Accumulate-Repeat-Accumulate Codes: Capacity-Achieving Ensembles of Systematic Codes for the Erasure Channel with Bounded Complexity

Henry D. Pfister, *Member*, Igal Sason, *Member*

Abstract — This paper introduces ensembles of systematic accumulate-repeat-accumulate (ARA) codes which asymptotically achieve capacity on the binary erasure channel (BEC) with *bounded complexity*, per information bit, of encoding and decoding. It also introduces symmetry properties which play a central role in the construction of new capacity-achieving ensembles for the BEC. The results here improve on the tradeoff between performance and complexity provided by previous constructions of capacity-achieving code ensembles defined on graphs. The superiority of ARA codes with moderate to large block length is exemplified by computer simulations which compare their performance with those of previously reported capacity-achieving ensembles of LDPC and IRA codes. The ARA codes also have the advantage of being systematic.

Index Terms — Binary erasure channel (BEC), capacity, complexity, degree distribution (d.d.), density evolution (DE), iterative decoding, irregular repeat-accumulate (IRA) codes, low-density parity-check (LDPC) codes, systematic codes.

I. INTRODUCTION

Error-correcting codes which employ iterative decoding algorithms now represent the state of the art in low-complexity coding techniques. There is already a large collection of iteratively decodable codes including low-density parity-check (LDPC), turbo, repeat-accumulate and product codes; all of these demonstrate a rather small gap (in rate) to capacity with feasible complexity [1].

The study of capacity-achieving (c.a.) sequences of LDPC codes for the binary erasure channel (BEC) was initiated by Luby et al. [2] and Shokrollahi [3]. They show that it is possible to closely approach the capacity of an erasure channel with a simple iterative procedure whose complexity is linear in the block length of the code [2], [3]. Following these works, Oswald and Shokrollahi presented in [4] a systematic study of c.a. sequences of LDPC codes for the BEC. Jin et al. introduced irregular repeat-accumulate (IRA) codes and presented a c.a. sequence of systematic IRA (SIRA) codes for the BEC [5]. A sequence of c.a. SIRA codes for the BEC with

lower encoding and decoding complexities was introduced in [6, Theorem 2]. All of the aforementioned codes have one drawback in common: their decoding complexity scales like the log of the inverse of the gap (in rate) to capacity [3], [4], [6], [7], [8], [9]; hence, under iterative message-passing decoding, these codes have *unbounded complexity* (per information bit) as the gap to capacity vanishes.

In [10], the authors presented for the first time two sequences of ensembles of non-systematic IRA (NSIRA) codes which asymptotically (i.e., as their block length tends to infinity) achieve capacity on the BEC with *bounded complexity* per information bit. This new result is achieved by puncturing bits and thereby introducing state nodes in the Tanner graph representing the codes. We note that for fixed complexity, these codes will eventually (for large enough block length) outperform any code proposed so far. However, the speed of convergence happens to be quite slow and, for small to moderate block lengths, the codes introduced in [10] are not record breaking.

In this paper, we are interested in the construction and analysis of c.a. codes for the BEC with bounded complexity that also perform well at moderate block lengths. We also would also like these codes to be systematic and to have reasonably low error floors. To this end, we make use of a new channel coding scheme, called "Accumulate-Repeat-Accumulate" (ARA) codes, which was recently introduced by Abbasfar et al. [11]. These codes are systematic and have both outstanding performance, as exemplified in [11], [12], [13], and a simple linear-time encoding. After presenting an appropriate ensemble of irregular ARA codes, we construct a number of c.a. degree distributions. Simulations show that some of these ensembles perform quite well on the BEC at moderate block lengths. We therefore expect that irregular ARA codes, optimized for general channels, also perform well at moderate block lengths (as is partially supported by some simulation results in [11]). This issue is regarded as a topic for further research, while this paper is focused on the BEC. Throughout the paper, we consider the encoding and decoding complexity *per information bit*.

Along the way, we study some symmetry properties of c.a. sequences for the BEC and discover a new code structure which we call "Accumulate-LDPC" (ALDPC) codes. We show that c.a. degree distributions for this structure can be easily constructed based on the results of [10, Theorems 1, 2]. This fact and structure was proposed independently by Hsu and Anastasopoulos [14].
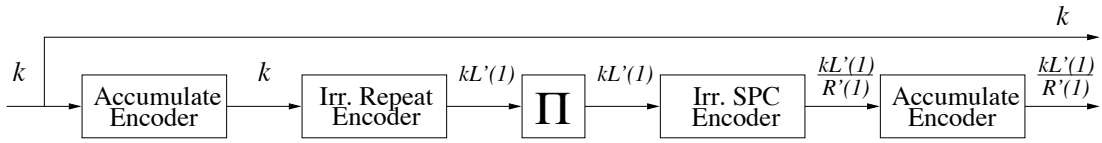
Fig. 1. Systematic encoder for the ARA ensemble ('Irr.' stands for 'irregular', 'SPC' means 'single-parity check', and Π represents a bit interleaver.)

