Successive Maximization for Systematic Design of Universally Capacity Approaching Rate-Compatible Sequences of LDPC Code Ensembles over Binary-Input Output-Symmetric Memoryless Channels

Hamid Saeedi, *Member, IEEE,* Hossein Pishro-Nik, *Member, IEEE,* and Amir H. Banihashemi, *Senior Member, IEEE*

***Abstract*—A systematic construction of capacity achieving low-density parity-check(LDPC) code ensemble sequences over the Binary Erasure Channel (BEC) has been proposed by Saeedi *et al.* based on a method, here referred to as Successive Maximization (SM). In SM, the fraction of degree-i nodes are successively maximized starting from $i = 2$ with the constraint that the ensemble remains convergent over the channel. In this paper, we propose SM to design universally capacity approaching rate-compatible LDPC code ensemble sequences over the general class of Binary-Input Output-Symmetric Memoryless (BIOSM) channels. This is achieved by first generalizing the SM method to other BIOSM channels to design a sequence of capacity approaching ensembles called the parent sequence. The SM principle is then applied to each ensemble within the parent sequence, this time to design rate-compatible puncturing schemes. As part of our results, we extend the stability condition which was previously derived for degree-2 variable nodes to other variable node degrees as well as to the case of rate-compatible codes. Consequently, we prove that using the SM principle, one is able to design universally capacity achieving rate-compatible LDPC code ensemble sequences over the BEC. Unlike the previous results in the literature, the proposed SM approach is naturally extendable to other BIOSM channels. The performance of the rate-compatible schemes designed based on our method is comparable to those designed by optimization.**

***Index Terms*—Low-density parity-check (LDPC) codes, rate-compatible LDPC codes, capacity-achieving codes, capacity-approaching codes, systematic design of LDPC codes, stability condition, binary erasure channel (BEC), binary symmetric channel (BSC), additive white Gaussian noise (AWGN) channel.**

Paper approved by T.-K. Truong, the Editor for Coding Theory and Techniques of the IEEE Communications Society. Manuscript received July 14, 2010; revised December 2, 2010.

A preliminary version of this work has been partly presented at the Information Theory Workshop, June 2009, Volos, Greece, and in part at the Allerton Conference on Communication, Control, and Computing, Sep. 2009.

This work was supported by the National Science Foundation under grants ECS-0636569 and CCF-0830614.

H. Saeedi was with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA, USA. He is now with the Wireless Innovations Lab, Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran (e-mail: hsaeedi@ieee.org).

H. Pishro-Nik is with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA, USA (e-mail: pishro@ecs.umass.edu).

A. H. Banihashemi is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada (e-mail: ahashemi@sce.carleton.ca).

Digital Object Identifier 10.1109/TCOMM.2011.060911.100409

I. Introduction

LOW-DENSITY Parity-Check (LDPC) codes have received much attention in the past decade. During this period there have been great achievements in the area of designing LDPC code ensembles with Belief Propagation (BP) decoding which exhibit an asymptotic performance practically close to the capacity over different types of channels, including the general class of Binary-Input Output-Symmetric Memoryless (BIOSM) channels [1]-[11]. In particular, for the Binary Erasure Channel (BEC), the performance analysis and code design have been addressed in both the asymptotic regime [3]-[9] and for finite block lengths [1], [2]. In [3]–[5], Shokrollahi *et al.* proposed a scheme to design sequences of LDPC code ensembles over the BEC, whose performance is proved to achieve the capacity for sufficiently large average check and variable node degrees. A more general category of capacity achieving sequences over the BEC were proposed in [12]–[14]. Construction and analysis of capacity achieving ensemble sequences of codes defined on graphs have also been studied in [6]–[9] for the BEC. A sequence of degree distributions with rate R is said to be *capacity achieving* over the BEC if the thresholds of the ensembles can be made arbitrarily close to $1 - R$, the capacity upper bound over the BEC, as the average check and variable node degrees tend to infinity. For BIOSM channels, it is easier to consider ensembles for a given channel parameter instead of a given rate. The results however are easily extendable to the case of fixed rate ensembles. We call a sequence of degree distributions capacity achieving over a BIOSM channel, if the rate of the ensembles within the sequence can be made arbitrarily close to the channel capacity while maintaining the reliable communication. The design of provably capacity achieving sequences over general BIOSM channels is still an open problem.

Another important problem of interest in LDPC codes is to design rate-compatible LDPC code schemes. In such a

0090-6778/11$25.00 © 2011 IEEE

scheme, starting from a given primary ensemble called the *parent code*, we are interested in obtaining a set of codes with higher transmission rates, which can provide reliable transmission when the channel condition improves, by puncturing the parent code. For rate-compatibility, the design must be such that for two consecutive rates, the code with the higher rate can be constructed by puncturing the code with the lower rate. Starting from a parent code with performance close to capacity, the important challenge in a rate-compatible design is to also keep the performance of the punctured codes close to the capacity. More specifically, if the parent code is chosen from a capacity achieving sequence, all punctured codes should be capacity achieving as average check node degree increases. To formulate the problem mathematically, imagine a parent code with rate R^n from a capacity achieving sequence which can provide reliable transmission over a channel with parameter θ^0. Our aim is to provide reliable transmission over a set of channels with parameters θ^j, $j = 1, ..., J$, while increasing the rate by puncturing the parent code in a rate-compatible fashion. For each θ^j, $j = 1, ..., J$, we need to choose a puncturing pattern that maximizes the corresponding reliable transmission rate $R^{n,j}$. Let $c(\theta^j)$ denote the capacity of the channel with parameter θ^j, and assume that $\theta^i < \theta^j$ and $c(\theta^i) > c(\theta^j)$ for $i > j$. We call a rate-compatible scheme *universally capacity achieving*, if $\lim_{n\to\infty} R^{n,j} = c(\theta^j)$ for $j = 1, .., J$. Analysis and design of rate-compatible LDPC codes have been addressed asymptotically in [15]-[19] and for finite block lengths in [20]-[23]. It is worth mentioning here that Raptor codes [24] can also achieve the capacity of the BEC at several rates but in a different framework than puncturing.

Unlike the BEC for which almost all aspects of conventional (unpunctured) and rate-compatible codes have been analytically investigated, for the general family of BIOSM channels, the contributions are mostly based on numerical methods and optimization. This usually provides little insight into the design method. In this respect, a fundamental open problem is to prove the existence of capacity achieving sequences of LDPC codes with BP decoding over BIOSM channels as well as to systematically construct such sequences. This can be seen as a sub-problem as well as a building block for the more general problem of designing universally capacity achieving rate-compatible LDPC coding schemes. In [10], it has been shown that capacity approaching LDPC codes over BIOSM channels can be designed using optimization[1]. A less complex optimization-based design method over the Binary-Input Additive White Gaussian Noise (BIAWGN) channel has been proposed in [25]. Several important analytical properties including the so-called stability condition have been proven for BIOSM channels in [10], [11]. For rate-compatible codes, it has been shown in [15], [17] that there is an upper bound on the puncturing ratio of LDPC codes over BIOSM channels, above which the code can not provide reliable transmission for any channel parameter. Moreover, it has been shown that over the BEC, the random puncturing maintains the ratio of rate to capacity at the same value as that of the parent code. Several bounds on the performance of punctured LDPC codes have been derived in [19]. For the case of maximum-likelihood decoding, capacity achieving codes have been designed based on puncturing in [18]. Among the results on the optimization-based design of rate-compatible codes over BIOSM channels, we can mention [16] for the asymptotic regime and [20]–[22] for finite block lengths.

In this paper, we systematically design sequences of universally capacity approaching rate-compatible LDPC code ensembles over BIOSM channels. We then provide some evidence suggesting that the designed sequences could in fact be universally capacity achieving. Starting from the unpunctured case, we extend some of the properties of capacity achieving sequences over the BEC [12], [13], to BIOSM channels. Among such properties, only the stability condition [10] has been shown to be extendable to BIOSM channels other than the BEC. We will analyze the case where the stability condition is satisfied with equality, i.e, the fraction of degree 2 edges (λ_2) is set equal to its upper bound, and show that this imposes an upper bound on the fraction of degree 3 edges (λ_3). Using a similar approach for the other degrees, we propose *Successive Maximization* (SM) of λ_i values as a systematic approach to design a sequence of LDPC code ensembles with performance approaching the capacity as the average check node degree increases. For the rate-compatible LDPC codes over BIOSM channels, we first prove a property similar to the stability condition. We show that for a given parent code, there is an upper bound on the fraction of punctured degree-2 variable nodes (Π_2) above which the probability of error of the punctured code is bounded away from zero and below which the probability of error tends to zero if it is made sufficiently small. We then consider the special case of the BEC and show that similar upper bounds can be obtained for variable nodes of all degrees in addition to degree-2 nodes. Using such upper bounds, we prove that applying the SM principle results in a universally capacity achieving rate-compatible scheme over the BEC. Moreover, for such a scheme, if puncturing fractions $\Pi_i^{n,j}$ are used to puncture the parent sequence (λ^n, ρ^n) over the channel with parameter θ^j, where i is the variable node degree, the values of $\Pi_i^{n,j}$ are independent of n. This result is consistent with the one obtained in [15], [17] based on a completely different approach. We then extend the results for the BEC to general BIOSM channels, and show that the SM principle can be applied to puncturing fractions of variable nodes to systematically design a coding scheme whose performance universally approaches the capacity in a rate-compatible fashion. This proposes a significantly different approach than the existing optimization-based methods in the literature. Our numerical results indicate that if the parent ensemble is chosen from the capacity approaching sequences designed based on the SM principle, the performance of the resulting rate-compatible schemes is similar to that of the existing optimization-based results in the literature. Moreover, we show that for a sequence of parent code ensembles (λ^n, ρ^n) designed based on the SM principle, the values of puncturing fractions $\Pi_i^{n,j}$ for degree 2 variable nodes ($i = 2$) are independent from the parent ensemble (n) and only depend

[1]We distinguish between "capacity approaching" and "capacity achieving" sequences. The former term is used when the performance of the ensemble sequence can be shown (probably numerically) to approach capacity without any guarantee to achieve it. The latter term is used if the performance provably tends to capacity as the average node degrees tend to infinity.

on the original channel parameter (θ^0) and the one for which the puncturing pattern is designed (θ^j). Our numerical results suggest that this property may in fact hold for other values of i. The importance of this property is that for a given channel parameter θ^j, the computed values of Π_i can universally be applied to any ensemble designed based on the SM method for a given original channel parameter θ^0 with an arbitrary check node distribution.

The paper organization is as follows. The next section is devoted to notations and some definitions. In Section III, after a short review on the construction of capacity achieving sequences over the BEC, we explain the successive maximization approach to devise capacity approaching sequences for other channels. Section IV provides some analytical results related to the proposed approach. In Section V, we focus on the puncturing of a given ensemble within a sequence that is designed based on the SM methodology. We also provide some properties of rate-compatible codes for the BEC and BIOSM channels. Moreover, we show that a similar SM principle to that of Section III can be used to devise a universally capacity approaching rate-compatible scheme. In Section VI, we show examples of our designs and Section VII concludes the paper. The details of the design algorithm and some numerical issues are discussed in the appendix.

