder.

One approach is to find so-called extreme channels that can be used to predict the LDPC code ensemble's performance under iterative decoding. Khandekar [13] showed that a code's behavior on the BEC can be used to predict its behavior on other channels. In particular, he showed that if for a BEC, the bit erasure probability of an LDPC code ensemble converges to zero under iterative message-passing decoding then the bit error probability will also converge to zero on any other MBIOS channel with the same Bhattacharyya parameter (B-parameter). Among the family of equi-capacity MBIOS channels, the BEC exhibits the smallest B-parameter whereas the binary symmetric channel (BSC) exhibits the greatest B-parameter [3]. Based on this observation, it was suggested that it may be possible to design a code for an arbitrary MBIOS channel by designing it for a BEC with a matching B-parameter. Further evidence of the extremes of the BEC and BSC can be found, e.g., in a study by Sutskover et al. (see [37] and [38]), which is based on an information-combining approach ([14]) to predict the behavior of LDPC code ensembles over various channels; they showed that the behavior of an LDPC code ensemble over a BEC and BSC can be used to provide bounds on its behavior over other MBIOS channels under iterative message-passing decoding. These works all use bounds that stem from the reduction of the density evolution equation to a single parameter. Such bounds first appeared in [2], albeit not from a universality standpoint.

Another approach stems from several researchers having noticed that LDPC codes exhibit similar performance under iterative message-passing decoding over a set of channels with similar parameters. Numerical evidence that equi-capacity and equi-B-parameters exhibit similar thresholds is provided in [3], and thus it was conjectured that the performance of an LDPC code over one MBIOS channel can be approximated by its performance on a different MBIOS channel but with the same capacity or B-parameter. In [6], the authors also provide supporting numerical evidence that LDPC code ensembles behave similarly on equi-capacity MBIOS channels. In [23], the authors conjecture that it is possible to design good LDPC codes based on a so-called "surrogate" channel, such as the BEC, so that they will exhibit good performance over other channels. Recently, Sanaei et al. [28] designed numerically some universal LDPC code ensembles that achieve a high fraction of capacity for a set of equi-capacity MBIOS channels. Based on some practical experiments, they have also conjectured that an LDPC code ensemble designed for two equi-capacity MBIOS channels will also converge under iterative message-passing decoding over any convex combination of these two channels[1].

We briefly mention several other avenues of research regarding universal LDPC codes; these works present approaches that are quite remote from the approach of our work on universality, and are presented here for the sake of completeness. Duyck et al. [5] numerically optimized LDPC code ensembles to be universal over Rician fading multiple-access channels. Miyake and Maruyama [21] studied universal properties of fixed length LDPC codes under the minimum-entropy decoding scheme. Universal codes with finite block lengths were addressed in [34];

[1]I.e., a channel formed by using one of these channels with probability θ, $0 \leq \theta \leq 1$, and the other channel with probability $1-\theta$.

this paper was focused on the performance of such code ensembles, in terms of bounds on the probability of error (and error exponents), for a class of channels the authors call "periodic erasure channels." Factor-graph decoding over a family of channels related by some unknown parameters was considered in [41] where this work defined several factor-graph based decoding schemes over channels with unknown parameters, and applied to codes that are represented by factor graphs, not necessarily LDPC codes. Finally, Yedla, et al. [44] consider the problem of universal joint source-channel coding; in their setting, there are two correlated sources transmitting over two channels with unknown parameters, to be jointly decoded by a single receiver.

In this paper, we consider the universality of LDPC code ensembles under both iterative message-passing decoding and maximum-likelihood decoding over MBIOS channels. For the case of belief-propagation (BP) decoding, we use density evolution to derive some conditions for the universality of LDPC codes over various MBIOS channels. These results serve to formulate an approach for the analytical design of universal LDPC code ensembles. Furthermore, we show that for any code ensemble, one can classify channels as good or bad (in the sense of convergence under BP decoding) based on the value of the B-parameter of the channel. For the ML decoding case, we show that regular LDPC code ensembles can be made universal both with and without puncturing over equi-capacity MBIOS channels.

This paper is structured as follows: Section II provides some preliminary material and notation, Section III explores the universality of LDPC code ensembles under BP decoding, Section IV contains some universality results for LDPC code ensembles under ML decoding, and Section V concludes this work with a summary and some directions for future research.

II. Preliminaries

This section follows the notation in [26, Chapter 4], and briefly introduces some preliminaries on MBIOS channels that are relevant for the analysis in this paper.

Consider an MBIOS channel whose input and 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 probability density function (*pdf*) of the random variable L given that the channel input is $X = 0$ (to be referred to as the L-density function). This density function satisfies the symmetry property $a(x) = e^x\, a(-x)$ for every $x \in \mathbb{R}$ (see [26, Theorem 4.26]).

This paper relies on the following three functionals (various other functionals are presented in [26, Section 4.1]).

Proposition 1. [Capacity functional] Consider an MBIOS channel whose symmetric L-density function is denoted by a. The capacity of this channel in units of bits per channel use, $C \triangleq C(a)$, is given by

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

This proposition is proved in [26, p. 193].

Definition 1. [The Bhattacharyya functional] The Bhattacharyya parameter (B-parameter), $B \triangleq \mathcal{B}(a)$, which is associated with the symmetric L-density function a is given by

$$B = \int_{-\infty}^{\infty} a(x) e^{-\frac{x}{2}}\,\mathrm{d}x. \tag{2}$$

The following is a direct consequence of a proposition that was introduced in [1, Proposition 1] (for a proof see [1, Appendix A]); it relates the capacity with the B-parameter of an MBIOS channel.

Proposition 2. For every MBIOS channel, let a be the L-density of the LLR at the channel output for an equi-probable binary input, and let B and C designate the B-parameter and channel capacity, respectively. Then, the following inequality holds:

$$\log_2\left(\frac{2}{1+B}\right) \leq C \leq \sqrt{1 - B^2}. \tag{3}$$

Proposition 2 implies that for a perfect MBIOS channel, whose capacity, C, approaches 1 bit per channel use, the corresponding B-parameter tends to zero. On the other hand, for a very noisy channel, whose capacity is close to zero, we have that the B-parameter tends to 1. This is consistent with the interpretation that the B-parameter forms an upper bound on the error probability under ML decoding when the channel is used only once to transmit a zero or a one.

The following proposition was introduced in [33, Lemma 8] (for a proof see [33, Appendix IV]), and it provides another property that relates the channel capacity and the B-parameter of an MBIOS channel. This proposition improves upon the lower bound in Proposition 2.

Proposition 3. For every MBIOS channel, the sum of its channel capacity and its B-parameter is greater than or equal to 1, i.e.,

$$B + C \geq 1$$

and equality is achieved for a BEC.

Note that Proposition 3 implies that among all equi-capacity MBIOS channels, the BEC possesses the minimal B-parameter.

Remark 1. From the lower bound of Proposition 2, it is implied that $C + B \geq 1 + (B - \log_2(1+B))$. It can easily be verified that $f(B) \triangleq 1 + (B - \log_2(1+B)) \leq 1$ for $B \in [0,1]$, with equality only at the end points, i.e., $B = 0$ or $B = 1$. To see this, we first note that indeed $f(0) = f(1) = 1$. The derivative of f is $f'(B) = 1 - ((1+B)\ln 2)^{-1}$, which has only one zero, at $B = -1 + 1/\ln 2 \approx 0.4427$. This is easily determined to be a minimum point of f, implying that $f(B) \leq 1$ for $B \in [0,1]$. On the other hand, Proposition 3 states that $B + C \geq 1$, thereby improving the lower bound in Proposition 2.

Following standard notation, the degree distributions of the LDPC codes under consideration from the edge perspective are denoted by $\lambda(x) = \sum_k \lambda_k x^{k-1}$ and $\rho(x) = \sum_k \rho_k x^{k-1}$, where λ_k and ρ_k designate the fraction of edges emanating from variable and parity-check nodes of degree k, respectively.

The analysis in this paper relies partially on the *stability condition* for LDPC code ensembles under BP decoding. This condition applies to the asymptotic case where we let the block length tend to infinity, and it forms a necessary condition for successful BP decoding in the sense that it requires that the fixed point of zero bit error rate be stable. Consider an LDPC code ensemble with a 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 assumes the form (see [26, Theorem 4.125])

$$\mathcal{B}(a)\lambda'(0)\rho'(1) < 1. \tag{4}$$

The reader is referred to [26, Section 4.9] for a proof.

Definition 2. [The error probability functional] The bit error probability that is associated with a symmetric L-density function a is given by

$$\begin{aligned}\mathcal{E}(a) &= \int_{-\infty}^{0^-} a(x)\,\mathrm{d}x + \frac{1}{2}\int_{0^-}^{0^+} a(x)\,\mathrm{d}x \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} a(x)\, e^{-(|\frac{x}{2}|+\frac{x}{2})}\,\mathrm{d}x.\end{aligned}$$

The following inequalities relate the Bhattacharyya and error probability functionals. Based on [26, Lemma 4.64], the following inequality holds for an arbitrary symmetric L-density a

$$2\mathcal{E}(a) \leq \mathcal{B}(a) \leq 2\sqrt{\mathcal{E}(a)\big(1-\mathcal{E}(a)\big)}. \tag{5}$$

Note that the lower and upper bounds on the B-parameter, as given in (5), are satisfied with equality for a BEC and BSC, respectively.

Convolutions of densities in the so-called L-domain and G-domain[2] are presented in [26, p. 181], and are denoted by $\circledast$ and $\boxtimes$, respectively. Using the density evolution approach for the asymptotic analysis of LDPC code ensembles over MBIOS channels, where we let the block length tend to infinity, the $\circledast$ convolution describes how the distribution of the (statistically independent) messages changes at the variable node under BP decoding at every single iteration, whereas the $\boxtimes$ convolution describes the change of this distribution at the parity-check node side.

III. Universality under Belief Propagation Decoding

A. Universal Achievability Results

In the following, we consider the suitability of LDPC code ensembles to operate reliably over a set of MBIOS channels under BP decoding. We rely here on the density evolution approach, and our goal is to construct LDPC code ensembles which achieve vanishing bit error probability, in the asymptotic case where the block length tends to infinity, uniformly over a set of MBIOS channels.

To this end, let us consider first an arbitrary MBIOS channel, and let a_0 denote the *pdf* of the LLR at the channel output given that the channel input is zero. Let λ and ρ designate the degree distributions of the variable and parity-checks, respectively, from the edge perspective. Based on density evolution, the densities at every iteration of the BP decoder satisfy the recursive equation

$$a_l = a_0 \circledast \lambda\Big(\Gamma^{-1}\Big(\rho\big(\Gamma(a_{l-1})\big)\Big)\Big), \quad l = 1, 2, \ldots \tag{6}$$

where the mapping Γ and its inverse Γ^{-1} are introduced in [25, p. 627]. The densities a_l are symmetric functions for every $l \geq 0$, i.e., $a_l(x) = e^x a_l(-x)$ for all $x \in \mathbb{R}$. Let $x_l = \mathcal{B}(a_l)$ for $l \geq 0$ where $\mathcal{B}(a)$ designates the B-parameter that is associated with the L-density a. Based on the proof of sufficiency in the stability condition (see [26, p. 234]), it follows that

$$x_l \leq \mathcal{B}(a_0)\, \lambda\big(1 - \rho(1 - x_{l-1})\big), \quad l = 1, 2, \ldots \tag{7}$$

where this inequality is proved in [9, Theorem 4.2] and [12, Theorem 2]. From (5), an LDPC code ensemble obtains asymptotically vanishing bit error probability as the number of iterations grows if and only if $\lim_{l\to\infty} x_l = 0$.

Let us now consider an arbitrary set of MBIOS channels, and let $\mathcal{A}$ designate the corresponding set of its L-densities. Suppose that one wishes to design an LDPC code ensemble with degree distributions (λ, ρ) in order to asymptotically achieve vanishing bit error probability under BP decoding for every channel in this set. Let us designate by B the maximal B-parameter over the MBIOS channels of the considered set, i.e.,

$$B \triangleq \max_{a \in \mathcal{A}} \mathcal{B}(a). \tag{8}$$

Let us consider the recursive equation

$$y_l = B\, \lambda\big(1 - \rho(1 - y_{l-1})\big), \quad l = 1, 2, \ldots \tag{9}$$

with the initial value $y_0 = B$. This recursive equation refers to the density evolution of a BEC whose erasure probability is equal to B. By comparing (7) and (9), it is straightforward to show (e.g., by induction) that $0 \leq x_l \leq y_l$ for every $l \geq 0$ and $a \in \mathcal{A}$. If the pair of degree distributions (λ, ρ) is selected in a way where $\lim_{l\to\infty} y_l = 0$, then we get that $\lim_{l\to\infty} x_l = 0$ in (7) for *every MBIOS channel* from the set $\mathcal{A}$. Hence, the universality of the LDPC code ensemble whose degree distribution is (λ, ρ) follows with respect to the considered set of channels.

One can thus rely on (9) to construct a sequence of LDPC code ensembles which achieves vanishing bit error probability, under BP decoding, for all the MBIOS channels of the considered set. In particular, to this end one can use the well-known explicit constructions of capacity-achieving sequences of LDPC code ensembles for the BEC (see, e.g., [22] and references therein). By this approach, the asymptotic design rate of this capacity-achieving sequence of LDPC code ensembles is equal to the capacity of the BEC, i.e.,

$$R_{\mathrm{d}} = 1 - B, \tag{10}$$

where B is given in (8). We study the following particular cases of this approach.

[2]As mentioned at the beginning of this section, an L-density is the pdf of the LLR $l(Y)$ given that the channel input is $X = 0$. A G-density is the result of the transformation $l(Y) \to (\operatorname{sgn} l(Y), \log\coth(|l(Y)|/2))$. The L- and G-domains are, respectively, the domains of the L and G densities.

1) Universal LDPC Code Ensembles for Equi-Capacity MBIOS Channels: Among all MBIOS channels which exhibit a given capacity C, the B-parameter that is associated with the L-densities of this set of channels attains its maximal and minimal values for the BSC and BEC, respectively (this follows readily from (5)). The B-parameter of a BSC whose crossover probability is p is equal to $\sqrt{4p(1-p)}$, and the capacity of this channel is equal to $C = 1 - h_2(p)$. By referring to the set of all equi-capacity MBIOS channels, one therefore gets from (8) that

$$B = \sqrt{4h_2^{-1}(1-C)\left(1-h_2^{-1}(1-C)\right)} \tag{11}$$

where h_2^{-1} designates the inverse of the binary entropy function on base 2. From (10), the asymptotic design rate of the corresponding sequence of LDPC code ensembles is equal to $R_{\mathrm{d}} = 1 - B$. As a consequence of Proposition 3, it follows that indeed $R_{\mathrm{d}} \leq C$, which is necessary for reliable communication. The fraction of the channel capacity that is achievable by this approach,

$$\mu_1(C) \triangleq \frac{R_{\mathrm{d}}}{C},$$

is therefore equal to

$$\mu_1(C) = \frac{1 - \sqrt{4h_2^{-1}(1-C)\left(1-h_2^{-1}(1-C)\right)}}{C}. \tag{12}$$

Lemma 1. The function μ_1 is monotonically increasing over the interval $(0, 1]$, and

$$\lim_{C \to 0} \mu_1(C) = \ln 2 \approx 69.3\%, \qquad \lim_{C \to 1} \mu_1(C) = 1.$$

Proof: See Appendix A. ■

This implies that as the value of the capacity is increased, a larger fraction of the channel capacity is achievable uniformly for the entire considered set of equi-capacity MBIOS channels, and the two extremes are 69.3% and 100% when the capacity varies between zero and 1 bits per channel use. For a value of the channel capacity which approaches 1, the channels are almost noiseless, so almost no coding is required. Hence, the uniform attainment of nearly 100% of the capacity for the entire set of channels is well expected. However, this convergence of the achievable fraction of capacity is rather slow as we let the code rate tend to 1 (as is evidenced in Fig. 1). To see this, note that if C is close to 1

$$\mu_1(C) \approx \frac{1 - 2\sqrt{h_2^{-1}(1-C)}}{C}$$

which tends to 1 quite slowly (e.g., for $C = 0.95$ bits per channel use, this approximation is equal to 0.895 which indeed coincides with Fig. 1).