The paper is organized as follows: Section II introduces ARA codes, describes their encoding and decoding, and provides a density evolution analysis for the BEC. Section III introduces symmetry properties that play a central role in the construction of c.a. sequences of ensembles for the BEC. Section IV serves as a preparatory step towards the construction of explicit c.a. sequences of ARA codes for the BEC, where their complexity of encoding and decoding stays bounded as the gap to capacity vanishes. Section V presents explicit constructions of c.a. sequences of bit-regular and check-regular ARA codes with bounded complexity. Section VI focuses on the construction of c.a. ensembles of ARA, NSIRA and ALDPC codes (with bounded complexity) based on the ensembles of self-matched LDPC codes introduced in the section. Computer simulations for the BEC are presented in Section VII, and the superiority of self-matched ARA codes with moderate to large block length is exemplified by comparing their performance with those of previously reported c.a. ensembles of LDPC and IRA codes from [3], [10]. Finally, Section VIII concludes our discussion.

II. ACCUMULATE-REPEAT-ACCUMULATE CODES

In this section, we present our ensemble of ARA codes. Density evolution (DE) analysis of this ensemble is presented in the second part of this section using two different approaches which lead to the same "DE fixed point equation"; this equation characterizes the fixed points of the iterative message-passing decoder. The connection between these two approaches is used later in this paper to state some symmetry properties which serve as an analytical tool for designing various c.a. ensembles for the BEC (e.g., ARA, IRA and ALDPC codes).

A. Description of ARA Codes

ARA codes can be viewed either as interleaved serially concatenated codes (i.e., turbo-like codes) or as sparse-graph codes (i.e., LDPC-like codes). From an encoding point of view, it is more natural to view them as interleaved serially concatenated codes (see Fig. 1) where the encoding process is described in Section II-B.

Since the decoding algorithm of ARA codes is simply belief propagation on the appropriate Tanner graph (see Fig. 2), this leads one to view them as sparse-graph codes from a decoding point of view. Treating these codes as sparse-graph codes also allows one to build large codes by "twisting" together many copies of a single small *protograph* [15], [16]. In general, this approach leads to very good codes with computationally efficient decoders.

In this work, we consider the ensemble of irregular ARA codes which is the natural generalization of the IRA codes from [5]. The ensemble of irregular ARA codes differs slightly from those proposed in [11], [12], [13]. For this ensemble, we find that DE for the BEC can be computed in closed form and that algebraic methods can be used to construct c.a. sequences.

B. Encoding of ARA Codes

We describe here briefly the encoding process of the ARA codes in Fig. 1. The encoding of ARA codes is done as follows: first, the information bits are accumulated (i.e., differentially encoded), and then the bits are repeated a varying number of times (by an irregular repetition code) and interleaved. The interleaved bits are partitioned into disjoint sets (whose size may also vary), and the parity of each set of bits is computed. Finally, the bits are accumulated for the second time. A codeword of a systematic ARA code is composed of the information bits and the parity bits at the output of the second accumulator.

Some slight modifications are used later for our simulations and these details are explained in Section VII. In this section, all references to the decoding graph should be taken to imply Fig. 2, and all sums are assumed to be modulo-2.

We will refer to the three layers of bit nodes in the decoding graph as systematic bits, punctured bits, and parity bits (the parity bits are named as "code bits" in Fig. 2). Referring to the Tanner graph of ARA codes, we designate the systematic bits from left to right by $(u_1, u_2, \ldots, u_k)$. The same convention is used for the punctured bits $(v_1, v_2, \ldots, v_k)$ and the parity bits $(z_1, z_2, \ldots, z_{n-k})$.

From the upper part of the graph, it follows that $v_j = u_j + v_{j-1}$ for $j \in \{2, \ldots, k\}$ and $v_1 = u_1$. This yields that

$$v_j = \sum_{i=1}^{j} u_i \qquad j = 1, 2, \ldots, k. \qquad (1)$$

Let $d(i)$ be the degree of the i-th "parity-check 2" node where the degree is w.r.t. the edges connecting the "punctured bit" nodes and the "parity-check 2" nodes, and $c(i, j)$ be the index of the punctured bit attached to the j-th edge of the i-th "parity-check 2" node. All the connections between the "punctured bit" nodes and "parity-check 2" nodes are described by these two sequences. Let the sequence $(w_1, w_2, \ldots, w_{n-k})$ be defined by

$$w_i = \sum_{j=1}^{d(i)} v_{c(i,j)} \qquad i = 1, 2, \ldots, n-k.$$

This can be thought as the sum of the punctured bits which are connected to the i-th "parity-check 2" node. From the lower
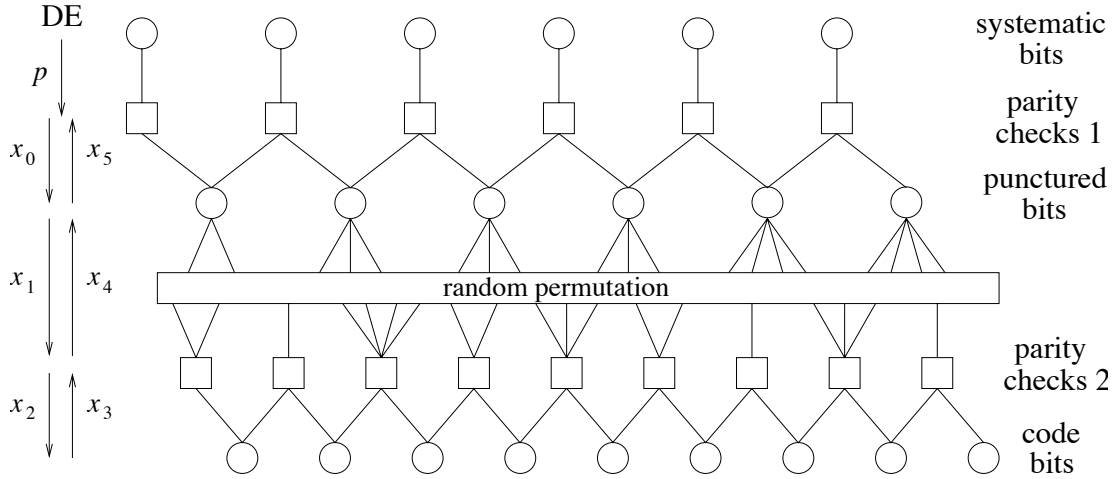
Fig. 2. Tanner graph for the ARA ensemble.