II. Definitions and Notations

In this section we present some definitions and properties which will be frequently used throughout the paper. We mainly follow the notations and definitions of [11], [17]. As our focus is on symmetric channels and a BP decoder, throughout the paper, without loss of generality, we assume that the all-one code word is transmitted. Moreover, we assume that the messages in the BP algorithm are in the log-likelihood ratio domain. We represent a (λ, ρ) LDPC code ensemble with its edge-based check and variable node degree distributions as $\rho(x) = \sum_{i=2}^{D_c} \rho_i x^{i-1}$ and $\lambda(x) = \sum_{i=2}^{D_v} \lambda_i x^{i-1}$, with constraints $\sum_{i=2}^{D_c} \rho_i = 1$ and $\sum_{i=2}^{D_v} \lambda_i = 1$, where the coefficient of x^i represents the fraction of edges connected to the nodes of degree $i+1$, and D_v and D_c represent the maximum variable node degree and the maximum check node degree, respectively. Average check node and variable node degrees are given by: $\overline{d}_c = 1/(\sum_{i=2}^{D_c} \rho_i/i)$ and $\overline{d}_v = 1/(\sum_{i=2}^{D_v} \lambda_i/i)$, respectively. The code rate R satisfies

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

We also define node-based degree distributions as $\overline{\rho}(x) = \sum_{i=2}^{D_c} \overline{\rho_i} x^{i-1}$ and $\overline{\lambda}(x) = \sum_{i=2}^{D_v} \overline{\lambda_i} x^{i-1}$, with constraints $\sum_{i=2}^{D_c} \overline{\rho_i} = 1$ and $\sum_{i=2}^{D_v} \overline{\lambda_i} = 1$, where the coefficient of x^i represents the fraction of nodes having degree $i+1$. We represent a BIOSM channel with parameter θ by $C(\theta)$ and define $c(\theta)$ as the Shannon capacity of that channel. We also assume that the channel is physically degraded when θ increases. For a sequence of degree distributions $(\lambda^n(x), \rho^n(x))$, λ_i^n and ρ_i^n indicate the ith coefficient of the nth member of the sequence for variable node and check node degree distributions, respectively. Notation T_i is used for the coefficient of x^{i-1} in the Taylor series expansion of $1 - \rho^{-1}(1-x)$ around $x = 0$, i.e., we have $1 - \rho^{-1}(1-x) = \sum_{i=2}^{\infty} T_i x^{i-1}$ (note that $T_1 = 0$). Similar to [5], we limit our discussions to check node degree distributions for which T_i's are positive. For example, check regular ensembles exhibit such a property.

Consider now the density evolution in the belief propagation algorithm for the channel $C(\theta)$, where we track the evolution of the initial channel density P_0 throughout iterations in the asymptotic regime, where the block length tends to infinity.[2] Based on [10], [11], Q_l, the probability density function (density) of the outgoing message from check nodes at iteration l can be written as $Q_l = \Gamma^{-1}\rho(\Gamma(P_{l-1}))$, where P_{l-1} is the density of the message from iteration $l-1$ entering the check nodes and Γ is the check node operator defined in [10], [11]. Also, P_l, the outgoing density from variable nodes at iteration l, can be written as $P_l = P_0 \otimes \lambda(Q_l)$, where $\otimes$ is the convolution operation, and the power of a density in variable node and check node degree distributions has been defined in [11]. Note that there is a one-to-one correspondence between the density P_0 and the channel parameter θ.

Let P be a symmetric density (as defined in [10]). For such a density, parameters $\mathbb{P}(P)$ and $\mathfrak{P}(P)$ are defined by:

$$\mathbb{P}(P) = 0.5 \int_{-\infty}^{\infty} P(x) e^{-(|x/2| + x/2)} dx,$$

and

$$\mathfrak{P}(P) = \int_{-\infty}^{\infty} P(x) e^{-(x/2)} dx.$$

Parameter $\mathbb{P}(P)$ is the probability that the random variable with density P is negative. In the setting of this paper, $\mathbb{P}(P)$ is the probability of error for the message with density P. Parameter $\mathfrak{P}(P)$ is called the Bhattacharyya constant. For any density P, $\mathfrak{P}(P)$ tends to zero if and only if (iff) $\mathbb{P}(P)$ tends to zero. Let $p_l = \mathfrak{P}(P_l)$ and $q_l = \mathfrak{P}(Q_l)$. Corresponding to the density evolution equations, we then have the following relationship [11]:

$$p_l \leq \mathfrak{P}(P_0)\lambda(1 - \rho(1 - p_{l-1})). \tag{2}$$

It is important to note that for the BEC, (2) is satisfied with equality. Moreover, $\mathfrak{P}(P_0)$ is equal to the channel erasure probability. The following important inequalities also hold [11]:

$$2\mathbb{P}(P) \leq \mathfrak{P}(P) \leq 2\sqrt{\mathbb{P}(P)(1 - \mathbb{P}(P))}. \tag{3}$$

A given degree distribution (λ, ρ) is called *stable* iff there exists $\xi > 0$ such that if $\mathbb{P}(P_l) < \xi$ then $\lim_{l\to\infty} \mathbb{P}(P_l) = 0$. In that respect, it is proven in [10], [11] that if $\lambda'(0)\rho'(1) > 1/\mathfrak{P}(P_0)$, then $\mathbb{P}(P_l)$ is bounded away form zero for every l and if $\lambda'(0)\rho'(1) < 1/\mathfrak{P}(P_0)$, then the ensemble is stable.

We call an ensemble (λ, ρ) convergent over $C(\theta)$, if starting from the initial density P_0, $\lim_{l\to\infty} \mathbb{P}(P_l) = 0$. The *threshold* of an ensemble over $C(\theta)$ is the supremum value of θ for which the ensemble is convergent. A sequence of degree distributions (λ^n, ρ^n) is called capacity achieving over a BIOSM channel $C(\theta)$, if the corresponding ensembles are convergent over $C(\theta)$ and if their rates R_n tend to $c(\theta)$ for sufficiently large average check node degrees as n tends to infinity.

[2]All the results presented in this paper are for the asymptotic regime where the block length tends to infinity.

Consider the $(k+2)$-tuple $(\lambda_2, \lambda_3, ..., \lambda_k, D_v, \rho(x); \theta)$ which corresponds to a degree distribution $(\lambda(x), \rho(x)) = (\sum_{i=2}^{k} \lambda_i x^{i-1} + (1 - \sum_{i=2}^{k} \lambda_i) x^{D_v - 1}, \rho(x))$ over $C(\theta)$ where $D_v > k$, and $0 \leq \lambda_i < 1, \forall i \in \{2, ..., k, D_v\}$. We call this setting a *code-channel pair*. With slight negligence, we call a code-channel pair convergent if the ensemble is convergent over the channel. We define $\Lambda_2(\rho(x), \theta) = \{\lambda_2 : \exists D_v; (\lambda_2, D_v, \rho(x); \theta)$ is convergent$\}$. Similarly, $\Lambda_3(\rho(x), \theta) = \{\lambda_3 : \forall \lambda_2 \in \Lambda_2(\rho, \theta), \exists D_v; (\lambda_2, \lambda_3, D_v, \rho(x); \theta)$ is convergent$\}$ and so on. We show the corresponding complement sets with respect to [0,1], with $\Lambda_i^C(\rho(x), \theta)$. It can be verified that these sets are continuous for BIOSM channels.

III. Capacity Achieving Sequences and the Principle of Successive Maximization

Capacity achieving sequences of LDPC code ensembles over the BEC were first designed by Shokrollahi *et al.* in [3]–[5]. In [12]–[14], the authors proposed new sequences of capacity achieving LDPC code ensembles over the BEC which contained Shokrollahi's sequences as a special case. Unlike Shokrollahi's sequences in which all variable node degrees from 2 to D_v have to be present, the sequences of [12]–[14] only contain variable node degrees from 2 to $k < D_v$ and degree D_v, where k is a strictly increasing function of D_v (and ultimately a function of $\rho(x)$ and θ).[3] For a given ensemble (λ^n, ρ^n) within the sequences proposed in [12]–[14], we have: $\lambda_i^n = T_i/\theta = \lambda_i^{n*}$, $2 \leq i \leq k$, and $\lambda_{D_v}^n = 1 - \sum_{i=2}^{k} \lambda_i^n$, where D_v is determined by θ and $\rho(x)$ such that

$$\sum_{i=2}^{D_v - 1} T_i \leq \theta < \sum_{i=2}^{D_v} T_i. \tag{4}$$

In the following, we generalize the definition of λ_i^{n*} to any type of BIOSM channel as

$$\lambda_i^{n*} \triangleq T_i / \mathfrak{P}(P_0). \tag{5}$$

In the method proposed in [12]–[14], the values of λ_i^n are computed based on the following principle: Starting from $i = 2$, and by sequentially going through the values of $i \geq 2$, set the value of λ_i^n to a maximum value $\widetilde{\lambda_i^n}$ such that the ensemble remains convergent for sufficiently large D_v. For the BEC, this process results in $\widetilde{\lambda_i^n} = \lambda_i^{n*}$ [13], [14]. Intuitively, the method of [12]–[14] works since based on (1), maximizing the rate is equivalent to maximizing $\overline{d_v}^{-1}$, which in turn implies that one should assign higher percentages to the lower degree coefficients as far as the constructed ensemble remains convergent.

In this paper, we generalize the method of [12]–[14] to BIOSM channels and refer to it as *successive maximization*. At the core of this method are the successive upper bounds $\widetilde{\lambda_i^n}$ on λ_i^n values. These upper bounds should fulfill a threshold property similar to that of λ_i^{n*} over the BEC as follows: Suppose that $\lambda_i^n = \widetilde{\lambda_i^n}$ for every $i < k$, then for any ensemble with $\lambda_k^n > \widetilde{\lambda_k^n}$, the probability of error is bounded away from zero. Moreover, for any ensemble with $\lambda_k^n < \widetilde{\lambda_k^n}$, if the probability of error is made sufficiently small, then it will tend to zero as the number of iterations tends to infinity, regardless of the values of λ_i^n, $i > k$. Unlike the BEC case, there is no proof that such upper bounds exist for other BIOSM channels (with the exception of the bound on λ_2), and even if they do exist, their values may not be easily obtained analytically (except for $i = 2$, where we have $\widetilde{\lambda_2^n} = \lambda_2^{n*}$). It should be noted that for the BIAWGN channel, the existence of such an upper bound on λ_3 was demonstrated empirically in [25]. In the next section, we prove the existence of a *positive* upper bound on λ_3 that fulfils the aforementioned properties over any BIOSM channel, and conjecture that similar upper bounds exist for other λ_i values. This makes it possible to apply the SM principle as a design tool for ensemble sequences. Numerical evidence presented in Section VI confirms that the resulting sequences are at least capacity approaching and may in fact be capacity achieving.

[3]Shokrollahi's sequences correspond to those of [12] with $k(D_v) = D_v - 1$.