The above analysis implies that at least 69.3% of the capacity of any MBIOS channel can be achieved by designing a capacity-achieving sequence of LDPC code ensembles for a BEC; the erasure probability of this BEC is set to be equal to the B-parameter of a BSC whose capacity matches our channel.

This presents an analytical approach for the design of universal LDPC code ensembles for equi-capacity MBIOS channels where a provable (non-vanishing) fraction of capacity is universally achieved, and the value of this fraction gets larger as the value of capacity is increased. We note however that numerical optimization via density evolution enables to design universal LDPC code ensembles in [28] achieving a significantly larger fraction of the channel capacity, though the considered approach here is purely analytical, and it is not subject to numerical optimizations.

2) Universal LDPC Code Ensembles for BEC and BIAWGNC with the Same Capacity: We consider here an achievable fraction of capacity when one wishes to design an LDPC code ensemble which achieves asymptotically vanishing bit error probability under BP decoding for both the BEC and the binary-input AWGN channel (BI-AWGNC) with the same capacity. Since among all equi-capacity MBIOS channels, the BEC possesses the minimal B-parameter (see Proposition 3), then the parameter B in (8) corresponds to the B-parameter of the BIAWGNC. The conversion from the channel capacity to the B-parameter for this channel is done numerically by first calculating the noise variance σ^2 via the following expression for its capacity (see [26, p. 194]):

$$C = 1 + \frac{1}{\ln 2}\left[\left(\frac{2}{\sigma^2} - 1\right) Q\Big(\frac{1}{\sigma}\Big) - \sqrt{\frac{2}{\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}} + \sum_{i=1}^{\infty} \frac{(-1)^i}{i(i+1)}\, e^{\frac{2i(i+1)}{\sigma^2}} Q\Big(\frac{1+2i}{\sigma}\Big)\right]$$

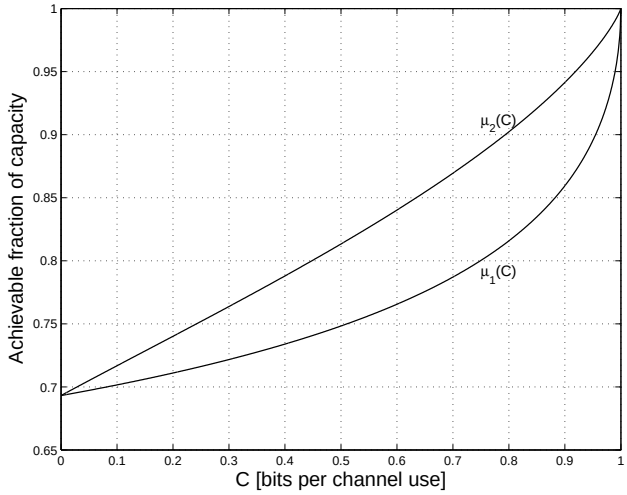

Fig. 1. Universal achievable fraction of capacity under BP decoding for two sets of MBIOS channels which exhibit a given capacity (see Theorem 1). The values of μ_1 in (12) and μ_2 in (13) correspond, respectively, to the entire set of equi-capacity MBIOS channels, and the subset of a BEC and BIAWGNC with capacity C bits per channel use.

(based on (1)), and then substituting the value of σ^2 to obtain the B-parameter $B = e^{-\frac{1}{2\sigma^2}}$ (based on (2)). From (10), the asymptotic achievable fraction of the capacity is equal to

$$\mu_2(C) = \frac{1-B}{C}. \tag{13}$$

Since the universality in this example applies to a subset of the equi-capacity MBIOS channels, the inequality $\mu_2(C) \geq \mu_1(C)$ is expected to hold for $0 \leq C \leq 1$. Our results so far are summarized in the following theorem. To this end, we denote by BEC(ε) the binary erasure channel whose erasure probability is ε:

Theorem 1. [Universality of LDPC Codes under BP Decoding for Equi-Capacity MBIOS Channels] Consider a set $\mathcal{A}$ of MBIOS channels that exhibit a given capacity C, and let B denote the maximal B-parameter over this set (see (8)). Let $\{(n, \lambda, \rho)\}$ form a capacity-achieving sequence of LDPC code ensembles for BEC(B), achieving vanishing bit erasure probability under BP decoding. Then, this sequence universally achieves vanishing bit error probability under BP decoding for the entire set $\mathcal{A}$, and the design rate of this sequence forms a fraction that is at least $\frac{1-B}{C}$ of the channel capacity. As a consequence, the following results hold:

- For the entire set of equi-capacity MBIOS channels, the universally achievable design rate forms at least a fraction $\mu_1(C)$ of capacity (see (12)). Moreover, μ_1 forms a monotonic increasing function of the capacity C (see Fig. 1), getting the extreme values $\ln 2 \approx 69.3\%$ and 100% at the endpoints where $C \to 0$ or $C \to 1$, respectively.
- For some sub-classes of equi-capacity MBIOS channels, the results for the universally achievable design rate significantly improve (see, e.g., (13) and μ_2 in Fig. 1).

Fig. 1 compares the achievable fractions of capacity, μ_1 and μ_2 as a function of the channel capacity.

B. Universal Lower Bound on the Achievable Gap to Capacity

The stability condition $\mathcal{B}(a)\lambda'(0)\rho'(1) < 1$ forms a necessary condition for asymptotically achieving vanishing bit error probability under BP decoding when the transmission takes place over an MBIOS channel.

We wish to find an upper bound on the achievable design rate of universal LDPC code ensembles over a set $\mathcal{A}$ of MBIOS channels, or alternatively, a universal lower bound on the achievable gap (in rate) to capacity. From the stability condition in (4) and also from (8), the inequality

$$B\lambda'(0)\rho'(1) \leq 1 \tag{14}$$

forms a necessary condition for achieving this goal universally over the set $\mathcal{A}$.

We consider in the following *right-regular* LDPC code ensembles where d_{c} designates the degree of the parity-check nodes (i.e., $\rho(x) = x^{d_{\mathrm{c}}-1}$).

Following the notation in [26, p. 181], let $a \circledast b$ denote the density which is the result of transforming both a and b from the L-domain to the G-domain, then performing the convolution in the G-domain, and then transforming it back to the L-domain. As mentioned earlier, under BP decoding, the operator $\circledast$ describes the change of the distributions at the check node side.

In the following, we introduce an additional necessary condition for universally achieving vanishing bit error probability under BP decoding with respect to a set $\mathcal{A}$ of MBIOS channels.

Theorem 2. [A Necessary Condition for Universality of LDPC Code Ensembles under BP Decoding] Let $\{(n, \lambda, \rho)\}$ be a right-regular sequence of LDPC code ensembles, universally achieving vanishing bit error probability under BP decoding for a set of MBIOS channels $\mathcal{A}$. Then, the following condition holds

$$B\lambda\big(\sqrt{1-\rho(1-x^2)}\big) < x, \quad \forall\, x \in (0, B] \tag{15}$$

where B designates the maximal B-parameter over the set $\mathcal{A}$.

Proof: For the derivation of this condition, we rely on the following inequality:

Lemma 2. Let $a^{\circledast k} \triangleq a \circledast a \circledast \cdots \circledast a$ denote the operator where a is convolved by itself $k-1$ times (i.e., a appears k times on the right-hand side of this equality), where the convolution here is in the G-domain (see [26, p. 181]). Then, for a symmetric L-density a with $\mathcal{B}(a) = \beta_a$

$$\mathcal{B}(a^{\circledast k}) \geq \sqrt{1-(1-\beta_a^2)^k} \tag{16}$$

for any integer $k \geq 2$.

Proof: Let a and b denote two symmetric L-densities with $\mathcal{B}(a) = \beta_a$ and $\mathcal{B}(b) = \beta_b$, then from [26, Problem 4.62]

$$\sqrt{\beta_a^2 + \beta_b^2 - \beta_a^2\beta_b^2} \leq \mathcal{B}(a \circledast b) \leq \beta_a + \beta_b - \beta_a\beta_b \tag{17}$$

where the upper and lower bounds are achieved with equality if a and b are from the family BEC or BSC, respectively. By setting $a = b$, we get the inequality in (16) for $k = 2$. The proof for a general $k \geq 2$ is completed by mathematical induction. Let us assume that (16) holds for a certain $k \geq 2$, then from (17)

$$\begin{aligned}
\mathcal{B}(a^{\circledast k+1}) &= \mathcal{B}(a^{\circledast k} \circledast a) \\
&\geq \sqrt{\mathcal{B}(a^{\circledast k})^2 + \mathcal{B}(a)^2 - \mathcal{B}(a^{\circledast k})^2\,\mathcal{B}(a)^2} \\
&= \sqrt{1-(1-\mathcal{B}(a^{\circledast k})^2)(1-\mathcal{B}(a)^2)} \\
&= \sqrt{1-(1-\mathcal{B}(a^{\circledast k})^2)(1-\beta_a^2)} \\
&\geq \sqrt{1-(1-\beta_a^2)^{k+1}}
\end{aligned}$$

which then implies that (16) also holds for $k+1$. ■

Corollary 1. For a *right-regular* LDPC code ensemble

$$\mathcal{B}\Big(\Gamma^{-1}\big(\rho(\Gamma(a))\big)\Big) \geq \sqrt{1-\rho(1-\mathcal{B}(a)^2)}. \tag{18}$$

Hence, by defining $x_l \triangleq \mathcal{B}(a_l)$ for all l in the density evolution equation in (6), we get the following chain of

equalities and inequalities

$$\begin{aligned} x_l &= \mathcal{B}(a_l) \\ &\overset{\text{(a)}}{=} \mathcal{B}(a_0)\,\mathcal{B}\biggl(\lambda\Bigl(\Gamma^{-1}\bigl(\rho(\Gamma(a_{l-1}))\bigr)\Bigr)\biggr) \\ &\overset{\text{(b)}}{=} \mathcal{B}(a_0)\,\lambda\biggl(\mathcal{B}\Bigl(\Gamma^{-1}\bigl(\rho(\Gamma(a_{l-1}))\bigr)\Bigr)\biggr) \\ &\overset{\text{(c)}}{\geq} \mathcal{B}(a_0)\,\lambda\Bigl(\sqrt{1-\rho(1-\mathcal{B}(a_{l-1})^2)}\,\Bigr) \\ &= \mathcal{B}(a_0)\,\lambda\Bigl(\sqrt{1-\rho(1-x_{l-1}^2)}\,\Bigr) \end{aligned} \tag{19}$$

where equality (a) follows from the recursive density evolution equation in (6) and since for two symmetric L-densities a and b

$$\mathcal{B}(a \circledast b) = \mathcal{B}(a)\,\mathcal{B}(b) \tag{20}$$

equality (b) follows since the linearity of the convolution operator and the last equality yield that

$$\begin{aligned} \mathcal{B}(\lambda(a)) &= \mathcal{B}\left(\sum_i \lambda_i a^{\circledast(i-1)}\right) \\ &= \sum_i \lambda_i \mathcal{B}\left(a^{\circledast(i-1)}\right) \\ &= \lambda\left(\mathcal{B}(a)\right) \end{aligned} \tag{21}$$

and inequality (c) follows from (18). By definition, the initial value x_0 is equal to the B-parameter of the symmetric L-density of the MBIOS channel. From (19), it follows that if the sequence $\{x_l\}$ tends asymptotically to zero, then the sequence

$$z_l = \mathcal{B}(a_0)\,\lambda\left(\sqrt{1-\rho(1-z_{l-1}^2)}\right), \quad l = 1, 2, \ldots \tag{22}$$

with the initial value $z_0 = \mathcal{B}(a_0)$, should also tend to zero. Note that the sequence $\{z_l\}$ is not the same as $\{x_l\}$: the sequence $\{x_l\}$, as shown in (19), is *greater than or equal to* the right-hand-side of (22). Further note that from (5), the convergence of the sequence $\{x_l\}$ forms a necessary and sufficient condition for achieving vanishing bit error probability as we let the number of iterations grow (recall that by the density evolution approach, we first let the block length tend to infinity, so that the tree assumption holds with probability 1 for any fixed number of iterations, and then we let the number of iterations grow).

Consider a sequence of right-regular LDPC code ensembles which universally achieves vanishing bit error probability under BP decoding over a set $\mathcal{A}$ of MBIOS channels. Let B be the maximal B-parameter over the entire set $\mathcal{A}$ (see (8)), then we obtain from (22) that the sequence defined recursively by

$$z_l = B\,\lambda\left(\sqrt{1-\rho(1-z_{l-1}^2)}\right), \quad l = 1, 2, \ldots \tag{23}$$

with the initial value $z_0 = B$ tends asymptotically to zero. Therefore, the satisfiability of the condition in (15) forms a necessary condition for universality. This completes the proof of Theorem 2. ■

For an extension of condition (15) for general LDPC code ensembles (not necessarily right-regular), see Appendix B.

In order to relate the condition in Theorem 2 to the stability condition, we calculate the derivative of the left-hand side of (15)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left\{B\,\lambda\left(\sqrt{1-\rho(1-x^2)}\right)\right\} = B\,\lambda'\left(\sqrt{1-\rho(1-x^2)}\right)\,x\left(1-\rho(1-x^2)\right)^{-\frac{1}{2}}\rho'(1-x^2)$$

and then require that this derivative be strictly less than 1 at the fixed point $x = 0$. Since d_{c} designates the fixed right degree of the right-regular LDPC code ensemble,

$$\begin{aligned}&\lim_{x\to 0} \frac{x}{\sqrt{1-\rho(1-x^2)}}\\ &= \lim_{x\to 0} \frac{x}{\sqrt{1-(1-x^2)^{d_{\mathrm{c}}-1}}}\\ &= \frac{1}{\sqrt{d_{\mathrm{c}}-1}}\end{aligned}$$

and therefore one gets the condition

$$\frac{B\lambda'(0)\rho'(1)}{\sqrt{d_{\mathrm{c}}-1}} < 1. \tag{24}$$

Interestingly, this coincides with the stability condition (4) up to a scaling factor that is equal to the reciprocal of the square root of $d_{\mathrm{c}}-1$; this scaling factor in (24) yields a weaker condition as compared to the stability condition. However, the condition in (15) provides a constraint on the interval $(0, B]$, and not just at a neighborhood of the fixed point at zero.