part of the graph we have $z_j = z_{j-1} + w_j$ (where $w_0 \triangleq 0$), and this gives

$$z_j = \sum_{i=1}^{j} w_i \qquad j = 1, 2, \ldots, n - k. \qquad (2)$$

From Fig. 2 and equations (1) and (2), one can see that an ARA code is the serial concatenation of four simple codes. The first is an accumulate code (upper part of the graph), the second is an interleaved irregular repetition code, the third is an irregular single parity-check (SPC) code (whose irregularity is due to the varying degrees of the "parity-check 2" nodes), and finally the fourth is a second accumulate code (lower part of the graph).

C. Density Evolution of Systematic ARA Ensembles

We consider here the asymptotic analysis of ensembles of ARA codes under the assumption that the codes are transmitted over a BEC and decoded with an iterative message-passing decoder. Based on the DE equations, derived in terms of the degree distributions of these ensembles, we consider the fixed points of the decoding process. In the following, we present two different approaches for the DE analysis of ARA codes for the BEC which, as expected, provide equivalent results. While the concept of the first approach is standard, the second one is helpful in establishing symmetry properties of c.a. ensembles for the BEC; these symmetries are discussed later in Section III.

1) Density Evolution via Message Passing: An irregular ensemble of ARA codes is defined by its degree distribution (d.d.). Nodes in the decoding graph will be referred to by the names given in Fig. 2. Let $L(x) = \sum_{i=1}^{\infty} L_i x^i$ be a power series where L_i denotes the fraction of "punctured bit" nodes with degree-i. Similarly, let $R(x) = \sum_{i=1}^{\infty} R_i x^i$ be a power series where R_i denotes the fraction of "parity-check 2" nodes with degree-i. In both cases, the degree refers only to the edges connecting the "punctured bit" nodes to the "parity-check 2" nodes. Similarly, let $\lambda(x) = \sum_{i=1}^{\infty} \lambda_i x^{i-1}$ and $\rho(x) = \sum_{i=1}^{\infty} \rho_i x^{i-1}$ form the d.d. pair from the edge

perspective where λ_i and ρ_i designate the fraction of the edges which are connected to "punctured bit" nodes and "parity-check 2" nodes with degree-i, respectively. We also assume that the permutation in Fig. 1 is chosen uniformly at random from the set of all permutations. The pair of degree distributions of an ARA ensemble is given by (λ, ρ).

It is easy to show the following connections between the d.d. pairs w.r.t. the nodes and edges in the graph:

$$\lambda(x) = \frac{L'(x)}{L'(1)}, \qquad \rho(x) = \frac{R'(x)}{R'(1)} \qquad (3)$$

or equivalently, since $L(0) = R(0) = 0$, then

$$L(x) = \frac{\int_0^x \lambda(t)\, dt}{\int_0^1 \lambda(t)\, dt}, \qquad R(x) = \frac{\int_0^x \rho(t)\, dt}{\int_0^1 \rho(t)\, dt}. \qquad (4)$$

The design rate R of the ensemble of ARA codes (see Fig. 1) is computed by expressing the block length n as the sum of k systematic bits and $kL'(1)/R'(1)$ parity bits which then yields

$$R = \frac{1}{1 + \frac{L'(1)}{R'(1)}}. \qquad (5)$$

A random code is chosen from the ensemble and a random codeword is transmitted over a BEC with erasure probability p. The asymptotic performance of the iterative message-passing decoder (as the block length of the code tends to infinity) is analyzed by tracking the average fraction of erasure messages which are passed in the graph of Fig. 2 during the lth iteration. The technique was introduced in [17] and is known as density evolution (DE). The main assumption of density evolution is that the messages passed on the edges of the Tanner graph are statistically independent. This assumption is justified by the fact that, for randomly chosen codes, the fraction of bits involved in finite-length cycles vanishes as the block length tends to infinity.

A single decoding iteration consists of six smaller steps which are performed on the Tanner graph of Fig. 2. Messages are first passed downward from the "systematic bit" nodes through each layer to the "code bit" nodes. Then, messages are passed back upwards from the "code bit" nodes through each

layer to the "systematic bit" nodes. Let l designate the iteration number. Referring to Fig. 2, let $x_0^{(l)}$ and $x_5^{(l)}$ designate the probabilities of an erasure message from the "parity-check 1" nodes to the "punctured bit" nodes and vice-versa, let $x_1^{(l)}$ and $x_4^{(l)}$ be the probabilities of an erasure message from the "punctured bit" nodes to the "parity-check 2" nodes and vice versa, and finally, let $x_2^{(l)}$ and $x_3^{(l)}$ be the probabilities of an erasure message from the "parity-check 2" nodes to "code bit" nodes and vice versa.