IV. Some Analytical Results on Unpunctured LDPC Codes

We first derive an inequality which helps in proving the statements of this section. For the code-channel pair $E \triangleq (\lambda_2, \lambda_3, ..., \lambda_k, D_v, \rho(x); \theta)$, define $f(y) = \mathfrak{P}(P_0)\lambda(1 - \rho(1 - y))$. Using (2), it is easy to see that the inequality $f(y) < y$, for $0 < y \leq \mathfrak{P}(P_0)$, is a sufficient condition for the convergence of E. Defining $x = 1 - \rho(1 - y)$, this condition can be reformulated as

$$\mathfrak{P}(P_0)\lambda(x) - 1 + \rho^{-1}(1 - x) < 0, \quad 0 < x \leq 1 - \rho(1 - \mathfrak{P}(P_0)) < 1.$$

Using the Taylor series expansion of $1 - \rho^{-1}(1 - x)$ and rearranging the terms, we thus have the following sufficient condition for convergence:

$$\sum_{i=2}^{k} (\mathfrak{P}(P_0)\lambda_i - T_i) x^{i-1} + (\mathfrak{P}(P_0)\lambda_{D_v} - T_{D_v}) x^{D_v - 1} -$$

$$\sum_{i=k+1, i \neq D_v}^{\infty} T_i x^{i-1} < 0, \quad 0 < x \leq 1 - \rho(1 - \mathfrak{P}(P_0)) < 1. \tag{6}$$

Recall that the stability condition in [10], [11] remains silent about the case where $\lambda'(0)\rho'(1)$ is exactly equal to $1/\mathfrak{P}(P_0)$. Here, we prove that for this case, a similar upper bound exists for λ_3. Before proving the main result, we prove some auxiliary lemmas and propositions.

Proposition 1: Consider a convergent code-channel pair $E_1 = (\lambda_2, \lambda_3, .., \lambda_a, ..., \lambda_b, ..., \lambda_k, D_v, \rho(x); \theta_0)$. Define $E_2 = (\lambda_2, \lambda_3, .., \lambda_a - \epsilon, ..., \lambda_b + \epsilon, ..., \lambda_k, D_v, \rho(x); \theta_0)$ such that $0 < \lambda_a - \epsilon, \lambda_b + \epsilon < 1$. Then the code-channel pair E_2 is also convergent.

Proof: The proof is similar to the proof of Lemma 1 of [30] and is thus omitted.

Corollary 1: Consider a code-channel pair $(\lambda_2, .., \lambda_k, \lambda_{D_v}, \rho(x); \theta)$ and the set $\Lambda_3(\rho, \theta)$ as defined in Section II. Select a $\lambda_3^\dagger \in \Lambda_3$. Then any $0 \leq \lambda_3 < \lambda_3^\dagger$ also belongs to Λ_3. In other words, the set Λ_3 is continuous.

Lemma 1 (Sufficient condition for stability): Let (λ, ρ)

be a degree distribution over $C(\theta)$ and P_0 be the initial channel density. Let $\lambda_2 = \lambda_2^* = 1/(\mathfrak{P}(P_0)\rho'(1))$, and $\lambda_3 < \lambda_3^* = \rho''(1)/(2\mathfrak{P}(P_0)\rho'(1)^3)$ (note that the right hand side is strictly positive), then the ensemble is stable, i.e., there exists $\xi > 0$ such that if $\mathbb{P}(P_l) < \xi$, then $\lim_{l\to\infty} \mathbb{P}(P_l) = 0$.
Proof: Starting from (2) and writing the Taylor expansion of the right hand side of the inequality at zero, we have:

$$p_l \le \mathfrak{P}(P_0)[\lambda'(0)\rho'(1)p_{l-1} + (-\lambda'(0)\rho''(1) + \lambda''(0)\rho'^2(1))p_{l-1}^2] + O(p_{l-1}^3).$$

Let $g(t) = -\rho''(1)/\rho'(1) + \mathfrak{P}(P_0)2t\rho'^2(1)$. Using $\lambda'(0)\rho'(1) = 1/\mathfrak{P}(P_0)$, we thus have:

$$\begin{aligned} p_l &\le p_{l-1} + \mathfrak{P}(P_0)[-\lambda'(0)\rho''(1) + \lambda''(0)\rho'^2(1)]p_{l-1}^2 + O(p_{l-1}^3) \\ &= p_{l-1} + [-\rho''(1)/\rho'(1) + \mathfrak{P}(P_0)2\lambda_3\rho'^2(1)]p_{l-1}^2 + O(p_{l-1}^3) \\ &= p_{l-1} + g(\lambda_3)p_{l-1}^2 + O(p_{l-1}^3) \le p_{l-1} + (g(\lambda_3)+\delta)p_{l-1}^2 \triangleq h(p_{l-1}), \end{aligned}$$

where the last inequality is valid for small enough p_{l-1}, and δ is a small positive number. Since $\lambda_3 < \rho''(1)/(2\mathfrak{P}(P_0)\rho'(1)^3)$, then $g(\lambda_3) < 0$. We choose a positive constant κ, such that if $p_{l-1} < \kappa$, then δ can be chosen small enough such that $g(\lambda_3) + \delta$, the coefficient of p_{l-1}^2, is negative. This means that $p_l \le h(p_{l-1}) < p_{l-1}$ for $0 < p_{l-1} < \kappa$. Note that for $t \in [0, \kappa]$, the only fixed point of function $h(t)$ is at $t = 0$. Since $h(t) < t$ for $t \in (0, \kappa)$, we can see that $\lim_{l\to\infty}(p_l) = 0$. Now if we let $\xi = \kappa^2/4$, using (3), if $\mathbb{P}(P_l) < \xi$, then $\lim_{l\to\infty}(p_l) = 0$. This means that $\lim_{l\to\infty} \mathbb{P}(P_l) = 0$. ■

Lemma 2: Consider the code-channel pair $(\lambda_2^*, \lambda_3, D_v, \rho(x); \theta)$ such that $\lambda_3 < \lambda_3^*$. Then for large enough D_v, the pair is convergent.
Proof: The sufficient convergence condition of (6) for the given ensemble reduces to

$$\begin{aligned} &(\mathfrak{P}(P_0)\lambda_3 - T_3)x^2 + (\mathfrak{P}(P_0)\lambda_{D_v} - T_{D_v})x^{D_v-1} - \sum_{i=4, i\neq D_v}^{\infty} T_i x^{i-1} \\ &\quad = x^2[(\mathfrak{P}(P_0)\lambda_3 - T_3) + (\mathfrak{P}(P_0)\lambda_{D_v} - T_{D_v})x^{D_v-3}] \\ &\quad - \sum_{i=4, i\neq D_v}^{\infty} T_i x^{i-1} < 0; \quad 0 < x \le 1 - \rho(1 - \mathfrak{P}(P_0)) < 1. \end{aligned}$$

Note that the first term in the coefficient of x^2 is negative based on the lemma assumption and since $0 < x < 1$, this term can be made dominant for sufficiently large D_v, making the term with x^2 negative. All the other terms including $x^i, i \ge 3$ are also negative as $T_i > 0, \forall i$. Therefore, the convergence condition holds for the given ensemble. ■

Lemma 3: Let (λ, ρ) be a convergent degree distribution over $C(\theta)$. Then we necessarily have $\lambda_3 \le \lambda_3^U$ where $\lambda_3^U = 3/(\overline{d_c}(1 - c(\theta))) - (3/2)\lambda_2$.[4]

[4]Note that other (possibly tighter) upper bounds have recently been proposed in [27].

Proof: For any convergent ensemble, we must have $R \le c(\theta)$. Using (1), we thus have $\overline{d_v}^{-1} \le 1/(\overline{d_c}(1 - c(\theta)))$. Also, $\lambda_2/2 + \lambda_3/3 \le \sum_2^{D_v} \lambda_i/i = 1/\overline{d_v}$. Therefore, $\lambda_2/2 + \lambda_3/3 \le 1/(\overline{d_c}(1 - c(\theta)))$, which reduces to $\lambda_3 \le \lambda_3^U$. ■

Theorem 1: Consider the code-channel pair $(\lambda_2^*, \lambda_3, D_v, \rho(x); \theta)$ where D_v can be arbitrarily large. There exists a threshold value $\widetilde{\lambda_3}$ in the interval $[\lambda_3^*, \lambda_3^U]$ such that if $\lambda_3 < \widetilde{\lambda_3}$, the ensemble is convergent for a sufficiently large value of D_v and if $\lambda_3 > \widetilde{\lambda_3}$, the probability of error is bounded away from zero regardless of the value of D_v.
Proof: Define $\widetilde{\lambda_3} = \inf(\Lambda_3^C(\rho(x), \theta))$. Based on this definition, $\forall \lambda_3 > \widetilde{\lambda_3}$, the probability of error of the resulting ensemble $(\lambda_2^*, \lambda_3, D_v, \rho(x))$ over $C(\theta)$ is bounded away from zero no matter how large D_v is. Also, there exists an arbitrarily small $\epsilon \ge 0$, such that $(\lambda_2^*, \widetilde{\lambda_3} - \epsilon, D_v, \rho(x); \theta)$ is convergent for a sufficiently large value of D_v. Now based on Proposition 1, if such a code-channel pair is convergent, any other pair $(\lambda_2^*, \lambda_3, D_v, \rho(x); \theta)$ for which $\lambda_3 < \widetilde{\lambda_3} - \epsilon$ is also convergent. Based on Lemma 2, we know that if $\lambda_3 < \lambda_3^*$, $(\lambda_2^*, \lambda_3, D_v, \rho(x), \theta)$ is convergent for sufficiently large D_v. Therefore $\widetilde{\lambda_3} \ge \lambda_3^*$. From Lemma 3, we know that $\widetilde{\lambda_3} \le \lambda_3^U$. This proves the theorem. ■
We expect the result of Theorem 1 to be generalizable to $\lambda_k, k > 3$, if $\lambda_i = \widetilde{\lambda_i} = \inf(\Lambda_i^C(\rho(x), \theta)), 2 \le i \le k - 1$. This, however, remains to be proved.

V. Universally Capacity Approaching Rate-Compatible LDPC Codes

In Section III, we discussed the design of sequences of degree distributions (λ^n, ρ^n) whose rates approach the capacity as n tends to infinity. In this section, we consider the problem of puncturing a degree distribution for a given n. For simplicity, we sometime drop the index n and refer to the ensemble as the *parent* ensemble. We use the notations (λ^p, ρ) and R^p for the parent ensemble and its rate, respectively. We show the fraction of the punctured bits (variable nodes) by Π. The resulting code rate in this case is equal to $R^p/(1 - \Pi)$. If the puncturing is performed randomly, we refer to it as *random puncturing*. Otherwise, the puncturing is called *intentional* [16]. In intentional puncturing, variable nodes of degree i can potentially have different puncturing fractions Π_i. The overall puncturing fraction Π can then be expressed as $\Pi = \sum_{i=2}^{D_v} \Pi_i \overline{\lambda_i}^p$ where $\{\overline{\lambda_i}^p\}$ is the node-based degree distribution of variable nodes for the parent ensemble.

In many situations, it is necessary to obtain more than one rate by puncturing. In this case, for a simple implementation, the puncturing pattern should be in such a way that for 2 consecutive rates, the punctured code with a higher rate can be constructed by puncturing the code with the lower rate. A puncturing pattern with this property is called *rate-compatible*. Let the set of channel parameters θ^j be ordered reversely by channel degradation (i.e., θ^0 is for the worst channel condition which corresponds to the parent code). For any $C(\theta^j)$, consider the set $\Phi^j = \{\Pi_i^j, 2 \le i \le D_v\}$.[5] For a

[5]For $\lambda_i = 0$, we assume $\Pi_i = 0$.

rate-compatible puncturing scheme, we must have $\Pi_i^m \leq \Pi_i^n$ for any $m < n$ and any i.