Let $d_{\mathrm{v}}^{\max}$ designate the maximal degree of the variable nodes. Since the design rate of a right-regular LDPC code ensemble is equal to

$$R_{\mathrm{d}} = 1 - \frac{1}{d_{\mathrm{c}} \sum_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i}} \tag{25}$$

then the maximization of R_{d} is equivalent to maximizing $\sum_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i}$.

Suppose that it is required to universally achieve vanishing bit error probability under BP decoding as the block length tends to infinity over a set $\mathcal{A}$ of equi-capacity MBIOS channels with capacity C. This requirement also implies that the bit error probability under MAP decoding vanishes. Thus, by combining [33, Eqs. (43), (44), and (53)], it follows that the design rate satisfies the inequality

$$0 \leq R_{\mathrm{d}} \leq 1 - \frac{1-C}{h_2\left(\frac{1-C^{\frac{d_{\mathrm{c}}}{2}}}{2}\right)} \tag{26}$$

and therefore, as the parity-check degree (d_{c}) is decreased, then R_{d} is more bounded away from capacity. Combining (25) and (26) gives that

$$\frac{1}{d_{\mathrm{c}}} \leq \sum_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i} \leq \frac{1}{(1-C)\, d_{\mathrm{c}}} \cdot h_2\left(\frac{1-C^{\frac{d_{\mathrm{c}}}{2}}}{2}\right). \tag{27}$$

By a maximization of $\sum\limits_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i}$ subject to

1) the necessary condition for vanishing bit error probability in Theorem 2,
2) the satisfiability of the stability condition for all the MBIOS channels in the set $\mathcal{A}$ (see (14)),
3) the inequality constraints in (27) that follow from the information-theoretic bounds in [33],

one obtains a linear programming (LP) universal upper bound on the achievable rate of LDPC code ensembles over the set $\mathcal{A}$ of equi-capacity MBIOS channels with capacity C under BP decoding. This gives the following LP bound where, practically, the values of $x \in (0, B]$ in the first inequality constraint are quantized uniformly over this interval in order to get a finite number of inequality constraints in the LP problem (to be referred to as the

TABLE I
LOWER BOUND ON THE UNIVERSAL ACHIEVABLE GAP TO CAPACITY ($\varepsilon \triangleq 1 - \frac{R_{\mathrm{d}}}{C}$) FOR EQUI-CAPACITY MBIOS CHANNELS UNDER BP DECODING; THE DEGREE OF THE PARITY-CHECK NODES IS FIXED (d_{c}), AND THE MAXIMAL DEGREE OF THE VARIABLE NODES IS SET TO $d_{\mathrm{v}}^{\max} = 200$. THESE NUMERICAL RESULTS REFER TO THE LP1 BOUND.

Capacity	Set of all Equi-Capacity Channels			BEC + BIAWGNC			BEC		
(C)	$d_{\mathrm{c}} = 8$	$d_{\mathrm{c}} = 10$	$d_{\mathrm{c}} = 12$	$d_{\mathrm{c}} = 8$	$d_{\mathrm{c}} = 10$	$d_{\mathrm{c}} = 12$	$d_{\mathrm{c}} = 8$	$d_{\mathrm{c}} = 10$	$d_{\mathrm{c}} = 12$
$\frac{1}{2}$	$2.83 \cdot 10^{-3}$	$7.05 \cdot 10^{-4}$	$1.76 \cdot 10^{-4}$	$2.83 \cdot 10^{-3}$	$7.05 \cdot 10^{-4}$	$1.76 \cdot 10^{-4}$	$2.83 \cdot 10^{-3}$	$7.05 \cdot 10^{-4}$	$1.76 \cdot 10^{-4}$
$\frac{3}{4}$	$9.09 \cdot 10^{-2}$	$1.79 \cdot 10^{-2}$	$7.84 \cdot 10^{-3}$	$7.90 \cdot 10^{-2}$	$1.43 \cdot 10^{-2}$	$7.84 \cdot 10^{-3}$	$5.56 \cdot 10^{-2}$	$1.43 \cdot 10^{-2}$	$7.84 \cdot 10^{-3}$
$\frac{9}{10}$	$2.06 \cdot 10^{-1}$	$1.57 \cdot 10^{-1}$	$1.20 \cdot 10^{-1}$	$1.73 \cdot 10^{-1}$	$1.33 \cdot 10^{-1}$	$1.03 \cdot 10^{-1}$	$1.67 \cdot 10^{-1}$	$1.11 \cdot 10^{-1}$	$7.99 \cdot 10^{-2}$

'LP1 bound'):

$$\begin{aligned}
&\text{maximize} \quad \sum_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i} \\
&\text{subject to} \\
&\begin{cases}
B\lambda\big(\sqrt{1-\rho(1-x^2)}\big) < x, \quad \forall\, x \in (0, B] \\
B\lambda_2 \rho'(1) \leq 1 \\
\sum_{i=2}^{\infty} \lambda_i = 1 \\
\lambda_i \geq 0, \quad i = 2, 3, \ldots \\
\frac{1}{d_{\mathrm{c}}} \leq \sum_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i} \leq \frac{1}{(1-C)\, d_{\mathrm{c}}} \cdot h_2\left(\frac{1-C^{\frac{d_{\mathrm{c}}}{2}}}{2}\right)
\end{cases}
\end{aligned}$$

Due to (25), LP1 also defines an upper bound on the design rate. This upper bound can also be translated into a universal lower bound on the achievable gap to capacity, $\varepsilon = 1 - R_{\mathrm{d}}/C$.

This LP problem is solved numerically with the aid of the CVX Matlab-based modelling system for convex optimization (see [8]). Numerical results for the lower bound on the achievable gap to capacity are provided in Table I for the cases where $\rho(x) = x^7, x^9$, and x^{11} (i.e., the parity-check degree is fixed to 8, 10, and 12, respectively), and the maximal degree of the variable nodes is set to $d_{\mathrm{v}}^{\max} = 200$.

In order to possibly improve the bound, let us consider the particular case where the set $\mathcal{A}$ forms a set of equi-capacity MBIOS channels that also includes the BEC. However, in the following case, the LDPC code ensembles are not restricted to be right-regular. For a BEC, the condition for vanishing bit erasure probability under BP decoding assumes the form

$$(1-C)\lambda\big(1-\rho(1-x)\big) < x, \quad \forall\, 0 < x \leq 1 - C. \tag{28}$$

This condition is used instead of the necessary condition in (15)[3]. Since in this LP the LDPC code ensemble is not assumed to be right-regular, the condition (27), which is a result of combining [33, Eqs. (43), (44), and (53)] and (25), assumes the form

$$\frac{1}{a_{\mathrm{R}}} \leq \sum_{i=2}^{d_{\mathrm{v}}^{\max}} \frac{\lambda_i}{i} \leq \frac{1}{(1-C)\, a_{\mathrm{R}}} \cdot h_2\left(\frac{1-C^{\frac{a_{\mathrm{R}}}{2}}}{2}\right),$$

where a_{R} is the average right degree of the LDPC code ensemble.

In this particular case where the set $\mathcal{A}$ includes the BEC, one gets the following LP problem (to be referred to

[3]It was verified numerically that adding condition (15) to the LP does not change the result. Thus, this condition is conjectured to be redundant in light of (28).

TABLE II
LOWER BOUND ON THE UNIVERSAL ACHIEVABLE GAP TO CAPACITY ($\varepsilon \triangleq 1 - \frac{R_{\rm d}}{C}$) FOR EQUI-CAPACITY MBIOS CHANNELS UNDER BP DECODING; THE DEGREE OF THE PARITY-CHECK NODES IS FIXED ($d_{\rm c}$), AND THE MAXIMAL DEGREE OF THE VARIABLE NODES IS SET TO $d_{\rm v}^{\max} = 200$. THESE NUMERICAL RESULTS REFER TO THE LP2 BOUND.

Capacity	Set of all Equi-Capacity Channels			BEC + BIAWGNC			BEC		
(C)	$d_{\rm c}=8$	$d_{\rm c}=10$	$d_{\rm c}=12$	$d_{\rm c}=8$	$d_{\rm c}=10$	$d_{\rm c}=12$	$d_{\rm c}=8$	$d_{\rm c}=10$	$d_{\rm c}=12$
$\frac{1}{2}$	$1.50\cdot10^{-2}$	$9.01\cdot10^{-3}$	$1.94\cdot10^{-2}$	$1.25\cdot10^{-2}$	$6.76\cdot10^{-3}$	$1.73\cdot10^{-2}$	$7.34\cdot10^{-3}$	$1.79\cdot10^{-3}$	$1.22\cdot10^{-2}$
$\frac{3}{4}$	$9.09\cdot10^{-2}$	$4.24\cdot10^{-2}$	$2.75\cdot10^{-2}$	$7.90\cdot10^{-2}$	$3.99\cdot10^{-2}$	$2.42\cdot10^{-2}$	$6.59\cdot10^{-2}$	$3.56\cdot10^{-2}$	$1.93\cdot10^{-2}$
$\frac{9}{10}$	$2.06\cdot10^{-1}$	$1.57\cdot10^{-1}$	$1.20\cdot10^{-1}$	$1.73\cdot10^{-1}$	$1.33\cdot10^{-1}$	$1.03\cdot10^{-1}$	$1.67\cdot10^{-1}$	$1.11\cdot10^{-1}$	$7.99\cdot10^{-2}$

as the 'LP2 bound'):

$$\begin{aligned}
&\text{maximize} \quad \sum_{i=2}^{d_{\rm v}^{\max}} \frac{\lambda_i}{i} \\
&\text{subject to} \\
&\begin{cases}
(1-C)\lambda\big(1-\rho(1-x)\big) < x, \quad \forall\, 0 < x \le 1-C \\
B\lambda_2\rho'(1) \le 1 \\
\sum_{i=2}^{\infty} \lambda_i = 1 \\
\lambda_i \ge 0, \quad i = 2,3,\ldots \\
\frac{1}{a_{\rm R}} \le \sum_{i=2}^{d_{\rm v}^{\max}} \frac{\lambda_i}{i} \le \frac{1}{(1-C)\,a_{\rm R}} \cdot h_2\left(\frac{1-C^{\frac{a_{\rm R}}{2}}}{2}\right)
\end{cases}
\end{aligned}$$

where the values of $x \in (0, 1-C]$ are quantized uniformly over this interval in order to get a finite number of inequality constraints in the LP problem; our implementation converts the first inequality constraint above to 1000 inequality constraints where x is equally spaced, and it gets the values $x_k = 0.001(1-C)k$ for $k = 1, \ldots, 1000$ (it was verified numerically that increasing the number of inequality constraints beyond one thousand, by a more refined uniform quantization of x over the interval $(0, 1-C]$, does not affect the numerical results of the LP2 bound). Numerical results for the lower bound on the achievable gap to capacity are provided in Table II for the same setting as in Table I[4].

By comparing Tables I and II, the values of the LP1 and LP2 bounds coincide for large values of the capacity C, whereas the LP2 bound shows an improved (larger) lower bound as compared to the LP1 bound for lower values of C. Note also that the two lower bounds become more significant (i.e., they become greater) as the value of capacity is increased. Let us mention that the possible improvement in the LP2 bound stems from the fact that it applies to a set of equi-capacity MBIOS channels that includes the BEC, whereas the LP1 bound applies to any set of equi-capacity MBIOS channels.

It was observed numerically that the LP2 bound on the achievable gap to capacity is sensitive the value of $d_{\rm v}^{\max}$, especially for large values of $d_{\rm c}$. For example, for $C = \frac{1}{2}$ and $d_{\rm c} = 12$, when $d_{\rm v}^{\max} = 200$, the LP2 lower bound is equal to $1.94 \cdot 10^{-2}$, but when $d_{\rm v}^{\max} = 500$ the LP2 lower bound becomes $7.38 \cdot 10^{-3}$.

C. Universal Conditions for Reliable Communications under Belief Propagation Decoding

We prove in this sub-section the following theorem and exemplify its use:

Theorem 3. [Universal Conditions on the B-parameter for Good/ Bad Communications under BP Decoding] Let $\{(n, \lambda, \rho)\}$ be a sequence of LDPC code ensembles whose block lengths tend to infinity. The following universal properties hold under BP decoding:

[4]Even though in the LP2 bound the LDPC code ensembles are not restricted to be right regular, for the purpose of comparing the results of the LP2 bound with those of the LP1 bound, we provide the numerical results for the same setting as for the LP1 bound.

- This sequence achieves vanishing bit error probability under BP decoding for *every* MBIOS channel whose B-parameter is less than
$$B_0(\lambda,\rho) \triangleq \inf_{x\in(0,1]} \frac{x}{\lambda\big(1-\rho(1-x)\big)}. \tag{29}$$
- For a *right-regular* sequence, it does not achieve reliable communications over *any* MBIOS channel whose B-parameter is greater than
$$B_1(\lambda,\rho) \triangleq \inf_{x\in(0,1]} \frac{x}{\lambda\big(\sqrt{1-\rho(1-x^2)}\big)}. \tag{30}$$
For every MBIOS channel whose B-parameter B satisfies $B > B_1(\lambda,\rho)$, BP decoding is not reliable in the sense that the left-to-right message error probability (i.e., the average probability of error for a message emanating from a variable node to a parity-check node) is greater than the positive value
$$\left(\frac{1}{2}\max\left\{x\in(0,1]:\ \frac{x}{\lambda\big(\sqrt{1-\rho(1-x^2)}\big)} \le B\right\}\right)^2 \tag{31}$$
irrespective of the number of iterations performed by the BP decoder.

Proof: We start by proving the first part of the theorem. Let $\{a_l\}$ be the sequence of symmetric L-densities that are obtained from the density evolution equation (6) (where $l \ge 0$ denotes the number of iterations). From (5), it follows that a necessary and sufficient condition for obtaining vanishing bit error probability under BP decoding is that the B-parameter that is associated with the *pdf* a_l tends to zero, i.e.,
$$\lim_{l\to\infty} \mathcal{B}(a_l) = 0. \tag{32}$$
From (7), it follows that if the sequence $\{y_l\}$ as defined in (9) by the recursive equation
$$y_l = B\lambda\big(1-\rho(1-y_{l-1})\big), \quad l = 1,2,\ldots$$
with the initial condition $y_0 = B$ tends to zero, then also the sequence $\{x_l\}$ where
$$x_l = \mathcal{B}(a_l)$$
tends to zero (since $0 \le x_l \le y_l$ for every integer $l \ge 0$, see (9) and the paragraph that follows). The sequence $\{y_l\}$ refers to the density evolution analysis for a BEC whose channel erasure probability is B. The threshold value, which determines a necessary and sufficient condition for the convergence of the sequence $\{y_l\}$ to zero, yields that if $B < B_0(\lambda,\rho)$ then $\lim_{l\to\infty} y_l = 0$ (this follows from [26, Theorem 3.59]). Hence, for every MBIOS channel, if the B-parameter is less than $B_0(\lambda,\rho)$, then the property in (32) is satisfied, and therefore the bit error probability vanishes under BP decoding. This completes the proof of the first part.