From the Tanner graph of ARA codes in Fig. 2, we see that an outgoing message from a "parity-check 1" node to a "punctured bit" node is an erasure if either the incoming message through the other edge (which connects a "punctured bit" node to the same "parity-check 1" node) is an erasure or the message received from the BEC for the systematic bit (which is connected to the same "parity-check 1" node) is an erasure. Using the statistical independence assumption, this yields the recursive equation

$$x_0^{(l)} = 1 - (1-p)\left(1 - x_5^{(l-1)}\right).$$

It is also clear from Fig. 2 that an outgoing message from a "punctured bit" node to a "parity-check 2" node is an erasure if and only if all the incoming messages passed through the other edges connected to this bit are erasures. The update rule of the iterative message-passing decoder on the BEC therefore implies that

$$x_1^{(l)} = \left(x_0^{(l)}\right)^2 \lambda\left(x_4^{(l-1)}\right).$$

From the graph in Fig. 2, we obtain in a similar manner the following DE equations of the iterative message-passing decoder:

$$
\begin{aligned}
x_0^{(l)} &= 1 - \left(1 - x_5^{(l-1)}\right)(1-p) \\
x_1^{(l)} &= \left(x_0^{(l)}\right)^2 \lambda\left(x_4^{(l-1)}\right) \\
x_2^{(l)} &= 1 - R\left(1 - x_1^{(l)}\right)\left(1 - x_3^{(l-1)}\right) \\
x_3^{(l)} &= p x_2^{(l)} \\
x_4^{(l)} &= 1 - \left(1 - x_3^{(l)}\right)^2 \rho\left(1 - x_1^{(l)}\right) \\
x_5^{(l)} &= x_0^{(l)} L\left(x_4^{(l)}\right)
\end{aligned}
$$

A fixed point is implied by

$$\lim_{l\to\infty} x_i^{(l)} \triangleq x_i \qquad i = 0, 1, \ldots, 5.$$

Now, we can solve for the fixed point by substituting x_2 into x_3, and then substituting the result into x_4 which gives the fixed point equation

$$x_4 = 1 - \left(\frac{1-p}{1 - p\,R(1-x_1)}\right)^2 \rho(1-x_1). \qquad (6)$$

Likewise, putting x_0 into x_5 gives the fixed point equation

$$x_5 = \frac{p\,L(x_4)}{1 - (1-p)\,L(x_4)}$$

and plugging this into x_0 gives

$$x_0 = 1 - (1-x_5)(1-p) = \frac{p}{1 - (1-p)L(x_4)}. \qquad (7)$$

Finally, Eqs. (6), (7) and the equality $x_1 = x_0^2 \lambda(x_4)$ give the following implicit equation for $x_1 \triangleq x$:

$$\frac{p^2\,\lambda\left(1 - \left(\frac{1-p}{1-pR(1-x)}\right)^2 \rho(1-x)\right)}{\left[1 - (1-p)\,L\left(1 - \left(\frac{1-p}{1-pR(1-x)}\right)^2 \rho(1-x)\right)\right]^2} = x. \qquad (8)$$

This equation provides the fixed points of the iterative message-passing decoder.

2) Density Evolution via Graph Reduction: For ensembles of ARA codes whose transmission takes place over a BEC, the DE fixed point equation (8) can be also derived using a *graph reduction* approach. This approach introduces two new operations on the Tanner graph which remove nodes and edges while preserving the information in the graph.
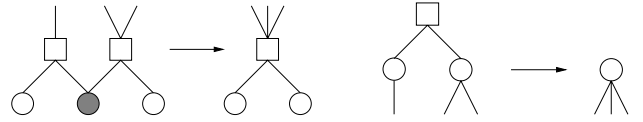

Fig. 3. Graph reduction operation applied to parity-check nodes (left) and bit nodes (right).

We start by noting that any "code bit" node whose value is not erased by the BEC can be removed from the graph by absorbing its value into its two "parity-check 2" nodes. On the other hand, when the value of a "code bit" node is erased, one can merge the two "parity-check 2" nodes which are connected to it (by summing the equations) and then remove the "code bit" node from the graph. This merging of two "parity-check 2" nodes causes their degrees to be summed and is shown on the left in Figure 3. Now, we consider the degree distribution (d.d.) of a single "parity-check 2" node in the reduced graph. This can be visualized as working from left to right in the graph, and assuming the value of the previous "code bit" node was known. The probability that there are k erasures before the next observed "code bit" is given by $p^k(1-p)$. The graph reduction associated with this event causes the degrees of $k+1$ "parity-check 2" nodes to be summed. The generating function for this sum of $k+1$ random variables, each chosen independently from the d.d. $R(x)$, is given by $R(x)^{k+1}$. Therefore, the new d.d. of the "parity-check 2" nodes after the graph reduction is given by

$$\widetilde{R}(x) = \sum_{k=0}^{\infty} p^k(1-p)R(x)^{k+1} = \frac{(1-p)R(x)}{1 - pR(x)}. \qquad (9)$$