To analyze the asymptotic behavior of a punctured ensemble, we model the puncturing of LDPC codes over a channel $C(\theta)$ as the transmission of the unpunctured bits over $C(\theta)$ while sending the punctured bits on an erasure channel with erasure probability of 1. Let $\mathcal{E}$ be the set of all edges in the graph. Also let $\mathcal{E}_i^{punc}$ be the set of edges in the graph which are connected to the variable nodes of degree i which are punctured. Also, let $\mathcal{E}^{punc}$ be the union of sets $\mathcal{E}_i^{punc}$. Similarly, define $\mathcal{E}^{un}$ and $\mathcal{E}_i^{un}$ for unpunctured edges. We define $\lambda^{punc}(x) = \sum \lambda_i^{punc} x^{i-1}$, where $\lambda_i^{punc} = \frac{|\mathcal{E}_i^{punc}|}{|\mathcal{E}^{punc}|}$. Notation $|.|$ denotes the cardinality of the set. We also define φ^{punc} as the fraction of punctured edges: $\varphi^{punc} = \frac{|\mathcal{E}^{punc}|}{|\mathcal{E}|}$. The polynomial $\lambda^{un}(x)$ and the fraction φ^{un} can be defined similarly for unpunctured edges. Based on the above definitions, we have:

$$\Pi_i = \frac{|\mathcal{E}_i^{punc}|}{|\mathcal{E}_i^{punc}| + |\mathcal{E}_i^{un}|} = \frac{\frac{|\mathcal{E}_i^{punc}|}{|\mathcal{E}|}}{\lambda_i^p} = \frac{\lambda_i^{punc}\varphi^{punc}}{\lambda_i^p}, \quad (7)$$

$$\varphi^{punc}\lambda_i^{punc} + \varphi^{un}\lambda_i^{un} = \lambda_i^p. \quad (8)$$

We can now derive the density evolution equations for our setting. Similar to the previous section, let Q_l be the probability density function of the outgoing message of the check nodes at iteration l. We define P_l^{punc} and P_l^{un} as the density at the output of the punctured and unpunctured variable nodes at iteration l, respectively. We then have

$$\begin{aligned} P_l^{punc} &= P_o^{punc} \otimes \lambda^{punc}(Q_l), \\ P_l^{un} &= P_o^{un} \otimes \lambda^{un}(Q_l), \\ P_l' &= \varphi^{punc} P_l^{punc} + \varphi^{un} P_l^{un}, \\ Q_l &= \Gamma^{-1}(\rho(\Gamma(P_{l-1}'))), \end{aligned} \quad (9)$$

in which $P_0^{punc} = \Delta_0$ where Δ_x is the Dirac delta function at x [11]. Consider a sequence of degree distributions $(\lambda^n(x), \rho^n(x))$. Consider also a set of channels with parameters $\theta^j, j = 0, 1, ..., J$, ordered increasingly by their quality. Now assume that the parent ensemble sequence $(\lambda^n(x), \rho^n(x))$ is punctured by the set $\phi^{n,j} = \{\Pi_i^{n,j}, 2 \leq i \leq D_v\}$ to create higher rate ensemble sequences that are convergent over the corresponding channels. This scheme is *universally capacity achieving* if $\lim_{n\to\infty} R^{n,j} = c(\theta^j)$ for all values of j. A universally capacity achieving scheme is called rate-compatible if the puncturing patterns $\phi^{n,j}$ are rate-compatible for every value of n.

In the following, we prove a theorem for puncturing a given degree distribution. Let (λ^p, ρ) be a convergent parent ensemble over a channel with parameter θ^0. The code-channel pair $(\lambda^p, \rho; \theta)$ is convergent for any $\theta \leq \theta^0$. Let P_0 be the channel density associated with θ. We define parameter $\widetilde{\Pi_2}$, corresponding to the parent code-channel pair, as:

$$\widetilde{\Pi_2} = \frac{[1 - \mathfrak{P}(P_0)\rho'(1)\lambda_2^p]}{[1 - \mathfrak{P}(P_0)]\rho'(1)\lambda_2^p}. \quad (10)$$

Note that if the pair is stable, i.e., if $\lambda_2 < \lambda_2^*$, then $\widetilde{\Pi_2} > 0$. The following lemma can be easily proved based on (7) and (8).

Lemma 4: Let

$$\vartheta = \rho\prime(1)(\varphi^{punc} p_0^{punc} \lambda_2^{punc} + \varphi^{un} p_0^{un} \lambda_2^{un}), \quad (11)$$

where $p_0^{un} = \mathfrak{P}(P_0^{un})$ and $p_0^{punc} = \mathfrak{P}(P_0^{punc}) = 1$. We have $\Pi_2 \lesseqqgtr \widetilde{\Pi_2}$ iff $\vartheta \lesseqqgtr 1$.

Theorem 2: Let (λ^p, ρ) be a parent ensemble convergent over $C(\theta)$ with $\lambda_2^p \neq 0$. Suppose that this code is punctured based on the set $\Phi = \{\Pi_i; i = 2, \ldots, D_v\}$ (note that $C(\theta)$ has a one-to-one correspondence with the channel density P_0). There exists a threshold value $\widetilde{\Pi_2}$, given by (10), such that if $\Pi_2 > \widetilde{\Pi_2}$, then for any l, $\mathbb{P}(P_l^{punc})$ and $\mathbb{P}(P_l^{un})$ are bounded away from zero and if $\Pi_2 < \widetilde{\Pi_2}$, there exists a strictly positive constant ξ such that if $\mathbb{P}(P_l^{punc}) < \xi$, and $\mathbb{P}(P_l^{un}) < \xi$ for some l, then $\lim_{l\to\infty} \mathbb{P}(P_l^{punc}) = 0$ and $\lim_{l\to\infty} \mathbb{P}(P_l^{un}) = 0$.

Proof: [*Sufficiency*] ($\Pi_2 < \widetilde{\Pi_2}$):
Let $p_l^{punc}, p_l^{un}, p_0^{punc}, p_0^{un}, p_l'$ and q_l denote $\mathfrak{P}(P_l^{punc}), \mathfrak{P}(P_l^{un}), \mathfrak{P}(P_0^{punc}), \mathfrak{P}(P_0^{un}), \mathfrak{P}(P_l')$ and $\mathfrak{P}(Q_l)$, receptively. Now for the density evolution equations, we have:

$$\begin{aligned} p_l^{punc} &= p_0^{punc}\lambda^{punc}(q_l), \\ p_l^{un} &= p_0^{un}\lambda^{un}(q_l), \\ p_l' &= \varphi^{punc} p_l^{punc} + \varphi^{un} p_l^{un}, \\ q_l &\leq 1 - \rho(1 - p_{l-1}'). \end{aligned} \quad (12)$$

Combining the equations, we obtain:

$$q_{l+1} \leq 1 - \rho(1 - \varphi^{punc} p_0^{punc} \lambda^{punc}(q_l) - \varphi^{un} p_0^{un} \lambda^{un}(q_l)). \quad (13)$$

By expanding the above formula into Taylor series at zero we have:

$$q_{l+1} \leq \rho\prime(1)(\varphi^{punc} p_0^{punc} \lambda_2^{punc} q_l + \varphi^{un} p_0^{un} \lambda_2^{un} q_l) + O(q_l^2),$$

or

$$q_{l+1} \leq \vartheta q_l + O(q_l^2),$$

where ϑ is defined in (11). Based on Lemma 4, $\vartheta < 1$ and thus we can find $\eta > 0$ such that $\vartheta + \eta < 1$. Note that since $\mathbb{P}(P_l^{punc}) < \xi$, and $\mathbb{P}(P_l^{un}) < \xi$ for an arbitrarily small ξ, based on (3), p_l^{un} and p_l^{punc} are also arbitrarily small. This makes p_l' arbitrarily small based on (12). Since $1-\rho(1-x)$ is a strictly increasing function and $\lim_{x\to 0} 1-\rho(1-x) = 0$, based on (12), we can make q_l arbitrarily small if we choose small enough ξ. For sufficiently small q_l, we can see that $q_{l+1} < (\vartheta + \eta)q_l < q_l$. With an argument similar to the one used in the proof of Lemma 1, we thus have $\lim_{l\to\infty} q_l = 0$. Based on (12), $\lim_{l\to\infty} p_l^{punc} = \lim_{l\to\infty} p_l^{un} = 0$, which implies $\lim_{l\to\infty} \mathbb{P}(P_l^{punc}) = 0$, and $\lim_{l\to\infty} \mathbb{P}(P_l^{un}) = 0$.

[*Necessity*]($\Pi_2 > \widetilde{\Pi_2}$):
To prove the necessity, it is enough to show that the probability of error of Q_n, ($\mathbb{P}(Q_n)$) is bounded away from zero as n tends to infinity. Suppose that $\mathbb{P}(Q_n)$ tends to zero. Then there exists an iteration l for which $\mathbb{P}(Q_l) = \epsilon/2$, where we can choose ϵ arbitrarily small. For the next iteration, similar to the proof of Theorem 5 of [10], assume that we replace this density with the BEC density of the same probability of error, i.e.,

$\Upsilon_0 = \epsilon\Delta_0 + (1-\epsilon)\Delta_\infty$. Now, if we indicate the resulting density, at the output of variable nodes, for the next iteration by Υ_1, we have:

$$\Upsilon_1 = \epsilon(\lambda_2^{punc}\varphi^{punc}\rho\prime(1)P_0^{punc} + \lambda_2^{un}\varphi^{un}\rho\prime(1)P_0^{un}) + \Delta_\infty(1-\epsilon(\lambda_2^{punc}\varphi^{punc}\rho\prime(1) + \lambda_2^{un}\varphi^{un}\rho\prime(1))) + O(\epsilon^2).$$

It can be verified that after i iterations, we obtain:

$$\Upsilon_i = \epsilon[\lambda_2^{punc}\varphi^{punc}\rho\prime(1)P_0^{punc} + \lambda_2^{un}\varphi^{un}\rho\prime(1)P_0^{un}]^{\otimes i} + \Delta_\infty(1-\epsilon[\lambda_2^{punc}\varphi^{punc}\rho\prime(1) + \lambda_2^{un}\varphi^{un}\rho\prime(1)]^i) + O(\epsilon^2),$$

where $\otimes$ denotes the i-fold convolution.

Define the sequence of densities

$$D_i = [\lambda_2^{punc}\varphi^{punc}\rho\prime(1)P_0^{punc} + \lambda_2^{un}\varphi^{un}\rho\prime(1)P_0^{un}]^{\otimes i}/D,$$

where

$$D = \int_{-\infty}^{\infty}[\lambda_2^{punc}\varphi^{punc}\rho\prime(1)P_0^{punc} + \lambda_2^{un}\varphi^{un}\rho\prime(1)P_0^{un}]^{\otimes i}dx.$$

Using the properties of $\mathfrak{P}(\cdot)$ in Section II, it is easy to verify that the value of $\mathfrak{P}(D_i)$ is equal to ϑ^i/D. We claim that there exists an n for which, if $i \geq n$, we have $1 < \frac{\vartheta^{2i}}{4D}$. To see this, note that if $4D \leq 1$, since based on Lemma 4 $\vartheta > 1$, the inequality is trivial. Otherwise, to satisfy the inequality, we set $n = \lceil \frac{\ln 4D}{2\ln\vartheta} \rceil$. Based on (3),

$$(\frac{\vartheta^i}{D})^2 \leq 4\mathbb{P}(D_i)(1-\mathbb{P}(D_i)) < 4\mathbb{P}(D_i).$$

Therefore, for $i \geq n$, we obtain $\frac{1}{D} < \mathbb{P}(D_i)$. For $i = n$, we can write

$$\begin{aligned}\mathbb{P}(\Upsilon_n) &= \epsilon\mathbb{P}([\lambda_2^{punc}\varphi^{punc}\rho\prime(1)P_0^{punc} \\ &+ \lambda_2^{un}\varphi^{un}\rho\prime(1)P_0^{un}]^{\otimes n}) + O(\epsilon^2).\end{aligned}$$

For sufficiently small ϵ, we then have $\mathbb{P}(\Upsilon_n) = \epsilon D\mathbb{P}(D_n) + O(\epsilon^2) > \epsilon > \epsilon/2 = \mathbb{P}(\Upsilon_0)$. Now since Q_l is physically degraded with respect to Υ_0 (see [11], Lemma 4.78), Q_{l+n} is physically degraded with respect to Υ_n (see [11], Lemma 4.80). Therefore, $\mathbb{P}(Q_{l+n}) \geq \mathbb{P}(\Upsilon_n) > \mathbb{P}(\Upsilon_0) = \mathbb{P}(Q_l)$. This contradicts the fact that the probability of error is a decreasing function of the number of iterations. In other words, ϵ can not become arbitrary small. This completes the proof. ■

This property is similar to the stability condition [10] for parent LDPC codes which provides an upper bound on λ_2.