In order to prove the second part of the theorem, which refers to a sequence of right-regular LDPC code ensembles, we rely on inequality (19). If the equality in (32) holds, then it follows from (19) that the sequence $\{z_l\}$ in (23) should necessarily tend to zero. Hence, from (5), if the sequence that is defined in (23) via the recursive equation
$$z_l = B\lambda\big(\sqrt{1-\rho(1-z_{l-1}^2)}\big), \quad l = 1,2,\ldots$$
stays bounded away from zero, with the initial value $z_0 = B$, then the communication is not reliable. More explicitly, for every MBIOS channel whose B-parameter is greater than $B_1(\lambda,\rho)$, the sequence $\{\mathcal{E}(a_l)\}$ that represents the left-to-right message error probabilities under BP decoding stays bounded away from zero (irrespective of the number of iterations). In order to proceed, the following lemma considers the convergence of the sequence $\{z_l\}$.

Lemma 3. Let $B_1(\lambda,\rho)$ be defined as in (30). If $B < B_1(\lambda,\rho)$ then the sequence $\{z_l\}$ in (23) tends to zero, and if $B > B_1(\lambda,\rho)$ then the sequence $\{z_l\}$ is lower bounded by the positive constant
$$x(B) \triangleq \max\left\{x\in(0,1]:\ \frac{x}{\lambda\big(\sqrt{1-\rho(1-x^2)}\big)} \le B\right\}. \tag{33}$$

Proof: See Appendix C. ■

From (19)

$$\mathcal{B}(a_l) \triangleq x_l \geq z_l, \quad l = 0, 1, \ldots$$

and therefore it follows from Lemma 3 that for every MBIOS channel whose B-parameter is greater than $B_1(\lambda, \rho)$

$$\mathcal{B}(a_l) \geq x(B), \quad l = 0, 1, \ldots.$$

From (5) we have

$$\begin{aligned} &\mathcal{B}(a_l) \leq 2\sqrt{\mathcal{E}(a_l)} \\ &\Rightarrow \mathcal{E}(a_l) \geq \left(\frac{\mathcal{B}(a_l)}{2}\right)^2 \geq \left(\frac{x(B)}{2}\right)^2 \end{aligned}$$

and therefore the left-to-right message error probability cannot be reduced below the positive value as above, irrespective of the number of iterations of the BP decoder. This completes the proof of Theorem 3. ■

Corollary 2. For every MBIOS channel with B-parameter $B > B_1(\lambda, \rho)$, let $x(B)$ be defined as in (33). Then, the (average) left-to-right message error probability is bounded away from zero by the universal bound

$$\eta \triangleq \lim_{B \to B_1(\lambda,\rho)^+} \left(\frac{x(B)}{2}\right)^2 \tag{34}$$

irrespective of the number of iterations of the BP decoder.

Proof: By definition, $x(B)$ in (33) is an increasing function of B, and therefore we take the limit $B \to B_1(\lambda, \rho)$, where the limit is from the right side, in order to obtain a lower bound on the left-to-right message error probability for the case where $B > B_1(\lambda, \rho)$. ■

Corollary 3. Let $\{(n, \lambda, \rho)\}$ be a sequence of right-regular LDPC code ensembles whose block lengths tend to infinity. Then, the left-to-right message error probability stays bounded away from zero under BP decoding for every MBIOS channel whose B-parameter is greater than

$$B_2(\lambda, \rho) \triangleq \min\left\{B_1(\lambda, \rho), \frac{1}{\lambda'(0)\rho'(1)}, \sqrt{1 - R_{\mathrm{d}}^2}\right\} \tag{35}$$

where B_1 is introduced in (30), and

$$R_{\mathrm{d}} \triangleq 1 - \frac{\int_0^1 \rho(x)\,\mathrm{d}x}{\int_0^1 \lambda(x)\,\mathrm{d}x}$$

designates the design rate.

Proof: If the B-parameter B is greater than $B_1(\lambda, \rho)$, then the statement follows from the second part of Theorem 3. Also, if $B > \frac{1}{\lambda'(0)\rho'(1)}$, then the communication under BP decoding is not reliable because the stability condition is not satisfied. Finally, if $B > \sqrt{1 - R_{\mathrm{d}}^2}$ then it follows from the right-hand side of (3) that $R_{\mathrm{d}} > C$ and error-free communication cannot be achieved when the design rate exceeds the channel capacity. Therefore, BP decoding is not reliable for any MBIOS channel whose B-parameter is greater than B_2 in (35). ■

Remark 2. In essence, the results of this section stem from one dimensional bounds based on the density evolution equation that utilize the B-parameter (namely, inequalities (7) and (19)). This approach is not new, and was introduced in [2]. In that paper, the authors derived iterative bounds on the expectation of messages transferred in BP decoding. These bounds enabled them to lower- and upper-bound the performance of BP decoding.

Remark 3. Unknown to the authors at the time of writing, Wang, et al. have proved inequalities (7) and (19) in [42]. Wang et al. mention that these inequalities can be used iteratively to derive upper and lower bounds on the decoding threshold based on the initial B-parameter of the channel, and that closed-form solutions for these bounds can be obtained, but do not derive them explicitly. In this paper, we have independently shown these inequalities and have also explicitly derived a closed-form solution of the bounds. Moreover, we have shown a lower bound on the decoding error probability when the B-parameter of the channel exceeds the value in (35).

Although here we concentrate only on channels with symmetric outputs, we note that [42, Theorem 4] extended inequalities (7) and (19) also to memoryless channels with binary input and non-symmetric output, under the assumption that the input distribution is uniform.

Remark 4. Note that if the B-parameter is above $B_2(\lambda, \rho)$ (see (35)), then Corollary 3 does not specify an explicit positive lower bound on the left-to-right message error probability under BP decoding. However, if $B > B_1(\lambda, \rho)$ (where it readily follows from (35) that $B_1(\lambda, \rho) \geq B_2(\lambda, \rho)$), then the second part of Theorem 3 determines an explicit positive lower bound on the left-to-right message error probability that is valid universally for all MBIOS channels. As shown in Examples 1 and 2 that follow, the value of this lower bound η (see (34)) is typically large, irrespective of the number of iterations of the BP decoder, and this lower bound holds for all MBIOS channels whose B-parameter is above $B_1(\lambda, \rho)$.

Remark 5. All channels in the convex hull[5] of equi-capacity MBIOS channels have the same capacity. Similarly, all channels in the convex hull of equi-B-parameter MBIOS channels have the same B-parameter. This is due to the linearity of the capacity and Bhattacharyya functionals in the L-density function, see (1) and (2). Therefore, the condition for good channels, under BP decoding, in the sense that $B < B_0(\lambda, \rho)$ (see the first part of Theorem 3) or the condition for bad channels in the sense that $B > B_2(\lambda, \rho)$ (see the second part of Theorem 3 and Corollaries 2 and 3) are both preserved, respectively, for the convex hull of good or bad channels. Although this conclusion does not prove [28, Conjecture 1] for equi-capacity MBIOS channels (since it does not cover the case where B is between B_0 and B_2 in case that $B_0 < B_2$), it supports this conjecture in the cases where $B < B_0$ or $B > B_2$.

Remark 6. From the two parts of Theorem 3, it follows directly that for right-regular codes, $B_1(\lambda, \rho) \geq B_0(\lambda, \rho)$. For a direct proof of this inequality, see Appendix D.

In the following, we exemplify the use of Theorem 3 and its corollaries:

Example 1 (Regular LDPC Code Ensembles)**.** In Table III, we show the numerical values of B_0 and B_1 in Theorem 3, and the value of η in Corollary 2 for some regular LDPC code ensembles whose design rate is one-half. The value of B_0 corresponds to the threshold for the BEC under BP decoding, and the value of B_1 (see (30)) refers to the value of the B-parameter where above it, the left-to-right message error probability is at least η, no matter how many iterations of the BP decoder are performed. For these regular LDPC code ensembles, $\lambda_2 = 0$,

TABLE III
THE NUMERICAL VALUES OF B_0 AND B_1 IN THEOREM 3, AND THE VALUE OF η IN COROLLARY 2 FOR SOME REGULAR LDPC CODE ENSEMBLES WHOSE DESIGN RATE IS ONE-HALF.

LDPC	B_0	B_1	η
(3,6)	0.4294	0.6553	$6.50 \cdot 10^{-2}$
(4,8)	0.3834	0.6192	$6.58 \cdot 10^{-2}$
(5,10)	0.3416	0.5884	$6.18 \cdot 10^{-2}$

and therefore the stability condition is useless. Also, since the design rate of these ensembles is equal to one-half, then $\sqrt{1 - R_{\mathrm{d}}^2} = \frac{\sqrt{3}}{2} \approx 0.8660$, hence the values of B_2 in (35) coincide with B_1 for these ensembles.

Example 2 (Optimized Right-Regular LDPC Code Ensembles for the BEC, and their Universal Properties)**.** In Table IV, right-regular LDPC code ensembles are optimized for the BEC under BP decoding; to this end, a linear program is solved as described in [26, Section 3.18] for a design rate of one-half ($R_{\mathrm{d}} = \frac{1}{2}$) and for a maximal degree of the variable nodes of one hundred ($d_{\mathrm{v}}^{\max} = 100$).

From Table IV, it can be seen that $B_0 \approx B_2$. Hence, for these optimized LDPC code ensembles, the first part of Theorem 3 states that these LDPC code ensembles are reliable under BP decoding, in the sense of achieving vanishing bit error probability, for every MBIOS channel whose B-parameter is below B_0; on the other hand, Corollary 3 implies that these code ensembles are not reliable under BP decoding for every MBIOS channel

[5]The convex hull of a set $\mathcal{A}$ of channels consists of all the channels that are convex combinations of channels in $\mathcal{A}$. A convex combination of several channels is the result of using each channel with probability θ_i, $0 \leq \theta_i \leq 1$, such that $\sum_i \theta_i = 1$.

TABLE IV

THE DEGREE DISTRIBUTIONS (FROM THE EDGE PERSPECTIVE), NUMERICAL VALUES OF B_0 AND B_1 IN THEOREM 3, THE VALUE OF η IN COROLLARY 2, AND THE VALUE OF B_2 IN COROLLARY 3 FOR SOME OPTIMIZED RIGHT-REGULAR LDPC CODE ENSEMBLES WHOSE DESIGN RATE IS ONE-HALF WITH A MAXIMAL DEGREE OF THE VARIABLE NODES THAT IS SET TO 100.

$\lambda(x) = \sum_i \lambda_i x^{i-1}$	$\rho(x) = \sum_i \rho_i x^{i-1}$	B_0	B_2	B_1	η
$\lambda_2 = 0.4127,\ \lambda_3 = 0.1762$ $\lambda_4 = 0.1177,\ \lambda_7 = 0.1202$ $\lambda_8 = 0.1731$	$\rho_6 = 1$	0.4816	0.4846	0.7066	$8.45{\cdot}10^{-2}$
$\lambda_2 = 0.2879,\ \lambda_3 = 0.1222$ $\lambda_4 = 0.0905,\ \lambda_6 = 0.1174$ $\lambda_7 = 0.0300,\ \lambda_{12} = 0.0807$ $\lambda_{13} = 0.0831,\ \lambda_{32} = 0.0050$ $\lambda_{33} = 0.1831$	$\rho_8 = 1$	0.4962	0.4962	0.7146	$1.02{\cdot}10^{-1}$
$\lambda_2 = 0.2226,\ \lambda_3 = 0.1013$ $\lambda_4 = 0.0504,\ \lambda_5 = 0.0646$ $\lambda_6 = 0.0445,\ \lambda_{10} = 0.1219$ $\lambda_{11} = 0.0117,\ \lambda_{24} = 0.0903$ $\lambda_{25} = 0.0678,\ \lambda_{100} = 0.2248$	$\rho_{10} = 1$	0.4988	0.4992	0.7123	$1.08{\cdot}10^{-1}$

whose B-parameter is slightly above B_0 or greater than this value. This is a universal result that applies to all MBIOS channels, and it separates them into two sets of good or bad channels for which the reliability of these code ensembles under BP decoding solely depends on the B-parameter of the communication channel *without any relevance to its channel model* (as long as it is MBIOS, and it exhibits a given B-parameter).

In contrast to the results in Table III that apply to regular LDPC code ensembles, for the right-regular LDPC code ensembles studied in this example, the value of B_1 is significantly greater than B_2, which here is given by the stability condition. In continuation to Remark 4, the lower bound on the left-to-right message error probabilities when the B-parameter is greater than B_1 is rather large (around 0.1), whereas such a measure is not provided here for the unreliability of the messages when the B-parameter is between B_1 and B_2.

The results of this paper imply that a family of degraded channels can be parameterized by the B-parameter. This is also supported by [26, Theorem 4.76], which states that a degraded channel has a higher B-parameter than the original (see also Proposition 2 in this paper). Moreover, in many cases there is a simple one-to-one correspondence between the channel parameter and the B-parameter. For example, for a BEC with erasure probability ϵ, we have $B = \epsilon$; for a BSC with crossover probability p, we have $B = \sqrt{4p(1-p)}$; and for a BIAWGN channel with noise variance σ^2, we have $B = e^{-\frac{1}{2\sigma^2}}$.

In order to obtain bounds on the B-parameter for any LDPC code ensemble, not necessarily right-regular, a simple modification of Theorem 3 and Corollary 3 yields the following:

Corollary 4. Let $\{(n, \lambda, \rho)\}$ be a sequence of (not necessarily right-regular) LDPC code ensembles whose block lengths tend to infinity. Then

- This sequence achieves vanishing bit error probability under BP decoding for *every* MBIOS channel whose B-parameter is less than
$$B_0(\lambda, \rho) \triangleq \inf_{x \in (0,1]} \frac{x}{\lambda\big(1 - \rho(1-x)\big)}.$$
- The left-to-right message error probability of this sequence stays bounded away from zero under BP decoding for every MBIOS channel whose B-parameter is greater than
$$B_3(\lambda, \rho) \triangleq \begin{cases} \min\left\{B_1(\lambda,\rho), \frac{1}{\lambda'(0)\rho'(1)}, \sqrt{1-R_{\mathrm{d}}^2}\right\}, \\ \qquad \text{if the sequence is right-regular.} \\ \min\left\{\frac{1}{\lambda'(0)\rho'(1)}, \sqrt{1-R_{\mathrm{d}}^2}\right\}, \\ \qquad \text{if the sequence is not right-regular.} \end{cases}$$

TABLE V
COMPARISON OF UNIVERSAL BOUNDS ON THRESHOLDS FOR VARIOUS LDPC CODE ENSEMBLES. THE BOUNDS BASED ON THE B-PARAMETER ARE COMPUTED BASED ON THE APPROACH PRESENTED IN THIS PAPER; THE BOUNDS BASED ON THE CAPACITY ARE COMPUTED ACCORDING TO [37].