A similar graph reduction can be also performed on the "systematic bit" nodes in Fig. 2. Since degree-1 bit nodes (e.g., the "systematic bit" nodes in Fig. 2) only provide channel information, erasures make them worthless. So they can be removed along with their parity-checks (i.e., the "parity-check 1" nodes in Fig. 2) without affecting the decoder. On the other hand, whenever the value of a "systematic bit" node is observed (assume the value is zero w.o.l.o.g.), it can be removed leaving a degree-2 parity-check. Of course, degree-2 parity-checks imply equality and allow the connected "punctured bit" nodes to be merged (effectively summing their

degrees). This operation is shown on the right in Figure 3. The symmetry between graph reduction on the information bits and the parity checks will become important later. Now, we consider the d.d. of a single "punctured bit" node in the reduced graph. This can be seen as working from left to right in the graph, and assuming the value of the previous "systematic bit" node was erased. The probability of the event where the values of k "systematic bit" nodes are observed and the value of the next "systematic bit" node is erased by the channel is given by $(1-p)^k p$. The graph reduction associated with this event causes the degrees of $k+1$ "punctured bit" nodes (from the d.d. $L(x)$) to be summed. Hence, the new d.d. of the "punctured bit" nodes after graph reduction is given by

$$\widetilde{L}(x) = \sum_{k=0}^{\infty} (1-p)^k p L(x)^{k+1} = \frac{pL(x)}{1-(1-p)L(x)}. \quad (10)$$

After the graph reduction, we are left with a standard LDPC code with new edge-perspective degree distributions given by

$$\widetilde{\lambda}(x) = \frac{\widetilde{L}'(x)}{\widetilde{L}'(1)} = \frac{p^2 \lambda(x)}{\big(1-(1-p)L(x)\big)^2} \quad (11)$$

$$\widetilde{\rho}(x) = \frac{\widetilde{R}'(x)}{\widetilde{R}'(1)} = \frac{(1-p)^2 \rho(x)}{\big(1-pR(x)\big)^2}. \quad (12)$$

After the aforementioned graph reduction, all the "systematic bit" nodes and "code bit" nodes are removed. Therefore the residual LDPC code effectively sees a BEC whose erasure probability is 1, and the DE fixed point equation is given by

$$\widetilde{\lambda}\big(1 - \widetilde{\rho}(1-x)\big) = x. \quad (13)$$

Based on (11) and (12), the last equation is equivalent to (8).

Remark 1 (The notation of tilted degree distributions): The tilted degree distributions $\widetilde{\lambda}$ and $\widetilde{\rho}$ which are given in (11) and (12), respectively, depend on the erasure probability p of the BEC. For simplicity of notation, we do not write this dependency explicitly in our notation. However, in Section III, when discussing symmetry properties and replacing p by $1-p$, the erasure probability is written explicitly in these tilted degree distributions.

D. The Stability Condition for ARA Ensembles

Like the NSIRA codes presented in [10], ARA codes have DE fixed points at both $x=0$ and $x=1$. One can see this by evaluating (8) at these points while assuming that each d.d. function f satisfies $f(0)=0$ and $f(1)=1$. To get decoding started, the d.d. is perturbed slightly by adding degree-1 parity-checks, pilot bits, and/or systematic bits. For successful completion of decoding, we need the fixed point at $x=0$ to be *stable*. To minimize the number of extra bits required to get decoding started, it is also useful for the fixed point (prior to the perturbation) at $x=1$ to be *unstable*. Although $x=1$ is not a fixed point after the perturbation, the *instability* condition helps prevent the decoder from getting stuck near $x=1$.

The stability and instability conditions are computed by taking the derivative of the LHS of (8) at $x=0$ and $x=1$. For the fixed point at $x=0$ to be stable, we need the derivative to be less than unity, and this gives

$$p^2 \lambda_2 \left(\rho'(1) + \frac{2pR'(1)}{1-p} \right) < 1. \quad (14)$$

Ensembles without degree-2 bits are unconditionally stable at $x=0$.

For the fixed point at $x=1$ to be unstable, we need the derivative to be greater than unity, and this gives

$$(1-p)^2 \rho_2 \left(\lambda'(1) + \frac{2(1-p)L'(1)}{p} \right) > 1. \quad (15)$$

This condition requires the presence of a non-vanishing fraction of degree-2 "parity-check 2" nodes; ensembles not having this property are unable to immediately create new degree-1 checks and may therefore get stuck shortly after starting. The instability condition guarantees that, on average, more new degree-1 checks are being created than lost when x is close to 1.

III. SYMMETRY OF CAPACITY-ACHIEVING ENSEMBLES

In this section, we discuss the symmetry between the bit and check degree distributions of c.a. ensembles for the BEC. First, we describe this relationship for LDPC codes, and then we extend it to ARA codes. The extension is based on analyzing the decoding of ARA codes in terms of graph reduction and the DE analysis of LDPC codes.

A. Symmetry of Capacity-Achieving LDPC Ensembles

The relationship between the bit d.d. and check d.d. of c.a. ensembles of LDPC codes can be expressed in a number of ways. Starting with the DE fixed point equation

$$p\lambda\big(1 - \rho(1-x)\big) = x \quad (16)$$

where p designates the erasure probability of the BEC, we see that picking either the d.d. λ or ρ determines the other d.d. exactly. In this section, we make this notion precise and use it to expose some of the symmetries of c.a. LDPC codes.

A few definitions are needed to discuss things properly. Following the notation in [4], let $\mathcal{P}$ be the set of d.d. functions (i.e., functions f with non-negative power series expansions around zero which satisfy $f(0)=0$ and $f(1)=1$); this set is defined as follows:

$$\mathcal{P} \triangleq \left\{ f : f(x) = \sum_{k=1}^{\infty} f_k x^k, \ x \in [0,1], \quad f_k \geq 0, \right.$$
$$\left. f(0) = 0, \ f(1) = 1 \right\}. \quad (17)$$

Let $\mathcal{T}$ be an operator which transforms invertible functions $f : [0,1] \to [0,1]$ according to the rule

$$\mathcal{T}f(x) \triangleq 1 - f^{-1}(1-x)$$

where f^{-1} is the inverse function of f. The function $\mathcal{T}f$ is well-defined on $[0,1]$ for any function f which is strictly monotonic on this interval, and therefore for any function in $\mathcal{P}$. We will say that two d.d. functions f and g are *matched* if $\mathcal{T}f = g$ (since $\mathcal{T}^2 f \equiv f$, the equality $\mathcal{T}f = g$ implies that

decoding graph can be constructed from the ARA decoding graph (see Fig. 2) by removing the bottom accumulate structure.