Corollary 2 (Independency of Π_2 from n for puncturing schemes with $\Pi_2 = \widetilde{\Pi_2}$): Consider a sequence of ensembles $(\lambda^n(x), \rho^n(x))$ which are convergent over $C(\theta^0)$ and let P_0^0 be the associated channel density. Now consider an improved channel $C(\theta^j)$, $j > 0$ and let P_0^j be the associated channel density. If for any ensemble within the sequence, the value of λ_2 satisfies the stability condition corresponding to θ^0 with equality, i.e., if $\lambda_2^n = \lambda_2^{n*}$, then the value of the upper bound $\widetilde{\Pi_2}$ corresponding to θ^j obtained in Theorem 2, is independent of n (in fact, it is independent of the parent ensemble sequence (λ^n, ρ^n)).

Proof: We have $\lambda_2^n = 1/(\mathfrak{P}(P_0^p)\rho'^n(1))$. Replacing this value in (10), we obtain $\widetilde{\Pi_2} = \frac{[1-\mathfrak{P}(P_0^j)/\mathfrak{P}(P_0^0)]}{[1-\mathfrak{P}(P_0^j)]/\mathfrak{P}(P_0^0)}$, which is independent of n. ■

Corollary 3 (Rate-compatibility of Π_2 for puncturing schemes with $\Pi_2 = \widetilde{\Pi_2}$): Consider a sequence of ensembles $(\lambda^n(x), \rho^n(x))$ which are convergent over $C(\theta^0)$ and let P_0^0 be the associated channel density. Now consider an improved channel $C(\theta^j)$, $j \geq 0$ and let P_0^j be the associated channel density. If for any ensemble within the sequence, the value of λ_2 satisfies the stability condition corresponding to θ^0 with equality, i.e., if $\lambda_2^n = \lambda_2^{n*}$, then the value of the upper bound $\widetilde{\Pi_2}$ is a decreasing function of θ^j.

Proof: From Corollary 2, we have $\widetilde{\Pi_2} = \frac{[1-\mathfrak{P}(P_0^j)/\mathfrak{P}(P_0^0)]}{[1-\mathfrak{P}(P_0^j)]/\mathfrak{P}(P_0^0)}$. It is easy to check that function $f(x) = \frac{[1-x/\mathfrak{P}(P_0^0)]}{[1-x]/\mathfrak{P}(P_0^0)}$ is a decreasing function of x for $x \in (0,1)$. This together with the fact that $\mathfrak{P}(P_0^j)$ is an increasing function of θ^j proves that $\widetilde{\Pi_2}$ is a decreasing function of θ^j. ■

In [12], [13], similar upper bounds to that of stability condition were obtained for other variable node degrees over the BEC. In the following, we prove a similar result for the case of rate-compatible codes over the BEC. Equation (13) can be rewritten as follows for the case of the BEC:

$$q_{l+1} = 1-\rho(1-\varphi^{punc}p_0^{punc}\lambda^{punc}(q_l) - \varphi^{un}p_0^{un}\lambda^{un}(q_l)) = g(q_l). \tag{14}$$

Similar to the unpunctured case, the convergence condition can be rewritten as

$$g(y) < y, 0 < y \leq 1-\rho(\varphi^{punc}p_0^{punc} + \varphi^{un}p_0^{un}).$$

Note that the above equation is a necessary and sufficient condition for convergence. Modifying the inequality we obtain:

$$\varphi^{punc}p_0^{punc}\lambda^{punc}(y) + \varphi^{un}p_0^{un}\lambda^{un}(y) - 1 + \rho^{-1}(1-y) < 0.$$

By expanding $1-\rho^{-1}(1-y)$ into Taylor series and rearranging the terms we have

$$\sum_{i=2}^{D_v}(\varphi^{punc}p_0^{punc}\lambda_i^{punc} + \varphi^{un}p_0^{un}\lambda_i^{un} - T_i)y^{i-1} - \sum_{i=D_v+1}^{\infty}T_i y^{i-1} < 0. \tag{15}$$

Now based on Theorem 2, there is an upper bound $\widetilde{\Pi_2}$ on the value of Π_2. Setting $\Pi_2 = \widetilde{\Pi_2}$ is equivalent to having $\vartheta = 1$. Putting this together with $T_2 = 1/\rho\prime(1)$, makes the coefficient of y in (15) equal to zero. Now to have convergence, the inequality must hold for all values of y including those close to zero. For those values, the dominant term is the term with degree 2. This implies that for convergence, we must have $(\varphi^{punc}p_0^{punc}\lambda_3^{punc} + \varphi^{un}p_0^{un}\lambda_3^{un}) \leq T_3$. This imposes an upper bound on the value of Π_3 above which the probability of erasure is bounded away from zero and below which the ensemble is convergent if the probability of erasure is made sufficiently small. The same method can be applied for other values of i. This is explained in the following proposition.

Proposition 2: Let (λ^p, ρ) be a convergent parent ensemble over the BEC with channel parameter ϵ^0. Suppose that the parent ensemble is punctured to be used over a channel with parameter $\epsilon^j < \epsilon^0$. Let p_0^{un} be the Bhattacharyya

constant for this channel, i.e., $p_0^{un} = \epsilon^j$. Also assume that $\lambda_i^p \neq 0, 2 \leq i \leq n \leq D_v$. Define

$$\Pi_i^* = \frac{1 - p_0^{un}\lambda_i^p/T_i}{(1 - p_0^{un})\lambda_i^p/T_i}. \quad (16)$$

Then if $\Pi_i = \Pi_i^*$ for $2 \leq i < n$, there exists an upper bound $\Pi_n^* = \frac{1-p_0^{un}\lambda_n^p/T_n}{(1-p_0^{un})\lambda_n^p/T_n}$ on Π_n above which the resulting punctured ensemble is not convergent over $C(\epsilon^j)$ and below which the ensemble is convergent over $C(\epsilon^j)$ if the probability of erasure can be made sufficiently small.[6]

Proof: As discussed before, to have a threshold upper bound on Π_n, terms with degree $i < n$ in (15) have to be removed. In order to do so, one has to set

$$(\varphi^{punc} p_0^{punc} \lambda_i^{punc} + \varphi^{un} p_0^{un} \lambda_i^{un}) = T_i, 2 \leq i < n.$$

Using (7) and (8) we can see that this is equivalent to having $\Pi_i = \frac{1-p_0^{un}\lambda_i^p/T_i}{(1-p_0^{un})\lambda_i^p/T_i} = \Pi_i^*$ for $2 \leq i < n$. This completes the proof. ■

We now would like to prove that the construction of universally capacity achieving rate-compatible LDPC codes over the BEC can be achieved by applying the SM principle to the values of Π_i, i.e., starting from a parent sequence and for each ensemble member of the sequence, we maximize Π_2 as far as the ensemble remains convergent and continue this procedure successively for other Π_i values. This will be performed for each of the J target channel parameters and we demonstrate that the resulting puncturing patterns are in fact rate-compatible. We also show that if the original parent sequence is capacity achieving, so will be all the J sequences of punctured ensembles.

Theorem 3: Consider a capacity achieving parent ensemble sequence (λ^n, ρ^n) over the BEC with parameter ϵ^0, constructed based on the method of [12]. For the set of channel erasure values ϵ^j ($\epsilon^1 > \epsilon^2 > ... > \epsilon^J$), we puncture each ensemble within the parent sequence based on the SM principle. The resulting scheme is then universally capacity achieving rate-compatible.

Proof: For a given n, assume that the ensemble includes constituent variable node degrees 2 to k and the maximum variable node degree D_v. For $2 \leq i \leq k$, the values of $\Pi_i^{n,j}$ resulting from the SM principle can be obtained based on Proposition 2. Also note that since the ensemble is designed based on [12], we have $\lambda_i^n = T_i^n/\epsilon^0$ for $2 \leq i \leq k$. Replacing the values of λ_i^n in (16), we have:

$$\Pi_i^{n,j} = \frac{1 - \epsilon^j/\epsilon^0}{(1 - \epsilon^j)/\epsilon^0} = \frac{\epsilon^0 - \epsilon^j}{(1 - \epsilon^j)} \triangleq P^{(j)}; 2 \leq i \leq k, 1 \leq j \leq J, \forall n, \quad (17)$$

which is only a function of j. We also set $\Pi_{D_v}^{n,j} = P^{(j)}$, $\forall n, \forall j$. In this case, we have (for simplification, the indices n and j are dropped for the puncturing-related parameters):

$$\varphi^{punc} p_0^{punc} \lambda_{D_v}^{punc} + \varphi^{un} p_0^{un} \lambda_{D_v}^{un} = \lambda_{D_v}^n P^{(j)} + \epsilon^j \lambda_{D_v}^n (1 - P^{(j)})$$

$$= \epsilon^0 \lambda_{D_v}^n = \epsilon^0 (1 - \sum_{i=2}^{k} T_i/\epsilon^0) = \epsilon^0 - \sum_{i=2}^{k} T_i, \quad (18)$$

[6]Note that for $i = 2$, we have $\Pi_2^* = \widetilde{\Pi_2}$.

where we use (7), (8) and (17) for the first equality and (17) for the second equality. To demonstrate that the resulting scheme is convergent over $C(\epsilon^j)$, it is enough to show that (15) holds for the given puncturing fractions $\Pi_i^{n,j}$. In (15), after applying the SM principle, the terms with $i = 2$ to k will all be equal to zero, and the terms with $i > D_v$ are equal to $-T_i y^{i-1}$ and will be negative. We show the sum of the remaining terms by S and will have:

$$S = (\sum_{i=k+1}^{D_v - 1} -T_i) y^{i-1} + ((\epsilon^0 - \sum_{i=2}^{k} T_i) - T_{D_v}) y^{D_v - 1} <$$

$$(\sum_{i=k+1}^{D_v - 1} -T_i + \epsilon^0 - \sum_{i=2}^{k} T_i - T_{D_v}) y^{D_v - 1} = (\epsilon^0 - \sum_{i=2}^{D_v} T_i) y^{D_v - 1} < 0,$$

where the first equality is obtained based on (18), the first inequality is a result of $y^{i-1} > y^{D_v - 1}$ for $i < D_v$ and $0 < y < 1$, and the last inequality holds based on (4). This completes the convergence proof.