$\lambda(x) = \sum_i \lambda_i x^{i-1}$	$\rho(x) = \sum_i \rho_i x^{i-1}$		Bounds based on B	Bounds based on C
$\lambda_3 = 1$	$\rho_6 = 1$		$0.4294 < B < 0.6553$	$0.4744 < C < 0.6350$
		BSC:	$0.0485 < p < 0.1223$	$0.0698 < p < 0.1187$
		BIAWGN:	$0.7691 < \sigma < 1.0877$	$0.8026 < \sigma < 1.0180$
$\lambda_4 = 1$	$\rho_8 = 1$		$0.3834 < B < 0.6192$	$0.5160 < C < 0.6630$
		BSC:	$0.0382 < p < 0.1074$	$0.0624 < p < 0.1048$
		BIAWGN:	$0.7222 < \sigma < 1.0214$	$0.7707 < \sigma < 0.9553$
$\lambda_5 = 1$	$\rho_{10} = 1$		$0.3416 < B < 0.5844$	$0.5564 < C < 0.6970$
		BSC:	$0.0301 < p < 0.0943$	$0.0540 < p < 0.0921$
		BIAWGN:	$0.6822 < \sigma < 0.9648$	$0.7333 < \sigma < 0.8996$
$\lambda_2 = 0.4127,\ \lambda_3 = 0.1762$ $\lambda_4 = 0.1177,\ \lambda_7 = 0.1202$ $\lambda_8 = 0.1731$	$\rho_6 = 1$		$0.4816 < B < 0.4846$	$0.4147 < C < 0.8980$
		BSC:	$0.0618 < p < 0.0626$	$0.0133 < p < 0.1404$
		BIAWGN:	$0.8272 < \sigma < 0.8308$	$0.5182 < \sigma < 1.1209$
$\lambda_2 = 0.2879,\ \lambda_3 = 0.1222$ $\lambda_4 = 0.0905,\ \lambda_6 = 0.1174$ $\lambda_7 = 0.0300,\ \lambda_{12} = 0.0807$ $\lambda_{13} = 0.0831,\ \lambda_{32} = 0.0050$ $\lambda_{33} = 0.1831$	$\rho_8 = 1$		$0.4962 \leq B \leq 0.4962$	$0.3989 < C < 0.8910$
		BSC:	$0.0659 \leq p \leq 0.0659$	$0.0144 < p < 0.1465$
		BIAWGN:	$0.8446 \leq \sigma \leq 0.8447$	$0.5265 < \sigma < 1.1513$

where B_1 is introduced in (30), and R_{d} designates the design rate.

It follows that for any family of MBIOS channels there exists a B_{th} between B_0 and B_3 (its exact value is dependent on the family) such that BP decoding achieves vanishing bit error probability for all channels of this family with $B < B_{\mathrm{th}}$ and does not achieve vanishing bit error probability for channels of the family with $B > B_{\mathrm{th}}$. Hence, $B_0(\lambda, \rho)$ and $B_3(\lambda, \rho)$ provide universal lower and upper bounds on the threshold B-parameter. In general, different channel families will have different thresholds. It should be noted that the lower bound is tight for the BEC. Furthermore, this universal bound is non-iterative, simple, and easy to compute.

Similar bounds on the B-parameters have been derived in [42]. In that paper, the authors have derived the same inequalities on the B-parameter evolution during BP decoding that have led to the bounds on the threshold in this section. Therefore, the lower bound $B_0(\lambda, \rho)$ and the upper bound $B_1(\lambda, \rho)$ are not new. In this paper, however, we have combined the upper bound $B_1(\lambda, \rho)$ with other upper bounds, such as the stability condition, to arrive at a tighter upper bound in some cases. Moreover, we have also demonstrated that these bounds can be tight in some cases, as shown in Table IV.

Other works, such as [14], [26, Section 4.10.2], and [37] used an information-combining approach to also provide universal bounds on the threshold. These bounds give upper and lower bounds on the capacity of the channel, another natural parameter for channel degradation (a degraded channel has lower capacity). These bounds are significantly more difficult to compute, requiring either an iterative process or involving computations that are numerically unstable for high left degrees.

In Table V, we compare the bounds suggested by this approach with the bounds of [37] for some of the ensembles considered in this paper. In order to make the comparison, we translate the bounds on the B-parameter to bounds on the channel parameters for a BSC and a BIAWGN channel. It is exemplified that the bounds in [37] are superior for the regular LDPC code ensembles, but the bounds of the approach presented here are more informative for the irregular LDPC code ensembles shown in Table V.

Comparing these information-combining results by Land et al. [14], and by Sutskover et al. ([37], [38]) with our bounds, there is one conceptual difference: for $B > B_1$, Theorem 3 and Corollary 2 provide an explicit lower bound on the left-to-right message error probability that is irrespective of the number of iterations, whereas this is not the case in these related works. Secondly, for the regular LDPC code ensembles, for which we have $B_3 = B_1$, we also have an explicit positive lower bound on the left-to-right message error probability for the case where the B-parameter is larger than B_3 (e.g., as shown in Table III, for the (3,6) LDPC code ensemble, a lower bound on the left-to-right message error probability around 6.5% applies to the cases where $p > 0.1223$ or $\sigma > 1.088$ for the BSC and the binary-input AWGN channel, respectively).

The observation made in Example 2, regarding the reliability of the optimized right-regular LDPC code ensembles over the entire set of MBIOS channels where this result solely depends on the B-parameter of the communication channel (but not on the specific channel model of the MBIOS channel) calls for analysis. Since these LDPC code ensembles were optimized numerically (via linear programming), closed forms for the degree distributions are not available, and we turn instead to consider the sequences of right-regular LDPC code ensembles as suggested by Shokrollahi [35]. In this respect, the following theorem demonstrates a universality property under BP decoding with respect to the entire set of MBIOS channels which exhibit a given B-parameter; the following theorem shows that not only the stability condition is common for the considered set of channels, but also a universality property exists for this set.

Theorem 4. [Universality of LDPC Code Ensembles under BP Decoding for MBIOS Channels with a Fixed B-Parameter] Consider the set of MBIOS channels that exhibit a fixed B-parameter (B). Then:

- Every capacity-achieving sequence designed for BEC(B), universally achieves the following fraction of capacity for the considered set of channels:

$$\mu_3(B) \triangleq \frac{1-B}{1-h_2\left(\frac{1-\sqrt{1-B^2}}{2}\right)}, \tag{36}$$

  where h_2 denotes the binary entropy function to the base 2. The function μ_3 is monotonic decreasing in B; it gets the values $\ln 2 \approx 69.3\%$ and 100% for the extreme cases where $B \to 1$ (i.e., a very noisy channel) and $B \to 0$ (i.e., a perfect channel), respectively.
- There exists an explicit construction of a sequence of right-regular LDPC code ensembles for which B satisfies

$$B \leq B_0 \leq B_2 \leq 1 - \left(\frac{d_{\mathrm{c}}-2}{d_{\mathrm{c}}-1}\right)^{\frac{\pi^2}{6}} e^{\frac{1}{d_{\mathrm{c}}-1}\left(\frac{\pi^2}{6}-\gamma\right)} (1-B) \tag{37}$$

  so B_0 and B_2 can be made arbitrarily close to B for large d_{c}. Here d_{c} denotes the fixed degree of parity-check nodes, B_0 and B_2 are introduced in (29) and (35) respectively, and $\gamma \approx 0.5772$ denotes Euler's constant.

Proof: Among all MBIOS channels which exhibit a given B-parameter B, the capacity is maximized or minimized for a BSC and BEC, respectively. For a BEC, $C = 1 - B$, and therefore the capacity is achieved (i.e., $R_{\mathrm{d}} = C$) because of (10). For a BSC whose crossover probability is p,

$$C = 1 - h_2(p), \quad B = \sqrt{4p(1-p)}$$

and therefore

$$C = 1 - h_2\left(\frac{1-\sqrt{1-B^2}}{2}\right).$$

From (10), the fraction of capacity that is universally achieved for the entire set of MBIOS channels which exhibit a given B-parameter B satisfies

$$\frac{1-B}{1-h_2\left(\frac{1-\sqrt{1-B^2}}{2}\right)} \leq \frac{R_{\mathrm{d}}}{C} \leq 1 \tag{38}$$

where the upper and lower bounds are obtained, respectively, for a BEC and BSC with a B-parameter B. Let us check the two extreme cases where $B = 0$ and $B \to 1$ (referring, respectively, to an ideal channel and a very noisy channel). In the case where $B = 0$, the upper and lower bounds coincide, and are equal to 1; hence, capacity is achievable. For examining the case where $B \to 1$, we rely on the following Taylor series expansion of the binary entropy function around $x = \frac{1}{2}$ (see [43, p. 575]):

$$h_2(x) = 1 - \frac{1}{2\ln 2} \sum_{q=1}^{\infty} \frac{(1-2x)^{2q}}{q(2q-1)}, \quad 0 \leq x \leq 1 \tag{39}$$

which enables to calculate the limit of the left-hand side in (38) when $B \to 1$ (from below). This gives

$$\begin{aligned}
&\lim_{B\to 1^-} \frac{1-B}{1-h_2\left(\frac{1-\sqrt{1-B^2}}{2}\right)} \\
&= \lim_{B\to 1^-} \frac{1-B}{\frac{1}{2\ln 2}\sum_{q=1}^{\infty}\frac{(1-B^2)^q}{q(2q-1)}} \\
&= \lim_{B\to 1^-} \frac{1-B}{\left(\frac{1-B^2}{2\ln 2}\right)} \\
&= \ln 2.
\end{aligned}$$

Moreover, it is easy to verify with (39) that the lower bound on $\frac{R_{\mathrm{d}}}{C}$ in (38) forms a monotonic decreasing function of B (where $0 \leq B < 1$); it varies from 1 to $\ln 2 \approx 0.693$ as the value of B is increased from zero to 1 bit per channel use. This shows that, for the entire set of MBIOS channels which exhibit a given B-parameter B, the achievable fraction (10) of capacity is at least 69.3%; this result is obtained by designing a capacity-achieving sequence of LDPC code ensembles for a BEC whose B-parameter matches our channel (as above). Interestingly, these two extreme values (i.e., 69.3% and 100%) coincide with those obtained in Theorem 1 for the entire set of equi-capacity MBIOS channels.

To prove the second part of the Theorem, we consider a sequence of right-regular LDPC code ensembles with a fixed right-degree d_{c}, and parameters of the degree distributions that are selected according to [30, Theorem 2.3] for a BEC with channel erasure probability B (see also [26, Section 3.15] and [33, Appendix VI]). These sequences are capacity-achieving as we let the right degree d_{c} tend to infinity. From [30, Theorems 2.1 and 2.3], this sequence is constructed to achieve at least a fraction $1-\varepsilon$ of the capacity of the BEC under BP decoding with a right degree d_{c} that scales logarithmically with the reciprocal of the gap to capacity, i.e., it behaves like $\log \frac{1}{\varepsilon}$.

For a BEC, the B-parameter of the channel is equal to the channel erasure probability. The sequence of right-regular LDPC code ensembles is designed to achieve vanishing bit erasure probability under BP decoding for a BEC whose channel erasure probability is set to B (since, by assumption, the parameters (α and N) of its degree distributions are selected according to [30, Theorem 2.3]). Hence, the threshold of this sequence, B_0, under BP decoding is greater than or equal to B. This proves the left-hand side of inequality (37).

We derive in the following the upper bound on B_2 in this inequality, based on [33, Appendix VI]. More explicitly, let $c(\alpha, N)$ be the function (see [33, Eq. (116)])

$$c(\alpha, N) \triangleq (1-\alpha)^{\frac{\pi^2}{6}}\, e^{\alpha\left(\frac{\pi^2}{6}-\gamma+\frac{1}{2N}\right)}. \tag{40}$$

for $0 < \alpha < 1$ and an integer $N \geq 1$ (on the right-hand side of this equality, $\gamma \approx 0.5772$ denotes Euler's constant). The fraction of edges attached to degree-2 variable nodes, for this right-regular sequence, satisfies (see [33, Eq. (117)])

$$\frac{\alpha}{1-c(\alpha,N)\,(1-B)} < \lambda_2 \leq \frac{\alpha}{B} \tag{41}$$

where $\alpha \triangleq \frac{1}{d_{\mathrm{c}}-1}$. From (35), (40) and (41)

$$\begin{aligned}
B_2 &\leq \frac{1}{\lambda'(0)\rho'(1)} \\
&= \frac{\alpha}{\lambda_2} \\
&\leq 1 - c(\alpha,N)(1-B) \\
&= 1 - (1-\alpha)^{\frac{\pi^2}{6}} e^{\alpha\left(\frac{\pi^2}{6}-\gamma+\frac{1}{2N}\right)}(1-B) \\
&\leq 1 - (1-\alpha)^{\frac{\pi^2}{6}} e^{\alpha\left(\frac{\pi^2}{6}-\gamma\right)}(1-B) \\
&= 1 - \left(\frac{d_{\mathrm{c}}-2}{d_{\mathrm{c}}-1}\right)^{\frac{\pi^2}{6}} e^{\frac{1}{d_{\mathrm{c}}-1}\left(\frac{\pi^2}{6}-\gamma\right)}(1-B).
\end{aligned}$$

This completes the proof of (37). Following the first part of Theorem 3 and Corollary 3, it follows that in the limit where $d_{\mathrm{c}} \to \infty$, $B_2 \leq 1-(1-B) = B$, so that (37) yields that $B_0 = B_2 = B$. Hence,

- the BP decoder achieves vanishing bit error probability for every MBIOS channel whose B-parameter is less than B,
- it is unreliable (i.e., the left-to-right message error probability is bounded away from zero) for every MBIOS channel whose B-parameter is greater than B.

This completes the proof of Theorem 4. ■

Remark 7. Another way to prove the first part of the above theorem is via use of the approach of sub-section III-A. In the setting of Theorem 4, the family of MBIOS channels being considered is the one that exhibits the same B-parameter (regardless of capacity). Over this family, the BSC and BEC exhibit the maximal and minimal capacities, respectively. Following the approach of sub-section III-A, we construct a capacity-achieving sequence of LDPC ensembles for a BEC with erasure probability B. The design rate of this ensemble is $R_{\mathrm{d}} = 1-B$. Since the BSC exhibits the maximal capacity over this set of MBIOS channels, the universally achievable fraction of capacity is $\frac{R_{\mathrm{d}}}{C}$, where C is the capacity of a BSC with B-parameter B. Using the expressions for R_{d} and C we obtain that the universally achievable fraction of capacity is indeed $\mu_3(B)$.

Table VI shows the resulting achievable fraction of capacity in (36) as a function of the B-parameter of the considered set of MBIOS channels.

TABLE VI
UNIVERSAL ACHIEVABLE FRACTION OF CAPACITY UNDER BP DECODING FOR THE ENTIRE SET OF MBIOS CHANNELS WHICH EXHIBIT A GIVEN B-PARAMETER B (SEE THEOREM 4).

B	$\mu_3(B)$
0	100%
0.250	85.0%
0.333	82.0%
0.500	77.5%
0.750	72.7%
1.000	69.3%

Corollary 5. In the limit where $d_{\mathrm{c}} \to \infty$, the BP decoder in Theorem 4 achieves vanishing bit error probability for all MBIOS channels whose B-parameter is less than B, and it is unreliable (i.e., the left-to-right message error probability is bounded away from zero) for every MBIOS channel whose B-parameter is greater than B. For finite d_{c}, the values of B_0 and B_2 differ from B by at most

$$(1-B)\left(\frac{\gamma}{d_{\mathrm{c}}-1} + \frac{\frac{\pi^2}{6}\left(\frac{\pi^2}{6}-\gamma\right)}{(d_{\mathrm{c}}-1)^2}\right),$$

and this difference tends uniformly to zero for $0 \leq B \leq 1$ as we let d_{c} tend to infinity.