Since an NSIRA code has no accumulate structure attached to the "punctured bit" nodes, the graph reduction process affects only the d.d. of the "parity-check 2" nodes. Therefore, graph reduction acts as a mapping $\mathcal{G}_{\text{NSIRA}}$ from the NSIRA d.d. pair (λ, ρ) to the LDPC d.d. pair $(\lambda, \widetilde{\rho}_p)$. This yields that for ensembles of NSIRA codes, the set of d.d. pairs which satisfy the DE fixed point equation is given by

$$\mathcal{C}_{\text{NSIRA}}(p) \triangleq \Big\{ (\lambda, \rho) \in \mathcal{P} \times \mathcal{P} \,|\, \lambda\big(1 - \widetilde{\rho}_p(1 - x)\big) = x \Big\}.$$

An ALDPC code has no accumulate structure attached to the "parity-check 2" nodes, and therefore the graph reduction process only affects the d.d. of the "punctured bit" nodes. Hence, graph reduction acts as a mapping $\mathcal{G}_{\text{ALDPC}}$ from the ALDPC d.d. pair (λ, ρ) to the LDPC d.d. pair $(\widetilde{\lambda}_{1-p}, \rho)$. For ALDPC ensembles, the set of d.d. pairs which satisfy the DE fixed point equation is therefore given by

$$\mathcal{C}_{\text{ALDPC}}(p) \triangleq \Big\{ (\lambda, \rho) \in \mathcal{P} \times \mathcal{P} \,|\, \widetilde{\lambda}_{1-p}\big(1 - \rho(1 - x)\big) = x \Big\}.$$

The symmetry between c.a. ensembles of NSIRA and ALDPC codes over the BEC follows from the symmetry relationship in (18), the equivalence between a d.d. pair (λ, ρ) for NSIRA codes and the d.d. pair $(\lambda, \widetilde{\rho}_p)$ for LDPC codes, and the equivalence between a d.d. pair (λ, ρ) for ALDPC codes and the d.d. pair $(\widetilde{\lambda}_{1-p}, \rho)$ for LDPC codes.

The symmetry relationship between c.a. NSIRA and ALDPC ensembles over the BEC is therefore given in the following diagram:

$$
\begin{array}{ccc}
(\lambda, \rho) \in \mathcal{C}_{\text{NSIRA}}(p) & \xleftrightarrow{\text{symmetry}} & (\rho, \lambda) \in \mathcal{C}_{\text{ALDPC}}(1 - p) \\
\big\downarrow \mathcal{G}_{\text{NSIRA}} & & \big\downarrow \mathcal{G}_{\text{ALDPC}} \\
(\lambda, \widetilde{\rho}_p) \in \mathcal{C}_{\text{LDPC}} & \xleftrightarrow{\text{symmetry}} & (\widetilde{\rho}_p, \lambda) \in \mathcal{C}_{\text{LDPC}}
\end{array}
$$

As before, the inverse of each graph reduction mapping is represented by a dashed arrow because this inverse transformation is only valid if it is known ahead of time that the power series expansions of λ and ρ are non-negative.

D. Connections with Forney's Transform

In [20], Forney introduces a graph transformation which maps the factor graph of any group code to the factor graph of the dual group code. For factor graphs of binary linear codes which only have equality and parity constraints (i.e., no trellis constraints), this operation is equivalent to swapping equality and parity constraints (e.g., bit and check nodes). Forney's approach represents observations by half-edges, and these remain attached to the original node even though the nature of that node has changed. For example, Forney's transform maps an LDPC code with parity-check matrix H to a low-density generator-matrix (LDGM) code with generator matrix H and the half-edges attached to the bit nodes of the LDPC code are attached to the parity-check nodes of the LDGM code.

Using Forney's transform, we see that the swapping of λ and ρ described by our symmetry mappings actually transforms the original ensemble into the dual ensemble. Let the design rate of the original ensemble be R, then the design rate of the dual ensemble is $1 - R$. This means that if we want to have any chance of achieving capacity, we must also map the channel erasure probability p to $1 - p$. Therefore, our symmetry relationships show that ARA, NSIRA, and ALDPC ensembles which are c.a. on BEC under iterative decoding also have dual ensembles which are c.a. on the BEC under iterative decoding.[1] In light of the area theorem and its relationship to the dual code [22], this result is not entirely surprising. Still, we had not considered the possibility that c.a. ensembles might automatically define c.a. dual ensembles.

Finally, we note that the basic structure of ARA codes is preserved under Forney's transform. In particular, this means that we can construct self-dual ARA codes by choosing the matrix which defines the connections between the "punctured bit" nodes and the "parity-check 2" nodes (see Fig. 2) to be symmetric. This property may also be useful for constructing quantum error-correcting codes based on classical codes which contain their duals [21].