Now we prove that

$$R^{n,j}/c(\epsilon^j) = R^{n,0}/c(\epsilon^0). \quad (19)$$

Note that $\Pi^{n,j} = \sum_{i=2}^{D_v} \overline{\lambda}_i^n \Pi_i^{n,j} = \Pi_i^{n,j}; 2 \leq i \leq D_v, 1 \leq j \leq J$. We then have

$$\frac{R^{n,j}}{c(\epsilon^j)} = \frac{R^{n,0}/(1 - \Pi^{n,j})}{1 - \epsilon^j} = \frac{R^{n,0}}{1 - \epsilon^0} = \frac{R^{n,0}}{c(\epsilon^0)},$$

where the 2nd equality is obtained based on (17). This proves (19). Now since the parent ensemble is capacity achieving, $\lim_{n\to\infty} R^{n,0} = c(\epsilon^0)$. Based on (19), this implies that $\lim_{n\to\infty} R^{n,j} = c(\epsilon^j)$ for any j. This proves the universally capacity achieving property. To see the rate-compatibility, similar to the argument in Corollary 3, one can see that $\Pi_i^{n,j}$ in (17) is a decreasing function of ϵ^j (increasing function of channel quality) for any n. Therefore, for $\epsilon^m < \epsilon^k$, we have $\Pi_i^{n,k} > \Pi_i^{n,m}$; $i \in \{2, ..., k, D_v\}$. This completes the proof. ■

This result is consistent with the one obtained in [17] stating that random puncturing of a parent ensemble over the BEC preserves the distance to capacity. The approach taken in [17] is, however, different and is based on the fact that one can model the puncturing of an ensemble over the BEC as the concatenation of the original BEC channel with another BEC channel with erasure rate equal to puncturing. Similar to the flatness condition, the approach of [17] is not extendable to other BIOSM channels. The importance of our approach is that in principle, it may be extendable to other BIOSM channels where we can expect that applying the SM principle to compute Π_i values, might also result in (a scheme performing close to) a universally capacity achieving rate-compatible scheme. Unlike the BEC case, however, the upper bounds on Π_i have to be estimated numerically (similar to the procedure we use to compute the upper bounds of $\lambda_i, i > 2$, for the unpunctured case) except for Π_2 whose upper bound is given by Theorem 2. Applying this procedure to the capacity approaching ensembles designed based on the method of Section III as parent ensembles, we have

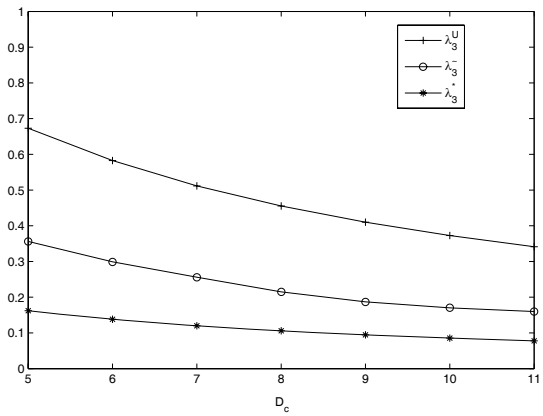

Fig. 1. The values of λ_3^U,λ_3^* and $\widetilde{\lambda_3}$ vs. D_c for a check-regular ensemble over the BIAWGN channel with $\sigma = .9574$.

in-fact been able to design universally capacity approaching rate-compatible ensembles over other BIOSM channels.

It is important to note that the values of $\Pi_i^{j,n}$ in Theorem 3 do not depend on i and n. While independency of i is a special property for the BEC, based on Corollary 2 these values are independent from n for $i = 2$ over any BIOSM channel. Our numerical results show that for a given $i > 2, i \neq D_v$ and j, the values of $\Pi_i^{j,n}$ are very close for different values of n, suggesting a general independency from n.

VI. Design Examples

We first present a plot of the upper and lower bounds on λ_3 related to Theorem 1, namely λ_3^U and λ_3^* as well as the numerically calculated values for $\widetilde{\lambda_3}$ for different values of D_c. The results are for check-regular ensembles over a BIAWGN channel with the noise standard deviation of $\sigma = .9574$ (capacity=1/2). As can be seen in Fig. 1, $\widetilde{\lambda_3}$ is a strictly decreasing function of D_c similar to $\widetilde{\lambda_2}$.

Example 1: We apply our method to design ensembles for a BIAWGN channel with capacity 1/2 ($\sigma = .9574$). As a point of reference, we consider the 4th ensemble of Table II in [10] with maximum variable node degree of 50, refereed to as C_{AWGN}. This ensemble has the best performance among the rate one-half ensembles designed in [10] for the BIAWGN channel. The check node degree distribution of this ensemble is $\rho_{AWGN}(x) = .33620x^8 + .0883x^9 + .57497x^{10}$. The threshold of this ensemble is $\sigma = .9718$. The capacity of a channel with $\sigma = .9718$ is equal to .5045, implying that the rate of this ensemble is 99.1% of the capacity. Keeping the same check node distribution, and setting k, the number of constituent variable node degrees to 24, we design the following variable node degree distribution based on the SM method:[7]

$$\lambda(x) = .1826x + .1602x^2 + .0297x^3 + .0306x^4 + .0306x^5 + .0307x^6 + .0306x^7 + .0297x^8 + .0296x^9 + .0288x^{10} + .0288x^{11} + .0279x^{12} + .0225x^{13} + .0171x^{14} + .0135x^{15} + .0099x^{16} + .0081x^{17} + .0054x^{18} + .0045x^{19} + .0036x^{20} + .0027x^{21} + .0018x^{22} + .0018x^{23} + .0009x^{24} + .2460x^{59}.$$

[7]The details of the design algorithm and some numerical issues are discussed in the appendix.

TABLE I
Performance of a Check-Regular Sequence Design Based on the SM Method Over a BIAWGN Channel With $\sigma = .9557$

D_c	$R_{AWGN}/c(.9557)$	k
5	.8902	3
6	.9386	3
7	.9520	4
8	.9653	6
9	.9756	10
10	.9884	18

The rate of this ensemble is 0.4950 which is 99.0% of the capacity, showing almost the same distance to capacity as C_{AWGN}. The disadvantage of this ensemble compared to C_{AWGN}, is having a larger maximum variable node degree and larger number of constituent variable node degrees. The average check node degree for this ensemble is about 10.1569.

We applied the SM method to design a check-regular ensemble with $D_c = 10$ (with the same values for σ and k), and we were able to design an ensemble whose rate was also 99.0% of the capacity. This suggests that at least for the designs based on SM, the important factor that determines the ensemble performance, is the average check node degree rather than the actual check node degree distribution. In other words, no optimization on the check node side would be necessary.

Example 2: Consider the following sequence design of check-regular ensembles for channel parameter $\sigma = .9557$. We start with $D_c = 5$ and $k = 3$, and for $D_c > 5$, we set $k = 2^{Dc-6} + 2$. This means that the number of constituent variable node degrees for an ensemble with check node degree D_c is roughly twice that of an ensemble with check node degree $D_c - 1$. As can be seen in Table I, the performance of the ensembles consistently improves as the average check node degree increases. The performance of the ensemble with $D_c = 10$ in Table I is slightly less than 99% of the capacity.

Example 3: For the Binary Symmetric Channel (BSC), we consider the ensemble C_{BSC} designed based on optimization in Example 2 of [10]. This rate one-half ensemble has check node degree distribution $\rho_{BSC}(x) = .25x^9 + .75x^{10}$ and threshold $\theta = .106$. This implies that for check node degree distribution $\rho_{BSC}(x)$ and channel parameter $\theta = .106$, the best achievable rate based on optimization is 0.5. We now apply the SM method to design an ensemble with the same check node degree distribution and channel parameter. The designed ensemble has the following variable node degree distribution where we have limited the number of constituent variable node degrees k to 22:

$$\lambda(x) = .1666x + .1644x^2 + .0171x^3 + .0190x^4 + .0219x^5 + .0228x^6 + .0238x^7 + .0257x^8 + .0266x^9 + .0285x^{10} + .0304x^{11} + .0323x^{12} + .0352x^{13} + .0390x^{14} + .0380x^{15} + .0314x^{16} + .0247x^{17} + .0200x^{18} + .0152x^{19} + .152x^{20} + .152x^{21} + .1874x^{65}.$$

The rate of this ensemble is .4988 which is very close

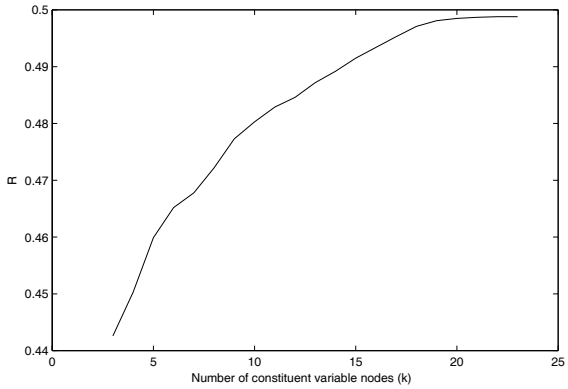

Fig. 2. The rate of the ensemble designed based on the SM method over the BSC with channel cross over probability .106 and check node distribution $\rho_{BSC}(x)$ versus the number of constituent variable node degrees.

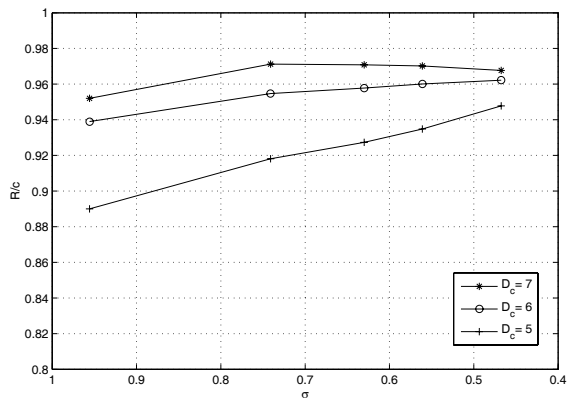

Fig. 3. The ratio of rate to capacity (R/C) for the rate-compatible ensemble sequence of Table II.

to .5, the best achievable rate based on optimization. The values of k and D_v for the designed ensemble are also close to (and smaller than) that of ensemble C_{BSC}. This code is able to achieve 97.4% of the capacity. This example suggests that the speed of convergence to capacity with respect to the average check node degree is faster for the BIAWGN channel compared to the BSC. The speed of convergence to capacity is considerably higher for the BEC. For the capacity value of 1/2 as an example, with $D_c = 10$, one is able to achieve 99.8% of the capacity of the BEC [12].