Proof: The first part of this corollary, for infinite d_{c} is immediate from (37). For finite d_{c}, subtracting B from (37) yields

$$\begin{aligned} 0 \leq B_0 - B &\leq B_2 - B \\ &\leq (1-B)\left(1 - \left(\frac{d_{\mathrm{c}}-2}{d_{\mathrm{c}}-1}\right)^{\frac{\pi^2}{6}} e^{\frac{1}{d_{\mathrm{c}}-1}\left(\frac{\pi^2}{6}-\gamma\right)}\right). \end{aligned} \tag{42}$$

Bernoulli's inequality states that $(1+x)^r \geq 1+rx$ for $x > -1$, $r \geq 1$. Thus,

$$\left(\frac{d_{\mathrm{c}}-2}{d_{\mathrm{c}}-1}\right)^{\frac{\pi^2}{6}} = \left(1 - \frac{1}{d_{\mathrm{c}}-1}\right)^{\frac{\pi^2}{6}} \geq 1 - \frac{\pi^2/6}{d_{\mathrm{c}}-1}. \tag{43}$$

Moreover, for every $y > 0$ we have $e^y \geq 1 + y$, so that

$$e^{\frac{1}{d_c - 1}\left(\frac{\pi^2}{6} - \gamma\right)} \geq 1 + \frac{\pi^2/6 - \gamma}{d_c - 1}\,. \tag{44}$$

Using (42)–(44) gives

$$\begin{aligned}
0 &\leq B_0 - B \\
&\leq B_2 - B \\
&\leq (1 - B)\left(1 - \left(1 - \frac{\pi^2/6}{d_c - 1}\right)\left(1 + \frac{\pi^2/6 - \gamma}{d_c - 1}\right)\right) \\
&= (1 - B)\left(\frac{\gamma}{d_c - 1} + \left(\frac{\pi^2}{6} - \gamma\right)\frac{\pi^2/6}{(d_c - 1)^2}\right).
\end{aligned}$$

From the above inequality it is clear that as we let $d_c \to \infty$, the differences $B_0 - B$ and $B_2 - B$ tend to zero uniformly. ■

Example 3. For ensemble no. 2 in Table IV, whose design rate is $R_d = \frac{1}{2}$ bits per channel use, the threshold under BP decoding corresponds *uniformly* to the B-parameter $B = 0.4962$ for every MBIOS channel. For the BEC, this corresponds to a capacity of $C = 1 - B = 0.5038$ bits per channel use, and therefore 99.3% of the capacity of the BEC is achieved under BP decoding with vanishing bit erasure probability. For the BIAWGN channel, this corresponds to channel capacity $C = 0.5977$ bits per channel use, and therefore this code ensemble achieves 83.4% of the capacity for this channel. The smallest fraction of capacity under BP decoding is achieved for the BSC. The B-parameter $B = 0.4962$ corresponds to $C = 0.6496$ for the BSC, which means that 77.0% of capacity is achieved under BP decoding.

Example 4. In this example, we consider a right-regular LDPC code ensemble, whose design rate is $R_d = 0.9$ bits per channel use, and which closely approaches the capacity of the BEC under BP decoding. To this end, we set the degree of the parity-check nodes to be 40, and the maximal variable node degree is set to 200. The following degree distributions are obtained by linear programming with the approach in [26, Section 3.18]:

$$\begin{aligned}
\lambda(x) &= 0.2638x + 0.1259x^2 + 0.1088x^3 + 0.0551x^5 \\
&\quad + 0.1589x^6 + 0.0278x^{15} + 0.2598x^{16}\,, \\
\rho(x) &= x^{39}.
\end{aligned}$$

From (29), (30), and (35)

$$B_0 = 0.0972, \quad B_1 = 0.3185, \quad B_2 = 0.0972$$

and therefore, since $B_0 = B_2$, then for every MBIOS channel, this LDPC code ensemble achieves vanishing bit error probability under BP decoding if the B-parameter is below $B = 0.0972$, and it is unstable if the B-parameter exceeds this value. This enables to calculate the threshold under BP decoding by transforming the B-parameter to the proper channel parameter. For the BEC, this corresponds to capacity of $C = 1 - B = 0.9028$ bits per channel use, and therefore 99.7% of the capacity of the BEC is asymptotically obtained under BP decoding with vanishing bit erasure probability. For the BIAWGN channel, this corresponds to channel capacity $C = 0.9400$ bits per channel use, and therefore this code ensemble achieves 95.7% of the capacity for this channel. The smallest fraction of capacity under BP decoding is achieved for the BSC, and it coincides with the lower bound $\mu_3(B) = 92.5\%$ as given in (36).

Example 5. We note that the approach presented in the examples above is not necessarily the best approach for obtaining universal LDPC code ensembles. Numerically optimized code ensembles may lead to better performance under BP over some channels. To demonstrate this, we consider the following LDPC code ensemble, obtained

using [16]:

$$\begin{aligned}\lambda(x) &= 0.244022x + 0.224973x^2 + 0.0476526x^5 \\ &\quad + 0.225756x^6 + 0.0270727x^{18} + 0.173877x^{19} \\ &\quad + 0.0515554x^{20} + 0.00509134x^{22}\ , \\ \rho(x) &= x^8.\end{aligned}$$

This code is numerically optimized for the BIAWGN channel, with a design rate of one-half; its threshold under BP decoding is $\sigma = 0.966293$. This corresponds to a capacity of $C = 0.5084$ bits per channel use. Therefore, this code achieves 98.35% of the capacity of the BIAWGN channel. The threshold of this code under BP decoding for the BEC computes to be $B = 0.4741$, which corresponds to a capacity of $C = 1 - B = 0.5259$ bits per channel use. I.e., this code achieves 95.08% of the capacity for the BEC.

The ensemble above and the ensemble considered in Example 3 share the same design rate. We see that the code ensemble considered here, when used over a BIAWGN channel, is superior to the ensemble of Example 3, achieving a much higher fraction of capacity. The performance of the two ensembles over the BEC, however, is similar, with a slight advantage to the ensemble of Example 3, recognizing that it was designed for a BEC.

This observation is also supported by the numerical results presented in [23]. In that work, the authors compared how LDPC code ensembles designed for one MBIOS channel performed over other MBIOS channels. The channels considered there were the BEC, the BIAWGN, and the flat-fading binary input Rayleigh channel. Their results show that the BEC can indeed be used as a so-called "surrogate" channel for the design of good LDPC code ensembles, while recognizing that better results can be obtained, at the expense of a higher computational load, with numerical optimization for the desired channel.

While the approach presented here may not be the optimal approach, it is *analytical and easy to compute*, and thus provides insight. For instance, we have shown how this approach can be used to obtain bounds on the thresholds of ensembles under BP decoding over *any* channel, which, as exemplified in Examples 3 and 4 above, are tight for some ensembles.

Remark 8. Universality results for LDPC code ensembles have been derived in this section with vanishing *bit* error probability under BP decoding. An extension of these results for vanishing *block* error probability can be made based on the results of [12] and [17]. These works showed that for a specific MBIOS channel, an LDPC code ensemble with $\lambda_2 = 0$ has the same threshold under vanishing block and bit error probabilities[6]. The threshold for vanishing block error probability, similar to the threshold for vanishing bit error probability, is defined as the maximal channel parameter for which the block error probability will converge to zero. This result is based on the union bound, $P_B \leq nP_b$, where P_B is the block error probability, P_b is the bit error probability, and n is the block length. The conditions on λ_2 ensure that the bit error probability decays fast enough, thus causing the block error probability to vanish as well.

An extension of this result to universality over a multitude of MBIOS channels is now straightforward. As an example, let us demonstrate this by extending the results of Theorem 1. In the setting of this theorem, we consider a set $\mathcal{A}$ of MBIOS channels exhibiting the same capacity, C, and maximal B-parameter B. If the capacity-achieving sequence of LDPC code ensembles $\{(n, \lambda, \rho)\}$ for BEC(B) also satisfies the above-mentioned conditions on λ_2, then this sequence is not only universal over this set in terms of vanishing *bit* error probability, but also in terms of vanishing *block* error probability.

Similarly, by imposing on λ_2 the conditions from [12] and [17], the other results of this section can be extended in a straight-forward manner for universality under vanishing *block* error probability.

IV. Universality under Maximum-Likelihood Decoding

In Section III we considered the universality of LDPC code ensembles under BP decoding. Though maximum-likelihood (ML) decoding is in general prohibitively complex, we show in the following that universality can be achieved under ML decoding for the entire set of equi-capacity MBIOS channels. The universality results proved in Section III under BP decoding automatically hold under ML decoding. In this section, we prove universality results

[6]In fact, [12] presents a stronger condition, also enabling $\lambda_2 > 0$ for code ensembles with a special structure.

for ML decoding that are stronger in the sense that capacity can be approached *arbitrarily closely* with vanishing *block* error probability for the entire set of channels under consideration.

A. Universality of Gallager's Regular LDPC Code Ensembles

In his monograph, Gallager introduced ensembles of regular LDPC codes, and also considered their performance under ML decoding via their distance properties (see [7, Chapters 2 and 3]). In the following, we rely on [30], and demonstrate that a proper selection of Gallager's regular LDPC code ensembles can be made to approach arbitrarily closely the channel capacity for the entire set of equi-capacity MBIOS channels with vanishing block error probability.

Theorem 5. [Universality of Regular LDPC Code Ensembles under ML Decoding for Equi-Capacity MBIOS Channels] Under ML decoding, Gallager's regular LDPC code ensembles can be made universal for the set $\mathcal{A}$ of MBIOS channels that exhibit a given capacity C. More explicitly, for any $\varepsilon > 0$ (that can be made arbitrarily small), there exists a sequence of these code ensembles whose design rate forms at least a fraction $1 - \varepsilon$ of the channel capacity with vanishing block error probability for the entire set $\mathcal{A}$. Moreover, the asymptotic parity-check density of this sequence scales like $\log \frac{1}{\varepsilon}$.

Proof: The proof of the first part of this theorem follows along the lines of the proof of [30, Theorem 2.2] by noticing that the way where the capacity-approaching sequence of regular LDPC code ensembles is determined only depends on the channel capacity. This therefore makes this sequence universal for the entire set of equi-capacity MBIOS channels $\mathcal{A}$, and it asymptotically achieves (as we let the block length of this sequence tend to infinity) vanishing block error probability under ML decoding with a design rate that is at least a fraction $1 - \varepsilon$ of the channel capacity. The asymptotic parity-check density scales like $\log \frac{1}{\varepsilon}$, which is a consequence of the upper and lower bounds on the parity-check density in [30, Theorem 2.2] and [30, Theorem 2.1], respectively, which both scale like $\log \frac{1}{\varepsilon}$. ■

Remark 9. [Capacity-Achieving Code Ensembles with Bounded Graphical Complexity under ML Decoding] In [11], some ensembles of codes defined on graphs are demonstrated to achieve capacity under ML decoding with *bounded graphical complexity*. This holds for every MBIOS channel, but the degree distributions are designed for a specific MBIOS channel, so the question of the universality over a family of MBIOS channels with fixed capacity is an open issue. We note, however, that the existence of such capacity-achieving code ensembles whose graphical complexity stays bounded, as opposed to the $\log \frac{1}{\varepsilon}$ result in [30, Theorem 2.1] and Theorem 5, is attributed to the existence of state nodes (e.g., punctured bits) in the Tanner graphs of irregular repeat-accumulate codes whereas LDPC code ensembles without puncturing are represented by bipartite graphs without these additional state nodes. It is an open question if the code ensembles discussed in [11] have the potential of being universal capacity-achieving code ensembles under ML decoding with respect to a multitude of capacity-achieving MBIOS channels. Note however that since the decoding complexity of ML decoding does not depend directly on the graphical complexity of these ensembles (as opposed to iterative BP decoding), and on the other hand $\log \frac{1}{\varepsilon}$ stays relatively small even for $\varepsilon = 10^{-3}$ (thus referring to the case of achieving 99.9% of the capacity), we choose to leave the analysis of the possible universality of the new capacity-achieving code ensembles in [11] under ML decoding outside the scope of this paper.

Example 6. In order to exemplify Theorem 5, consider lower bounds on the error exponents of some expurgated Gallager's regular LDPC code ensembles under ML decoding. Fig. 2 shows lower bounds on the error exponent for several expurgated Gallager's LDPC code ensembles of length $n = 100000$ and design rate $\frac{1}{2}$. The expurgation followed the approach in [7, Chapter 2]. The bounds were computed for three MBIOS channels of different capacities. For the BSC and BEC, the Shulman-Feder bound was used (see [29, section 4.4.1]). For the BIAWGN channel, the error exponent was computed based on [40, Theorem 3.1]. The distance spectra of the ensembles were computed according to the asymptotic results in [7, Chapter 2]. It is noted that for this block length, the asymptotic results are very close to the exact distance spectra (see [39]). It is evident from Fig. 2 that as we increase the degrees of the variable and check nodes while maintaining a constant design rate, the point where the error exponent vanishes gets closer to the channel capacity, regardless of the MBIOS channel in question. Thus, this demonstrates that this sequence of ensembles becomes universal under maximum-likelihood decoding for equi-capacity MBIOS channels.

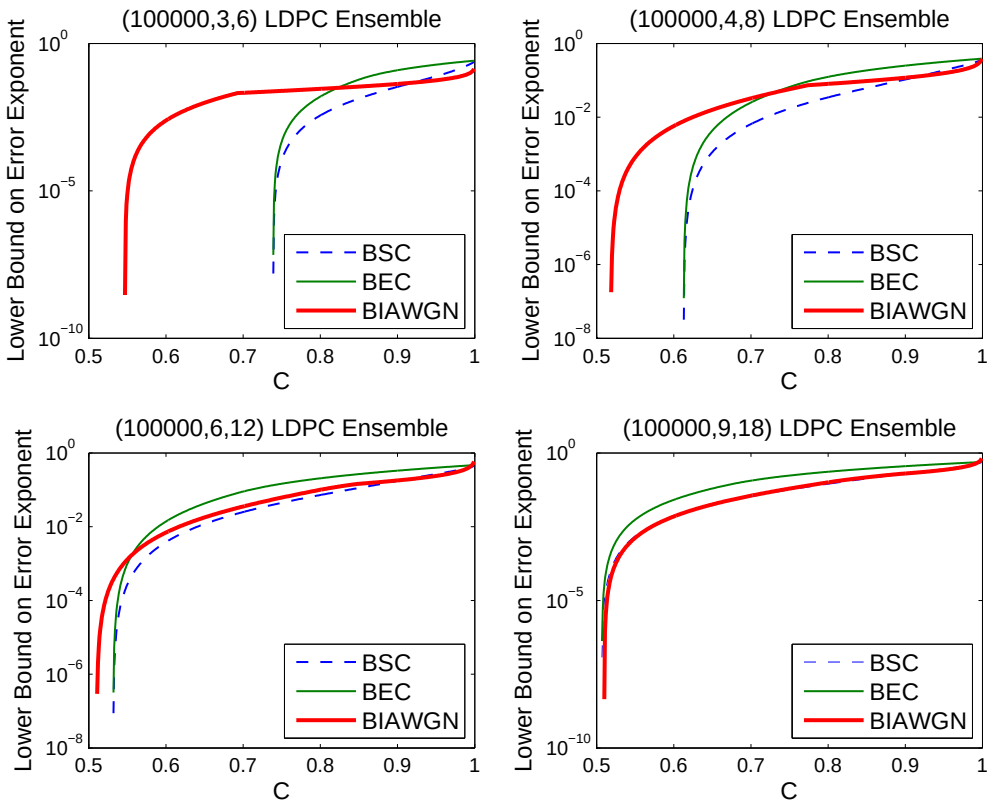

Fig. 2. Lower bounds on the error exponent for expurgated Gallager's LDPC code ensembles on various MBIOS channels. The results were computed for block length $n = 100000$, and for codes with constant design rate $1/2$ and increasing variable and check node degrees.