IV. Capacity-Achieving ARA Ensembles

This section serves as a preparatory step towards the construction of explicit c.a. ARA ensembles for the BEC, whose decoding complexities stay bounded as the gap to capacity vanishes. Later in Section V we will present explicit constructions of bit-regular and check-regular ARA ensembles which are based on a similar approach due to the symmetry properties provided in the previous section. Section VI introduces another approach for the construction of c.a. ensembles of ARA codes with bounded complexity over the BEC. The concepts used for these constructions are based on the symmetry properties in the previous section, and the material presented in this section.

A. A Starting Point

Using the tilted degree distributions after graph reduction given in Section II-C.2, we apply the DE equation for LDPC codes (13) to derive c.a. sequences. This property is proved in the following lemma.

Lemma 1: If equality (13) is satisfied for all $x \in [0, 1]$, then the design rate of the corresponding ensemble of ARA codes is equal to the capacity of the BEC.

Proof: From the condition in (13), it follows that

$$1 - \widetilde{\lambda}^{-1}(1 - x) = \widetilde{\rho}(x). \tag{21}$$

Since $\widetilde{\lambda}(0) = 0$, $\widetilde{\lambda}(1) = 1$ and $\widetilde{\lambda}$ is a monotonic differentiable function on the interval $[0, 1]$, then from (21), the substitution

---

$x = \lambda(u)$ and integration by parts give

$$
\begin{aligned}
\int_0^1 \widetilde{\rho}(x)\, dx &= \int_0^1 \left(1 - \widetilde{\lambda}^{-1}(1-x)\right) dx \\
&= 1 - \int_0^1 \widetilde{\lambda}^{-1}(x)\, dx \\
&= 1 - \int_0^1 u \widetilde{\lambda}'(u)\, du \\
&= 1 - \left[u\, \widetilde{\lambda}(u)\Big|_0^1 - \int_0^1 \widetilde{\lambda}(u)\, du \right] \\
&= \int_0^1 \widetilde{\lambda}(u)\, du.
\end{aligned} \tag{22}
$$

From (3), (11), and the equalities $L(0) = 0$ and $L(1) = 1$, we get

$$
\begin{aligned}
\int_0^1 \widetilde{\lambda}(x)\, dx &= \int_0^1 \frac{p^2 \lambda(x)}{\left(1 - (1-p)L(x)\right)^2}\, dx \\
&= \frac{1}{L'(1)} \int_0^1 \frac{p^2 L'(x)}{\left(1 - (1-p)L(x)\right)^2}\, dx \\
&= \frac{1}{L'(1)} \int_0^1 \frac{p^2\, du}{\left(1 - (1-p)u\right)^2} \\
&= \frac{p}{L'(1)}.
\end{aligned} \tag{23}
$$

Similarly, from (3), (12) and the equalities $R(0) = 0$ and $R(1) = 1$, we obtain

$$
\int_0^1 \widetilde{\rho}(x)\, dx = \frac{1-p}{R'(1)}. \tag{24}
$$

By combining (22)–(24), we obtain the equality

$$
\frac{L'(1)}{R'(1)} = \frac{p}{1-p} \tag{25}
$$

and hence, from (5), the design rate of the ensemble of ARA codes is equal to $1-p$ (i.e., the ensemble achieves the capacity of the BEC). ∎

Now, consider the DE fixed point equation (13) (or equivalently (8)). Using this equation, we see that the condition $\rho(0) = 0$ (i.e., no degree-1 "parity-check 2" nodes) is necessary and sufficient to guarantee that (13) is always satisfied at $x = 1$. Likewise, the condition $\lambda(0) = 0$ (i.e., no degree-1 "punctured-bit" nodes) is necessary and sufficient to guarantee that (13) is always satisfied at $x = 0$. From Lemma 1, we conclude that if there exists a d.d. pair (λ, ρ) with bounded average degree that satisfies (8), then there is a c.a. sequence of ARA ensembles with bounded complexity for the BEC. This conclusion is also based on the truncation discussed in the next section.

B. Truncating Degree Distributions

After finding a d.d. pair which satisfies the DE equation (8), a suitable truncation can be used to exhibit a sequence of ensembles that achieves capacity. Consider, for example, a sequence of d.d. pairs $\{(\lambda^{(M)}, \rho^{(M)})\}_{M \in \mathbb{N}}$ indexed by the maximum degree M. Since the effect of each truncation is negligible as M goes to infinity, Lemma 1 shows that the design rate approaches capacity in this case. If the truncations are chosen properly, then we can also show that each truncated d.d. pair in the sequence has no DE fixed points for $x \in (0, 1]$. This implies that, for any $\delta > 0$, there exists a block length $n_0(\delta, M)$ such that the probability of decoding failure is less than δ for all block lengths $n > n_0$.

Since our DE equations depend on both the edge and node degree distributions, the truncation must be chosen carefully to simultaneously bound both. For the check d.d., we want modified degree distributions $\hat{R}$ and $\hat{\rho}$ such that $\hat{R}(x) > R(x)$ and $\hat{\rho}(x) > \rho(x)$ for $x \in [0, 1)$. In particular, we replace large degree checks by degree-1 checks and this gives

$$
\begin{aligned}
\hat{\rho}(x) &= \left(\rho_1 + \sum_{i=M+1}^{\infty} \rho_i \right) + \sum_{i=2}^{M} \rho_i x^{i-1} \\
\hat{R}(x) &= \frac{\int_0^x \hat{\rho}(t)\, dt}{\int_0^1 \hat{\rho}(t)\, dt}.
\end{aligned}
$$

This truncation was introduced in [10] and proven to satisfy the desired conditions.