One advantage of our method is its simple implementation. Taking advantage of this property, we investigate the effect of the number of constituent variable node degrees k on the achievable code rate. We again consider the check node degree distribution $\rho_{BSC}(x)$ and for different values of k from 3 to 24, design ensembles using the SM method. In Fig. 2, we have plotted the rate of the designed ensembles versus k. As can be seen, for $k > 22$, the curve starts to saturate, implying that there is not any advantage of choosing k greater than about 22. The values of maximum variable node degrees for the designed ensembles range from 23 to 81. For the rate-compatible codes, we consider the sequence of Table I and puncture the first three ensembles for a set of four channels with noise powers smaller than that of the parent ensemble. The details are provided in Table II, where we have the puncturing polynomial $\Pi(x) = \sum_{i=2}^{D_v} \Pi_i x^{i-1}$ to represent the puncturing fractions in the last four columns. In Fig. 3, we have plotted the ratio $R^j/c(\theta^j)$ of the ensembles of Table II versus the channel noise standard deviation σ. As can be seen, the performance of the punctured codes for a given parent ensemble is similar to or better than the parent ensemble. In fact, we expect the punctured ensemble to perform almost the same as the parent ensemble, similar to the case of the BEC. To justify the improvement resulting from puncturing, we note that although the parent ensembles have been constructed based on the SM method, for finite values of D_c, they are not necessarily optimal in that they may not provide the best possible rate for the given channel parameter. This leaves the door open for further improvement with puncturing. From Fig. 3, it is also observed that for any given channel parameter, the performance of punctured ensembles approaches the capacity as the average check node degree increases. Based on Table II, the designed sequence also fulfills the rate-compatibility property. Note however that unlike, for example the approach of [16], we did not impose any constraint to guarantee rate-compatibility and our empirical results suggest that this property is inherent in the proposed method. For the case of Π_2, we analytically proved this fact in Corollary 3. To compare the performance of the schemes designed based on the SM principle and those obtained by optimization, we consider the ensemble used in [16] as a reference. This ensemble (C) has been optimized for the rate one half and has a threshold of $\sigma = .9557$. We can assume that ensemble C has been optimized for the highest rate when the channel parameter σ is set to .9557. The degree distribution of C is:

$$\lambda_C(x) = .25105x + .30094x^2 + .00104x^3 + .43853x^9,$$

$$\rho_C(x) = .63676x^6 + .36324x^7 .$$

Keeping the check node degree distribution of ensemble C intact, we design an ensemble C_{SM} with the same number of constituent variable nodes using the SM method:

$$\lambda_{SM}(x) = .2717x + .2442x^2 + .0371x^3 + .4471x^9.$$

We then apply the SM method again, this time to puncture C_{SM}. The puncturing polynomials for the same four channels considered in Table II are given in Table III. The distance to capacity (in dB) for the parent ensemble and its punctured versions is reported in Fig. 4. As can be seen in Fig. 4, the scheme performs very closely to the scheme obtained by optimization-based puncturing of the ensemble C. In fact, the proposed scheme even slightly outperforms the scheme of [16] on channels with $\sigma = .6300$ and $\sigma = .7410$. The proposed scheme performs inferior only on the best channel parameter ($\sigma = .4675$) and even for this channel parameter, the performance gap is less than .08 dB.[8] We have also demonstrated the performance of random puncturing of the ensemble C for comparison. Also note again that unlike [16], we did

[8]Note that our parent code itself performs close to .1dB worse than C and the gap in performance is always less than this gap for different puncturing rates.

TABLE II
THE VALUES OF Π_i USED TO PUNCTURE THE FIRST 3 ENSEMBLES OF THE SEQUENCE OF TABLE I

D_c	$\sigma = 0.9557$	$\sigma = 0.7410$	$\sigma = 0.6300$	$\sigma = 0.5609$	$\sigma = 0.4675$
5	$\lambda(x) = .4322x$ $+.3534x^2 + .2144x^5$	$\Pi(x) = .2947x+$ $.2037x^2 + .2231x^5$	$\Pi(x) = .4115x+$ $.2697x^2 + .3473x^5$	$\Pi(x) = .4703x+$ $.2949x^2 + .4676x^5$	$\Pi(x) = .5308x+$ $.3026x^2 + .7460x^5$
6	$\lambda(x) = 0.3457x$ $+0.2974x^2 + 0.3569x^6$	$\Pi(x) = .2947x+$ $.2124x^2 + .2357x^6$	$\Pi(x) = .4115x+$ $.2755x^2 + .3483x^6$	$\Pi(x) = .4703x+$ $.2997x^2 + .4142x^6$	$\Pi(x) = .5308x+$ $.3065x^2 + .5015x^5$
7	$\lambda(x) = 0.2881x$ $+0.2556x^2 + 0.0380x^3$ $+0.4183x^9$	$\Pi(x) = .2947x+$ $.1921x^2 + .3240x^3$ $+.2862x^9$	$\Pi(x) = .4115x+$ $.2590x^2 + .4472x^3$ $+.3638x^9$	$\Pi(x) = .4703x+$ $.2862x^2 + .5063x^3$ $+.4094x^9$	$\Pi(x) = .5308x+$ $.2997x^2 + .5772x^3$ $+.5950x^9$

TABLE III
THE VALUES OF Π_i'S USED TO PUNCTURE THE ENSEMBLE C_{SM}

$\sigma = 0.7410$	$\Pi(x) = 0.2947x + .1921x^2 + .3376x^3 + .2222x^9$
$\sigma = 0.6300$	$\Pi(x) = 0.4115x + .2600x^2 + .4462x^3 + .3470x^9$
$\sigma = 0.5609$	$\Pi(x) = 0.4703x + .2871x^2 + .5063x^3 + .3628x^9$
$\sigma = 0.4675$	$\Pi(x) = 0.5308x + .3007x^2 + .5820x^3 + .4200x^9$

not impose any constraint to guarantee rate-compatibility. This reduces the design complexity significantly.

It is interesting to see that based on Tables II and III, except for $i = D_v$, the values of Π_i are almost independent (for Π_2 provably independent based on Corollary 2) of the parent ensembles and only depend on the channel parameter for which the puncturing is applied. In other words, for a given channel parameter θ^j, the computed values of Π_i can universally be applied to any ensemble designed based on the SM method for a given original channel parameter θ^0 and any arbitrary check node distribution.

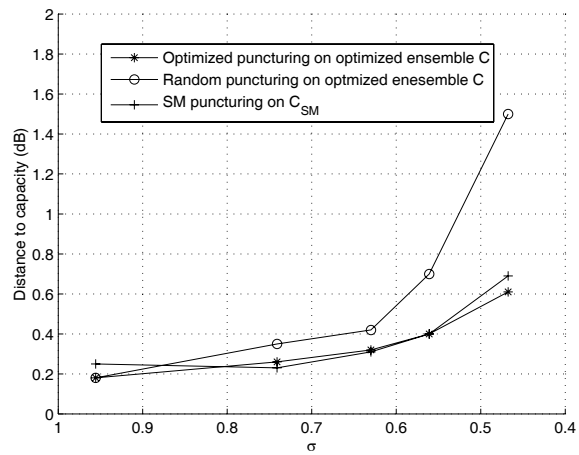

Fig. 4. Performance comparison of schemes constructed based on the proposed SM method and those constructed based on optimization method of [16].

VII. CONCLUSION

In this paper, we proposed the method of successive maximization (SM) for the systematic design of universally capacity approaching rate-compatible LDPC code ensemble sequences over BIOSM channels. The SM method was first applied to design a sequence of capacity approaching parent ensembles. It was then applied to each parent ensemble, this time to design rate-compatible puncturing schemes. As part of our results, we were able to extend the stability condition which was previously derived for degree-2 variable nodes to other variable node degrees as well as to the case of rate-compatible codes. Consequently, we rigorously proved that using the SM principle, one is able to design universally capacity achieving rate-compatible LDPC code ensemble sequences over the BEC. Unlike the previous results on such schemes over the BEC in the literature, the proposed SM approach can be naturally extended to other BIOSM channels. Using such an extension, we designed rate-compatible codes over the BIAWGN channel and the BSC whose performance universally approaches the capacity as the average check node degree increases. We demonstrated that for finite values of D_c, the performance of the ensembles designed by our method is comparable to those designed based on optimization. One important direction in the continuation of this work is to analytically compute the values of $\widetilde{\lambda_i}$ or to devise design algorithms which are more robust against numerical errors. This can pave the road for demonstrating that the proposed sequences can in fact achieve the capacity of BIOSM channels.

APPENDIX
DESIGN ALGORITHM AND NUMERICAL STABILITY ISSUES

The design method of Section III can be formulated into the following algorithm which will be referred to as Algorithm 1. Let k be the number of constituent variable node degrees. Starting from $i = 2$, we need to find $\widetilde{\lambda_i} = \sup\{\Lambda_i(\rho(x), \theta^0)\}$ for different values of i successively. We know that $\widetilde{\lambda_2} = \lambda_2^*$, and thus start from $i = 3$. The calculation of $\widetilde{\lambda_i}$ can be performed by continuously increasing the value of λ_i from zero and in each step, checking if the ensemble with sufficiently large D_v is convergent using density evolution. In each step i, we always set $\lambda_{D_v} = 1 - \sum_{j=2}^{i} \lambda_j$. We repeat this process for successive values of i until either $\lambda_{D_v} < 0$ for a given value of i, or $i > k + 1$. Then, we decrease the value of D_v as far as the ensemble remains convergent. We, however, remind the reader that since the computations are performed in critical values of λ_i (i.e., at the border of stability/instability), this algorithm is very sensitive to numerical errors. To mitigate the effect of such errors, one has to use density evolution with very high precision as well as very small increments in the values of λ_i in the vicinity of the threshold $\widetilde{\lambda_i}$. This in turn, increases the computational complexity. Consequently, reducing the precision may result in numerical errors which usually propagate to other steps. This issue will be discussed in this appendix.

Concentrating on the computation of $\widetilde{\lambda_3}$, we note that this computation directly depends on the value of λ_2 and is performed under the assumption that $\lambda_2 = \widetilde{\lambda_2}$. The important

question, however, is that what happens if the value of λ_2 is slightly different from $\widetilde{\lambda_2}$ due to numerical errors. Is the upper bound on λ_3 a continuous function of λ_2 such that the computed upper bound for λ_3 tends to $\widetilde{\lambda_3}$ if λ_2 tends to $\widetilde{\lambda_2}$? Based on our numerical results, the answer to this question is positive. In fact, we have been able to prove this for the case of the BEC in the next proposition. We conjecture that such continuity also exists for other channels as well as other variable node degrees.

Proposition 3: Over the BEC, consider a code-channel pair $E = (\lambda_2, \lambda_3, D_v, \rho(x); \epsilon)$ where D_v can be made arbitrarily large. For any given value of λ_2 in $\Lambda_2(\rho(x), \epsilon)$, define the set $A = \{\lambda_3 : \exists D_v, E \text{ is convergent}\}$ and $\Im(\lambda_2) = \sup(A)$. Then, $\Im$ is a continuous function of λ_2. In particular, $\Im$ is continuous at $\lambda_2 = \widetilde{\lambda_2}$.