For short block lengths, we compare the lower bound for the expurgated $(6, 12)$ ensemble and block length $n = 1008$ computed using the exact distance spectrum [39] and the upper bound from [7, Chapter 2]. Fig. 3 shows the comparison. Clearly, the vanishing point of the error exponent is closer to capacity when computed using the exact distance spectrum. The calculation of the lower bound on the error exponent that uses the upper bound on the distance spectrum provides, however, a reasonable estimate of the lower bound on the error exponent that is calculated via the exact distance spectrum.

B. Universality of Punctured Regular LDPC Code Ensembles

The performance of punctured LDPC code ensembles under BP decoding was addressed extensively (see, e.g., [27, Chapter 5] and references therein). The potential performance of punctured LDPC code ensembles under ML decoding was studied, e.g., in [10] and [32]. These works show the remarkable performance of some punctured LDPC code ensembles for various channel models. In the following, we rely on [10], and consider the universality of some randomly punctured regular LDPC code ensembles under ML decoding over the set of equi-capacity MBIOS channels.

Consider a linear block code, which will be referred to as a mother code. By introducing the option of possibly puncturing various fractions of the code bits of the mother code, one generates a set of new linear block codes with some higher rates. The advantage of puncturing lies in the flexibility of the selected rates of the punctured codes, and in the ability to use the same decoder as for the mother code to decode all of these punctured codes. Specifically, by puncturing nq bits of a mother code of length n and rate R, one obtains a punctured code of length $n(1-q)$ and rate at most $\frac{R}{1-q}$. A lower rate occurs whenever at least two different codewords of the mother code are mapped to the same codeword after puncturing; this phenomenon is called *rate reduction* (see [10, Section III]).

In [10], the authors analyze the performance of punctured LDPC code ensembles under ML decoding. Specifically, they consider puncturing Gallager's ensemble of regular (n, j, k) LDPC codes, and provide conditions on the original ensemble (before puncturing) for asymptotically obtaining zero rate reduction with probability 1 as we let the block

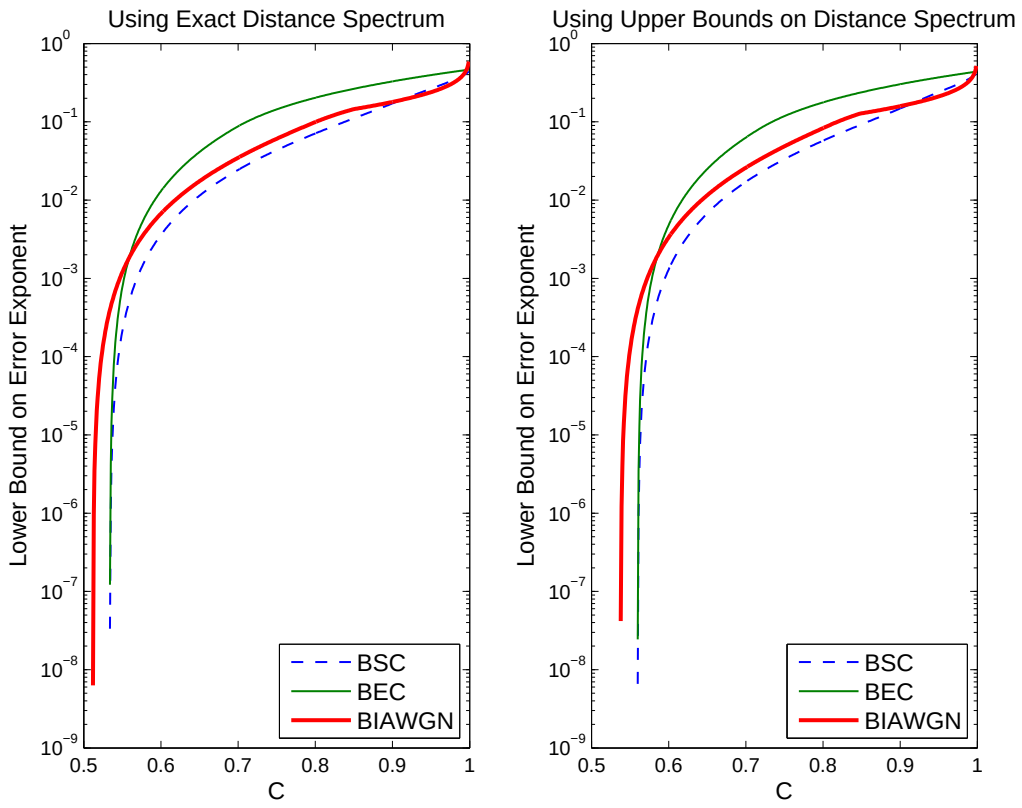

Fig. 3. A comparison of the lower bounds on the error exponent of some expurgated Gallager's LDPC code ensembles for various MBIOS channels. The results were computed for the expurgated $(6, 12)$ ensemble with block length $n = 1008$, using both the exact distance spectrum as found in [39] (left-hand plot) and the upper bound from [7, Chapter 2] (right-hand plot).

length n tend to infinity. Consider an ensemble whose design rate is R_{d}, then in the case where there is no rate reduction due to puncturing, the design rate of the punctured ensemble is $\frac{R_{\mathrm{d}}}{1-q}$. It is also shown in [10] that under the condition of zero rate reduction, if the original sequence of code ensembles achieves a fraction $1-\varepsilon$ of capacity (note that Theorem 5 ensures the existence of such a sequence), then so does the sequence of punctured code ensembles. This leads to the following theorem:

Theorem 6. [Universality of Punctured Regular LDPC Code Ensembles under ML Decoding for Equi-Capacity MBIOS Channels] Under ML decoding, punctured regular LDPC code ensembles can be made universal for the set of MBIOS channels that exhibit a given capacity. More explicitly, let $\varepsilon > 0$ (which can be set arbitrarily close to zero), and consider a sequence of regular (n, j, k) LDPC code ensembles whose design rate R_{d} forms a fraction of at least $1 - \varepsilon$ of the capacity C. Assume that this sequence achieves vanishing block error probability under ML decoding for the entire set of MBIOS channels $\mathcal{A}$ which exhibit a channel capacity C. Random puncturing of a fraction q of code bits from this sequence of ensembles produces a new sequence of punctured code ensembles with any desired design rate $R'_{\mathrm{d}} > R_{\mathrm{d}}$ with the following properties:

- It achieves vanishing block error probability under ML decoding over the entire set of equi-capacity MBIOS channels with capacity $C' = \frac{C}{1-q}$.
- It achieves a fraction of at least $1 - \varepsilon$ of the capacity C'.

Proof: From Theorem 5, there exists a sequence of regular LDPC code ensembles that universally achieves, under ML decoding, a fraction $1-\varepsilon$ of the channel capacity with vanishing block error probability for the entire set $\mathcal{A}$. The idea is to first construct such a sequence with a low enough design rate, and then increase the design rate via (random) puncturing to obtain the new universal code ensemble. More specifically, according to [10, Theorem 1], if the design rate of the mother ensemble is low enough, then the rate reduction due to puncturing is zero. Note that the proof of this theorem is based solely on the distance properties of the original (mother) code ensemble and the desired design rate. Since the proofs of [10, Theorems 2 and 3] rely only on the capacity of the MBIOS channel

and the condition for zero rate reduction, this implies that the new sequence of punctured LDPC code ensembles has vanishing block error probability under ML decoding over all MBIOS channels with capacity $C' = \frac{C}{1-q}$. Note that since

$$\frac{R'_{\mathrm{d}}}{1-\varepsilon} = \frac{R_{\mathrm{d}}}{(1-\varepsilon)(1-q)} \geq \frac{C}{1-q} = C'$$

then this new sequence of punctured LDPC code ensembles has a design rate that is at least a fraction $1-\varepsilon$ of C'. ■

V. Summary and Outlook

A. Summary

In this work we considered the universality of LDPC code ensembles under both BP and ML decoding. We have focused on obtaining closed-form analytical results, even though better performance can be obtained by numerical design of LDPC code ensembles (see, e.g., [3], [6], [23], [28] and also the discussion in Example 5 of this paper).

Under BP decoding we derived an analytical approach for designing LDPC code ensembles that achieve vanishing bit error probability over every channel in a family of MBIOS channels. This approach was applied to several families of MBIOS channels, such as the family of equi-capacity MBIOS channels. We also derived a necessary condition for a sequence of LDPC code ensembles to universally achieve vanishing bit error probability under BP decoding over an arbitrary set of MBIOS channels. This condition forms the basis for a linear programming universal upper bound on the achievable rate of LDPC code ensembles over a set of equi-capacity MBIOS channels. The analytical design method and the necessary condition above were used to derive bounds on the threshold of LDPC code ensembles under BP decoding. These bounds were compared to existing information-combining bounds on the threshold ([37]) in Table V). In some cases, our bounds were more informative.

For the ML decoding case, we used the results of [30] to show that Gallager's regular LDPC code ensembles can be made universally capacity achieving over the set of equi-capacity MBIOS channels (in the sense of vanishing block error probability). The result was extended to randomly punctured LDPC code ensembles as well, based on [10]. It is noted that the ML decoding results are an improvement over the BP decoding case, where the universally achievable fraction of capacity over the family of equi-capacity MBIOS channels depended on the channel capacity and could be as low as 69.3%.

B. Topics for Further Research

In this sub-section, we propose some directions for future research:

- The LP bounds derived in Section III are not tight in general, since the universal achievable gap to capacity does not always decrease for increasing values of d_{c} (contrary to the expected experimental behavior of optimized LDPC code ensembles under BP decoding). Finding some new constraints in these optimization problems may enhance the tightness of these bounds. Moreover, our bounds refer to fixed right-degree ensembles (note that typically LDPC code ensembles are designed to be right-regular or almost right-regular). Extending the bounds to the case where the parity-check degree is not fixed is also of interest.
- The Bhattacharyya parameter (B-parameter) for equi-capacity MBIOS channels can vary in a large range. As a result, the universal LDPC code ensembles designed in this work achieve, e.g., 75% of capacity if the channel capacity is 0.5 bit per channel use (see Fig. 1). Nonetheless, the fact that these ensembles are provably universal and are designed by simple analytical tools is important. Since, in practice, numerical optimizations enable to design LDPC code ensembles which universally achieve a larger fraction of capacity for some classes of equi-capacity MBIOS channels (see [28]), further analysis in this direction is of interest.
- The ideas of universality in this paper can be developed to consider other sets of communication channels (for example, the universality of LDPC code ensembles for the set of MBIOS channels with the same uncoded bit error probability can be considered in a similar approach to Section III-A via the use of density evolution and (5)).
- The approach for universality in Section III of this paper stems from the asymptotic analysis of BP decoding via density evolution; it has resulted in an analytical design of a universal decoder that is based on code design for a BEC. This universal LDPC code ensemble achieves vanishing bit error probability under BP decoding

but does not achieve full capacity when used over other channels in the family it was designed for, as Fig. 1 demonstrates. That said, one should not infer that this is the penalty of universality, as these results are merely an artifact of the approach presented here. Numerical evidence in [6] and [23] suggests that better results are possible (see also Example 5 in this paper). One possible approach to obtain better analytical results may rely on analytic properties of GEXIT charts [19], instead of the suggested approach in this paper that relies on density evolution for the BEC as a starting point for the analysis. Another possible approach may be to investigate universal LDPC code ensemble design under other suboptimal decoding methods for LDPC codes (e.g., a study of the universality of LDPC code ensembles under LP decoding).

- Although ML decoding is prohibitively complex for codes of large block lengths, the fact that (regular) LDPC code ensembles are capacity-achieving under ML decoding for the set of equi-capacity MBIOS channels is interesting (see Theorems 5 and 6). As a continuation of the previous item, it would be interesting to investigate the universality issue for some near-ML decoding algorithms that provide a better tradeoff between performance and complexity than the ML decoding algorithm.
- The analysis in this paper can be easily adapted to other families of code ensembles defined on graphs, e.g., to irregular repeat-accumulate codes (see [36]).

APPENDIX A
PROOF OF LEMMA 1

In the following, we prove the monotonicity of μ_1 (see (12)) over the interval $[0,1)$, and then calculate the limits of $\mu_1(C)$ as the channel capacity C tends either to zero or 1 bit per channel use.

Let $x \triangleq h_2^{-1}(1-C)$, then we get from (12) that

$$\mu_1(C) = \frac{1-\sqrt{4x(1-x)}}{1-h_2(x)}.$$

We note that μ_1, as a function of x, monotonically decreases when $0 \leq x \leq \frac{1}{2}$. This is readily seen by taking the derivative of μ_1 with respect to x, which remains negative when $0 \leq x \leq \frac{1}{2}$. The substitution of a Taylor series expansion of $h_2(x)$ about $x = \frac{1}{2}$ (see (39)) in the denominator gives

$$\begin{aligned}
\mu_1(C) &= \frac{1-\sqrt{1-(1-2x)^2}}{\frac{1}{2\ln 2}\sum_{q=1}^{\infty}\frac{(1-2x)^{2q}}{q(2q-1)}} \\
&= \frac{(1-2x)^2}{1+\sqrt{1-(1-2x)^2}}\ \frac{2\ln 2}{\sum_{q=1}^{\infty}\frac{(1-2x)^{2q}}{q(2q-1)}} \\
&= \frac{2\ln 2}{1+\sqrt{1-(1-2x)^2}}\ \frac{1}{\sum_{q=1}^{\infty}\frac{(1-2x)^{2(q-1)}}{q(2q-1)}}\,.
\end{aligned}$$

If C is increased from 0 to 1, then x, which was defined above as $x \triangleq h_2^{-1}(1-C)$, decreases from $\frac{1}{2}$ to 0 and therefore $\mu_1(C)$ is increasing with C, and

$$\lim_{C\to 1}\mu_1(C) = 1.$$

On the other hand, the limit of $\mu_1(C)$ when we let the capacity tend to zero is equal to

$$\begin{aligned}
\lim_{C\to 0}\mu_1(C) &= \lim_{x\to\frac{1}{2}}\frac{2\ln 2}{1+\sqrt{1-(1-2x)^2}}\ \frac{1}{\sum_{q=1}^{\infty}\frac{(1-2x)^{2(q-1)}}{q(2q-1)}} \\
&= \ln 2\ \lim_{x\to\frac{1}{2}}\frac{1}{\sum_{q=1}^{\infty}\frac{(1-2x)^{2(q-1)}}{q(2q-1)}} \\
&= \ln 2.
\end{aligned}$$

This completes the proof of Lemma 1.