Proof: Using the convergence condition (6) for $k = 3$ and rearranging the terms, we have

$$\lambda_3 < \sum_{i=3}^{\infty} \frac{T_i}{\epsilon} x^{i-3} + \frac{\widetilde{\lambda}_2 - \lambda_2}{x} - \frac{1 - \lambda_2 - \lambda_3}{\epsilon} x^{D_v - 3}. \quad (20)$$

Now note that if $\lambda_3 \in A$ for a certain D_v, for any greater value of D_v, we still have $\lambda_3 \in A$. Therefore the value of $\Im(\lambda_2)$ does not change if D_v tends to infinity. Therefore we have:

$$\Im(\lambda_2) = \min_{0<x<x_0} \left\{ \sum_{i=3}^{\infty} \frac{T_i}{\epsilon} x^{i-3} + \frac{\widetilde{\lambda}_2 - \lambda_2}{x} \right\},$$

where $x_0 = 1 - \rho(1 - \mathfrak{P}(P_0))$, and we have neglected the last term of (20) assuming that D_v tends to infinity. First, we prove that $\lim_{\lambda_2 \to \widetilde{\lambda}_2} \Im(\lambda_2) = \Im(\widetilde{\lambda_2}) = \widetilde{\lambda_3}$. We define $f_1(x) = \sum_{i=3}^{\infty} \frac{T_i}{\epsilon} x^{i-3} + \frac{b}{x}$, where $b = \widetilde{\lambda}_2 - \lambda_2, 0 \le b \le \widetilde{\lambda}_2$. For $b = 0$ ($\lambda_2 = \widetilde{\lambda_2}$), the minimum of this function in the interval $[0, \infty)$ is at $x = 0$. Now assume that $b > 0$. The second derivative of this function is strictly positive in the interval $(0, \infty)$. In other words, this function has a local minimum in this interval which is the root of equation $f_1'(x) = 0$ where $f_1'(x) = \sum_{i=4}^{\infty} \frac{T_i}{\epsilon}(i-3)x^{i-4} - \frac{b}{x^2}$. This equation has only one root in the interval of $(0, \infty)$. It can easily be seen that if b tends to zero, the root of the equation also tends to zero. In other words, $argmin(f_1(x))$ tends to zero as b tends to zero. Thus we can conclude that $\Im(\lambda_2)$ is continuous at $\lambda_2 = \widetilde{\lambda}_2$. The proof of continuity for other points is straight forward. ■

As previously mentioned, Algorithm 1 requires very high precision to mitigate the effect of numerical errors. The reason is that the value of $\widetilde{\lambda}_i$ is very sensitive to the value of $\widetilde{\lambda}_{i-1}$. For example, although we proved in Proposition 3 that for the BEC, function $\Im$ is continuous at $\lambda_2 = \widetilde{\lambda}_2$, one can verify that $\frac{d}{d\lambda_2}\Im(\lambda_2)|_{\lambda_2 = \widetilde{\lambda}_2}$ can be very large, implying that any small deviation from $\widetilde{\lambda}_2$ will cause a significant deviation from the value of $\widetilde{\lambda}_3$. Moreover, the computation of the exact value of $\widetilde{\lambda}_{i-1}$ by density evolution requires very high precision. For example, using a similar method to that of Algorithm 1 to compute $\widetilde{\lambda}_2$ with a reasonable complexity (in our case, dynamic range of [-50,50] and 13-bit quantization) would result in a value for $\widetilde{\lambda}_2$ which is non-negligibly different from (and usually greater than) λ_2^* (the difference can some times be as high as half a percent of $\widetilde{\lambda}_2$). Now if we set $\lambda_2 = \widetilde{\lambda}_2 = \lambda_2^*$ and compute the value of $\widetilde{\lambda_3}$ and subsequently the value of $\widetilde{\lambda_4}$, the computed value for $\widetilde{\lambda}_4$ will appear to be close to zero. This, however, is not the correct value, at least for the case of the BEC where we already know that $\widetilde{\lambda}_4 = \lambda_4^*$. The close to zero value of $\widetilde{\lambda}_4$ is caused by the fact that the computed value of $\widetilde{\lambda}_3$ is sightly larger than its true value (this is confirmed for the case of the BEC where we already know the true value of $\widetilde{\lambda_3}$). To prevent this, we need to slightly reduce the value of λ_3 from its computed upper bound. At the same time, the amount of reduction in the value of λ_3 is very critical and may make the computed value for $\widetilde{\lambda}_4$ too large. This in turn will cause the computed value for $\widetilde{\lambda}_5$ in the next step to be close to zero. In general, a numerical error at one step propagates to the following steps. One way to prevent λ_{i+1} to tend to zero at step $i+1$ due to an over-estimated value of λ_i in the previous step (step i), is to let λ_{i+1} to also increase to a fraction of λ_i while increasing λ_i. This joint increment can also be applied to more than 2 consecutive degree coefficients. This means that at step i, while increasing λ_i, we also increase $\lambda_{i+1}, .., \lambda_{i+K}$, to a fraction of λ_i. We then increase i by 1, and set all coefficients with index greater than i to zero, and repeat the process. At each step, after the maximization is performed, we still multiply the resulting value by a constant α less than or equal to 1.[9] Using this joint increment technique, the values of λ_i to λ_{i+K} at step i become less dependent on the value of α used at step $i-1$.

REFERENCES

[1] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, June 2002.

[2] M. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. Spielman, and V. Stemann, "Practical loss-resilient codes," in *Proc. 29th ACM Symposium Theory Computing*, 1997, pp. 150–159.

[3] A. Shokrollahi, "Capacity-achieving sequences," in *Codes, Systems, and Graphical Models (IMA Volumes in Mathematics and Its Applications)*, B. Marcus and J. Rosenthal, editors. Springer-Verlag, 2000, vol. 123, pp. 153–166.

[4] A. Shokrollahi, "New sequences of linear time erasure codes approaching the channel capacity," in *Proc. 13th International Symposium Applied Algebra, Algebraic Algorithms, Error Correcting Codes* (M. Fossorier, H. Imai, S. Lin, and A. Poli), no. 1719, Lecture Notes in Computer Science, pp. 65–76, 1999.

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

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

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

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

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

[10] T. J. Richardson, A. Shokrollahi, and R. L. Urbanke, "Design of capacity approaching irregular low density parity check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[9] In our simulations, we have used $\alpha = 0.95$ and K=2 which is applied for $i > 3$.

[11] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.
[12] H. Saeedi and A. H. Banihashemi, "New sequences of capacity achieving LDPC Code ensembles for the binary erasure channel," in *Proc. International Symp. Inf. Theory*, July 2008.
[13] H. Saeedi and A. H. Banihashemi, "New sequences of capacity achieving LDPC code ensembles over the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6332–6346, Dec. 2010.
[14] H. Saeedi and A. H. Banihashemi, "Systematic design of LDPC codes for binary erasure channels," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 118–127, Jan. 2010.
[15] H. Pishro-Nik and F. Fekri, "Results on punctured low-density parity-check codes and improved iterative decoding techniques," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 599–614, Feb. 2007.
[16] J. Ha, J. Kim, and S. W. McLaughlin, "Rate-compatible puncturing of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2824–2836, Nov. 2004.
[17] H. Pishro-Nik, N. Rahnavard, and F. Fekri, "Nonuniform error correction using low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2702–2714, July 2005.
[18] H. Chun-Hao and A. Anastasopoulos, "Capacity achieving LDPC codes through puncturing," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, Oct. 2008.
[19] I. Sason and G. Wiechman, "On achievable rates and complexity of LDPC codes over parallel channels: bounds and applications," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 580–598, Feb. 2007.
[20] J. Ha, J. Kim, D. Klinc, S. W. McLaughlin, "Rate-compatible punctured low-density parity-check codes with short block lengths," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 728–738, Feb. 2006.
[21] B. N. Vellambi and F. Fekri, "Finite-length rate-compatible LDPC codes: a novel puncturing scheme," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 297–301, Feb. 2009.
[22] M. R. Yazdani and A. H. Banihashemi, "On construction of rate-compatible low-density parity-check codes," *IEEE Commun. Lett.*, vol. 8, no. 3, pp. 159–161, Mar. 2004.
[23] Y. Guosen Yue, W. Xiaodong, and M. Madihian, "Design of rate-compatible irregular repeat accumulate codes," *IEEE Trans. Commun.*, vol. 55, no. 6, pp. 1153–1163, June 2007.
[24] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, June 2006.
[25] H. Saeedi and A. H. Banihashemi, "On the design of LDPC codes ensembles for BIAWGN channels," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1376–1382, May 2010.
[26] C. Measson, A. Montanari, and R. Urbanke. "Why we can not surpass capacity: the matching condition," in *Proc. Allerton Conf. Commun., Control Computing*, Oct. 2005.
[27] I. Sason, "On universal properties of capacity-approaching LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 2956–2990, July 2009.
[28] E. Arikan, "Channel polarization: a method for constructing capacity-achieving codes," in *Proc. International Symp. Inf. Theory*, July 2008.
[29] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2657–2673, Nov. 2004.
[30] H. Pishro-Nik and F. Fekri, "Performance of low-density parity-check codes with linear minimum distance," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 292–300, Jan. 2006.

Hamid Saeedi (S'01-M'08) received the B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran, in 1999 and 2001, respectively, and the Ph.D. degree from Carleton University, Ottawa, ON, Canada, in 2007, all in electrical engineering. In 2008–2009, he was a postdoctoral fellow in the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst. In 2010, he joined the Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran, where he is now an Assistant Professor. He is also affiliated with the Advanced Communications Research Institute (ACRI), Sharif University of Technology, Tehran, Iran. His research interests include coding and information theory, wireless communications, and signal processing for broadband communications.

Dr. Saeedi is the recipient of several awards including the Carleton University Senate Medal for Outstanding Academic Achievement, a Natural Sciences and Engineering Council of Canada (NSERC) Industrial Research and Development Fellowship, and an Ontario Graduate Scholarship.

Hossein Pishro-Nik (S'01-M'06) is an Associate Professor of electrical and computer engineering at the University of Massachusetts, Amherst. He received a B.S. degree from Sharif University of Technology, and M.Sc. and Ph.D. degrees from the Georgia Institute of Technology, all in electrical and computer engineering. His research interests include the mathematical analysis of communication systems, in particular, error control coding, wireless networks, and vehicular ad hoc networks. His awards include an NSF Faculty Early Career Development (CAREER) Award, an Outstanding Junior Faculty Award from UMass, and an Outstanding Graduate Research Award from the Georgia Institute of Technology.

Amir H. Banihashemi (S'90-A'98-M'03-SM'04) received the B.A.Sc. degree in electrical engineering from Isfahan University of Technology, Isfahan, Iran in 1988, and the M.A.Sc. degree in communications engineering from the the University of Tehran, Tehran, Iran, in 1991, with the highest academic rank in both classes.

From 1991 to 1994, he was with the Electrical Engineering Research Center and the Department of Electrical and Computer Engineering, Isfahan University of Technology. During 1994–1997, he was with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, working towards the Ph.D. degree. He held two Ontario Graduate Scholarships for international students during this period. In 1997, he joined the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada, where he worked as a Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellow. He joined the Faculty of Engineering at Carleton University in 1998, where at present he is a Professor in the Department of Systems and Computer Engineering. His research interests include coding and information theory, digital and wireless communications, theory and implementation of communications algorithms, and compressive sensing and sampling. He has published more than 130 papers in refereed journals and conferences.

Dr. Banihashemi served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2003 to 2009. He is a member of the Board of Directors for the Canadian Society of Information Theory and the Director of the Broadband Communications and Wireless Systems (BCWS) Centre at Carleton University. He has been involved in many international conferences as chair, member of technical program committee, and member of organizing committee. These include co-chair for the Communication Theory Symposium of Globecom'07, TPC co-chair of the Information Theory Workshop (ITW) 2007, member of the organizing committee for the International Symposium on Information Theory (ISIT) 2008, and co-chair of the Canadian Workshop on Information Theory (CWIT) 2009. Dr. Banihashemi is a recipient of Carleton's Research Achievement Award in 2006, and is awarded one of the NSERC's one hundred 2008 Discovery Accelerator Supplements (DAS).