APPENDIX B
EXTENSION OF (15) FOR A GENERAL RIGHT-DEGREE DISTRIBUTION

The condition in (15) is stated for a right-regular channel. For a general right-degree distribution,

$$\rho(x) = \sum_i \rho_i x^{i-1}.$$

This condition can be readily extended to the case at hand, although it takes a more involved form. As in (19) we begin with (6) to obtain

$$\begin{aligned} x_l &\triangleq \mathcal{B}(a_l) \\ &\overset{\text{(a)}}{=} \mathcal{B}(a_0)\,\mathcal{B}\bigg(\lambda\Big(\Gamma^{-1}\Big(\rho(\Gamma(a_{l-1}))\Big)\Big)\bigg) \\ &\overset{\text{(b)}}{=} \mathcal{B}(a_0)\,\lambda\bigg(\mathcal{B}\Big(\Gamma^{-1}\Big(\rho(\Gamma(a_{l-1}))\Big)\Big)\bigg) \\ &\overset{\text{(c)}}{=} \mathcal{B}(a_0)\,\lambda\left(\sum_i \rho_i \mathcal{B}\left(a_{l-1}^{\circledast i-1}\right)\right) \\ &\overset{\text{(d)}}{\geq} \mathcal{B}(a_0)\,\lambda\bigg(\sum_i \rho_i \sqrt{1-\big(1-\mathcal{B}(a_{l-1})^2\big)^{i-1}}\bigg), \end{aligned} \tag{45}$$

where equality (a) follows from the recursive density evolution equation in (6) and the property in (20) of the B-functional, equalities (b) and (c) follow from the linearity of the convolution operator and of the B-functional (see (21)), and inequality (d) follows from (18).

The extension of (15) for a general LDPC code ensemble readily follows by replacing z_l in (22) with the sequence

$$z'_l = \mathcal{B}(a_0)\,\lambda\bigg(\sum_i \rho_i \sqrt{1-\big(1-\mathcal{B}(z'_{l-1})^2\big)^{i-1}}\bigg),$$

with initial condition $z'_0 = \mathcal{B}(a_0)$.

Thus, the generalized version of (15) assumes the form

$$B\,\lambda\bigg(\sum_i \rho_i \sqrt{1-\big(1-x^2\big)^{i-1}}\bigg) \leq x, \quad \forall x \in (0, B]. \tag{46}$$

APPENDIX C
PROOF OF LEMMA 3

Let us consider the sequence

$$z_l = z_0 \lambda\big(\sqrt{1-\rho(1-z_{l-1}^2)}\big), \quad l = 1, 2, \ldots$$

for $z_0 < B_1(\lambda, \rho)$ where

$$B_1(\lambda, \rho) \triangleq \inf_{x \in (0,1]} \frac{x}{\lambda\big(\sqrt{1-\rho(1-x^2)}\big)}$$

is introduced in (30). By substituting $x = 1$ on the right-hand side above, it follows readily that $B_1(\lambda, \rho) \leq 1$ and therefore $z_0 < 1$. In the following, it is proved by induction that the sequence is monotonically decreasing and bounded between 0 and 1: let us assume that $0 \leq z_{l-1} < 1$ holds for a specific $l \geq 1$, then

$$\begin{aligned} z_l &= z_0 \lambda\big(\sqrt{1-\rho(1-z_{l-1}^2)}\big) \\ &\leq B_1(\lambda, \rho)\ \lambda\big(\sqrt{1-\rho(1-z_{l-1}^2)}\big) \\ &\leq z_{l-1} \end{aligned}$$

where the last inequality follows from the definition of B_1 and the above assumption for z_{l-1}. It therefore follows by induction that the sequence $\{z_l\}$ is monotonically decreasing and bounded between 0 and 1, hence it is a convergent sequence. Let $z^* \in [0,1]$ denote the limit of this sequence, then due to the continuity of λ and ρ over the interval $[0,1]$, it follows (by letting l tend to infinity in the recursive equation for the sequence $\{z_l\}$) that the limit $z = z^*$ satisfies the equation

$$z = z_0 \lambda\big(\sqrt{1-\rho(1-z^2)}\big).$$

For $z \in (0,1]$

$$\begin{aligned} & z_0 < B_1 \leq \frac{z}{\lambda\big(\sqrt{1-\rho(1-z^2)}\big)} \\ & \Rightarrow z_0 \lambda\big(\sqrt{1-\rho(1-z^2)}\big) < z \end{aligned}$$

and therefore the limit z should be necessarily zero for the case where the initial value z_0 is less than $B_1(\lambda,\rho)$.

For the proof of the second part of the lemma, we consider the case where $B_1(\lambda,\rho) < z_0 \leq 1$. From the way B_1 is defined in (30), it follows that the set

$$\mathcal{F}_{z_0} \triangleq \left\{ x \in (0,1] : \ \frac{x}{\lambda\big(\sqrt{1-\rho(1-x^2)}\big)} \leq z_0 \right\} \tag{47}$$

is non-empty. Let $x(z_0)$ designate the maximal value of this set (note that $0 < x(z_0) \leq 1$).

Let us define the function $g(u,v) \triangleq u\lambda\big(\sqrt{1-\rho(1-v^2)}\big)$ over the square $\{(u,v) : 0 \leq u \leq 1, 0 \leq v \leq 1\}$. Note that the function g is monotonically increasing in its two variables; the monotonicity in u is due to its linearity in u and the non-negativity of λ; and the monotonicity in v is due to the monotonicity of the degree distribution λ and ρ over the interval $[0,1]$ and since they are mapped to the same interval. We show in the following, by induction, that $z_l \in [x(z_0), z_0]$ for every integer $l \geq 0$. For $l = 0$, the inequality $x(z_0) \leq z_0 \leq 1$ holds since for $x \in (z_0, 1]$

$$\frac{x}{\lambda\big(\sqrt{1-\rho(1-x^2)}\big)} \geq x > z_0.$$

Let us assume that $z_{l-1} \in [x(z_0), z_0]$ for a specific $l \geq 1$ then

$$\begin{aligned} z_l &= g(z_0, z_{l-1}) \\ &\overset{\text{(a)}}{\geq} g(z_0, x(z_0)) \\ &= z_0 \lambda\big(\sqrt{1-\rho(1-x(z_0)^2)}\big) \\ &\overset{\text{(b)}}{\geq} x(z_0) \end{aligned}$$

where inequality (a) is due to the monotonicity of g, and inequality (b) follows from the way $x(z_0)$ is defined above (or, more generally, this inequality holds for every $x \in \mathcal{F}_B$ where the set $\mathcal{F}_{z_0}$ is defined in (47)). Also, from the above assumption for z_{l-1}

$$\begin{aligned} z_l &= g(z_0, z_{l-1}) \\ &\leq g(z_0, 1) \\ &= z_0 \end{aligned}$$

and therefore, it follows by induction that

$$x(z_0) \leq z_l \leq z_0, \quad l = 0, 1, \ldots$$

and the sequence $\{z_l\}$ is bounded away from zero (since $x(z_0) > 0$). This completes the proof of Lemma 3.

APPENDIX D
PROOF OF THE INEQUALITY IN REMARK 6

From the definitions of B_0 and B_1 in (29) and (30), respectively, in order to prove that $B_1(\lambda,\rho) \geq B_0(\lambda,\rho)$, it is sufficient to show that

$$\lambda\big(\sqrt{1-\rho(1-x^2)}\big) \leq \lambda\big(1-\rho(1-x)\big), \quad \forall\, x \in [0,1].$$

Since $\lambda(0)=0$, $\lambda(1)=1$, and λ is monotonic increasing over the interval [0, 1], then this inequality is equivalent to

$$\sqrt{1-\rho(1-x^2)} \leq 1-\rho(1-x), \quad \forall\, x \in [0,1].$$

By squaring and rearranging terms, we need to prove that

$$h(x) \triangleq \rho(1-x^2)+\rho^2(1-x)-2\rho(1-x) \geq 0, \quad \forall\, x \in [0,1].$$

Note that h is zero at the endpoints of this interval (since $\rho(0)=0$ and $\rho(1)=1$). From the assumption of right-regularity then $\rho(x)=x^{d_{\mathrm{c}}-1}$. Let $\gamma \triangleq d_{\mathrm{c}}-1$ (where $\gamma \geq 1$), then

$$\begin{aligned} h(x) &= (1-x^2)^{\gamma}+(1-x)^{2\gamma}-2(1-x)^{\gamma} \\ &= 2(1-x)^{\gamma}\left[\frac{(1+x)^{\gamma}+(1-x)^{\gamma}}{2}-1\right] \\ &\geq 0 \end{aligned}$$

where the last transition follows from the non-negativity of both terms over the interval $x \in [0,1]$ (the second term is non-negative due to the convexity of the function $f(x)=x^{\gamma}$ for $x \geq 0$ (note that $\gamma \geq 1$)). This completes the proof of the inequality in Remark 6.

Acknowledgment

The authors would like to thank the associate editor, Pascal Vontobel, and the anonymous reviewers for their meticulous reading of the submitted version of this manuscript, and detailed comments which helped in general to improve the clarity of the presentation. A stimulating discussion with Henry D. Pfister at an early stage of this work is also acknowledged.

REFERENCES

[1] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. on Information Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[2] D. Burshtein and G. Miller, "Bounds on the performance of belief propagation decoding," *IEEE Trans. on Information Theory*, vol. 48, no. 1, pp. 112–122, January 2002.

[3] S. Y. Chung, *On the Construction of Some Capacity-Approaching Coding Schemes*, PhD dissertation, MIT, Department of Electrical Engineering and Computer Science, 2000. [Online]. Available: http://wicl.kaist.ac.kr/pdf/sychungphdthesis.pdf.

[4] S. Y. Chung, G. D. Forney, Jr., T. 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, pp. 58–60, February 2001.

[5] D. Duyck, M.H. Azmi, J. Yuan, J.J. Boutros and M. Moeneclaey, "Universal LDPC codes for Cooperative Communications," *Proceedings 6th International Symposium on Turbo Codes and Iterative Information Processing*, pp. 83–87, Brest, France, September 2010

[6] 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.

[7] R. G. Gallager, *Low-Density Parity-Check Codes*, Cambridge, MA, USA, MIT Press, 1963. [Online]. Available: http://web.mit.edu/gallager/www/pages/ldpc.pdf.

[8] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming*, June 2009. [Online]. Available: http://stanford.edu/~boyd/cvx.

[9] C. H. Hsu, *Design and Analysis of Capacity-Achieving Codes and Optimal Receivers with Low Complexity,* PhD dissertation, University of Michigan, USA, Department of Electrical Engineering, 2006. [Online]. Available: http://www.eecs.umich.edu/~anastas/docs/chunhao_thesis.pdf.

[10] 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.

[11] C. H. Hsu and A. Anastasopoulos, "Capacity-achieving codes with bounded graphical complexity and maximum-likelihood decoding," *IEEE Trans. on Information Theory*, vol. 56, no. 3, pp. 992–1006, March 2010.

[12] H. Jin and T. J. Richardson, "Block error iterative decoding capacity for LDPC codes," *Proceedings 2005 IEEE International Symposium on Information Theory (ISIT 2005)*, pp. 52–56, Adelaide, Australia, September 2005.

[13] A. Khandekar, *Graph-based codes and iterative decoding*, PhD dissertation, Caltech, Pasadena, CA, USA, Department of Engineering and Applied Science, 2002. [Online]. Available: http://resolver.caltech.edu/CaltechETD: etd-06202002-170522.

[14] I. Land and J. Huber, "Information combining," *Foundations and Trends in Communications and Information Theory*, vol. 3, no. 3, pp. 227–330, November 2006.

[15] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. on Information Theory*, vol. 44, no. 6, pp. 2148–2177, October 1998.

[16] *LDPCopt – A Fast and Accurate Degree Distribution Optimizer for LDPC Code Ensembles*, [Online]. Available: http: //ipgdemos.epfl.ch/ldpcopt/

[17] M. Lentmaier, D. V. Truhachev, K. Sh. Zigangirov, and D. J. Costello, "An analysis of the block error probability performance of iterative decoding," *IEEE Trans. on Information Theory*, vol. 51, no. 11, pp. 3834–3855, November 2005.

[18] 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.

[19] C. Méasson, A. Montanari, T. J. Richardson and R. L. Urbanke, "The generalized area theorem and some of its consequences," *IEEE Trans. on Information Theory*, vol. 55, no. 11, pp. 4793–4821, November 2009.

[20] G. Miller and D. Burshtein, "Bounds on the maximum-likelihood decoding error probability of LDPC codes," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2696–2710, November 2001.

[21] S. Miyake and M. Maruyama, "Construction of universal codes using LDPC matrices and their error exponents," *IEICE Trans. Fundamentals*, vol. E90-A, no. 9, pp. 1830–1839, September 2007.

[22] 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.

[23] 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.

[24] 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.

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

[26] T. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[27] W. E. Ryan and S. Lin, *Channel Codes: Classical and Modern*, Cambridge University Press, 2009.

[28] A. Sanaei, M. Ramezani and M. Ardakani, "Identical-capacity channel decomposition for design of universal LDPC codes," *IEEE Trans. on Communications*, vol. 57, no. 7, pp. 1972–1981, July 2009.

[29] 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, June 2006.

[30] 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.

[31] I. Sason and R. Urbanke, "Complexity versus performance of capacity-achieving irregular repeat-accumulate codes on the erasure channel," *IEEE Trans. on Information Theory*, vol. 50, no. 6, pp. 1247–1256, June 2004.

[32] 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.

[33] I. Sason, "On universal properties of capacity-approaching LDPC code ensembles," *IEEE Trans. on Information Theory*, vol. 55, no. 7, pp. 2956–2990, July 2009.

[34] 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.

[35] A. Shokrollahi, "Capacity-achieving sequences," *IMA Volume in Mathematics and its Applications*, vol. 123, pp. 153–166, 2000.

[36] B. Shuval, *On Universal LDPC Code Ensembles over Memoryless Symmetric Channels*, Master Thesis, Technion–Israel Institute of Technology, Haifa, Israel, February 2011.

[37] I. Sutskover, S. Shamai and J. Ziv, "Extremes of information combining," *IEEE Trans. on Information Theory*, vol. 51, no. 4, pp. 1313–1325, April 2005.

[38] 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.

[39] S. Tong, "Tangential-sphere bounds on the ensemble performance of ML decoded Gallager codes via their exact ensemble

distance spectrum," *Proceeding of the 2008 IEEE International Conference on Communications (ICC 2008)*, pp. 1150–1154, Beijing, China, May 2008.

[40] M. Twitto and I. Sason, "On the error exponents of improved tangential-sphere bounds," *IEEE Trans. on Information Theory*, vol. 53, no. 3, pp. 1196–1210, March 2007.

[41] P. O. Vontobel, "A factor-graph approach to universal decoding," *Proceedings 44th Annual Allerton Conference on Communication, Control, and Computing*, pp. 23–30, Monticello, Illinois, USA, September 2006.

[42] C. C. Wang, S.R. Kulkarni, and H.V. Poor, "Finite-dimensional bounds on $\mathbb{Z}_m$ and binary LDPC codes with Belief Propagation decoders," *IEEE Trans. on Information Theory*, vol. 53, no. 1, pp. 56–81, January 2007.

[43] 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.

[44] A. Yedla, H. D. Pfister and K. R. Narayanan, "Can iterative decoding for erasure decoded sources be universal?," *Proceedings 47th Annual Allerton Conference on Communication, Control and Computing*, pp. 408–415, Monticello, Illinois, USA, September 2009.