For the bit d.d., we want truncated degree distributions $\hat{L}$ and $\hat{\lambda}$ that satisfy $\hat{L}(x) < L(x)$ and $\hat{\lambda}(x) < \lambda(x)$ for $x \in (0, 1]$. In this case, we replace large degree bits by pilot bits (e.g., these bits are forced to zero and known at the receiver); this gives

$$
\hat{\lambda}(x) = \sum_{i=1}^{M} \lambda_i x^{i-1}, \qquad \hat{L}(x) = \sum_{i=1}^{M} L_i x^i.
$$

We note that this truncation satisfies the desired conditions (as long as $L_i > 0$ for some $i > M$) because it simply removes positive terms.

C. Encoding and Decoding Complexity

When transmission takes place over a BEC, the encoding/decoding complexity under iterative message-passing decoding is defined to be the average number of edges per information bit in the Tanner graph of the code (see Fig. 2). The motivation for measuring the complexity in this way is because the encoder and the iterative decoder can be both designed to use every edge in the graph exactly one time (due to the absolute reliability of information provided by the BEC).

From the Tanner graph of ARA codes in Fig. 2, it can be verified that the encoding complexity (χ_{E}) and the decoding complexity (χ_{D}) are both equal to

$$
\chi_{\mathrm{E}} = \chi_{\mathrm{D}} = 3 + L'(1) + \frac{2(1-R)}{R} \tag{26}
$$

where R is the design rate of the ensemble.

The complexity of NSIRA codes can also be computed from Fig. 2 by ignoring the accumulate structure for the systematic bits. This shows that

$$
\chi_{\mathrm{E}} = \chi_{\mathrm{D}} = L'(1) + \frac{2}{R}. \tag{27}
$$

Likewise, the complexity of ALDPC codes can be computed from Fig. 2 by ignoring the accumulate structure for the parity bits. This shows that

$$
\chi_{\mathrm{D}} = \frac{3 + L'(1)}{R}. \tag{28}
$$

In general, the encoding complexity of ALDPC codes does not grow linearly with the block length because the brute-force encoding of an LDPC code is quadratic in the block length. We can, however, apply fast encoding methods for LDPC codes (e.g., [18] and [19]) to ALDPC codes. These methods will result in essentially linear-time encoding algorithms for c.a. ALDPC codes.

D. The Effect of Puncturing

Puncturing is a well-known technique that allows one to design for one code rate and adaptively increase that rate to match channel conditions. Strictly speaking, we note that punctured ARA ensembles are no longer systematic because some information bits may not be transmitted as a result of the puncturing. This technique, however, can be used to extend the range of p for which certain d.d. pairs are c.a. with bounded complexity.

For example, consider any code construction which is provably c.a. with bounded complexity for $p > p_0$ (e.g., the check-regular ARA ensemble which will be introduced in Section V). This construction can be made to achieve capacity for $0 < p < 1$ by simply puncturing bits at random before transmission (i.e., all bits have the same puncturing rate). Let α be the fraction of bits transmitted, then the effective erasure rate of the channel is given by $p_{\text{eff}} = 1 - \alpha(1 - p)$. Picking $\alpha < 1 - p_0$ guarantees that $p_{\text{eff}} > p_0$ and that the ensemble achieves capacity. This operation does increase the complexity by a factor of $\frac{1}{\alpha}$ because the punctured bits must be retained as part of the decoding graph. We apply this method in our computer simulations to increase the code rate of a particularly good ensemble of rate $\frac{1}{2}$ codes (see Fig. 13).

Codes with two classes of bits (e.g., ARA codes) may also benefit from asymmetric puncturing of the two classes. For example, puncturing all of the systematic bits of an ARA code converts that code into a NSIRA code [10]. So we find that sending a fraction α of the systematic bits of an ARA code gives a smooth transition between ARA codes and NSIRA codes for $\alpha \in [0, 1]$.

V. Bit and Check Regular Capacity-Achieving Ensembles with Bounded Complexity

This section gives explicit constructions of c.a. ensembles for the BEC, which are either bit-regular or check-regular. As will be observed, these ensembles have bounded complexity (per information bit) as the gap to capacity vanishes.

The symmetry property in Section III-B allows one for example to design an ensemble of high-rate ARA codes, and get automatically (by switching between the pair of degree distributions) a new ensemble of ARA codes which is suited for low-rate applications. We will rely on this symmetry property in Section V-B when we transform a bit-regular ARA ensemble designed for a BEC with erasure probability $p \in (0, p^*]$ into a check-regular ensemble designed for $p \in [1 - p^*, 1)$. We also rely on the fact that the method in Section V-A for computing the function R given the function L can be easily inverted using the symmetry property. This means that given an algorithm to solve for $R(x)$ in terms of

$L(x)$ for a certain p_0, the inverse algorithm which solves $L(x)$ in terms of $R(x)$ is exactly the same, except that p_0 is replaced by $1 - p_0$.

A. Solving for $R(x)$ in terms of $L(x)$

Given $L(x)$, we start with the calculation of $\lambda(x) = \frac{L'(x)}{L'(1)}$. Then $\widetilde{\lambda}(x)$ is calculated from (11), and $\widetilde{\rho}(x) = 1 - \widetilde{\lambda}^{-1}(1 - x)$ is calculated from (13). Combining (3) and (12) gives

$$\widetilde{\rho}(x) = \frac{(1-p)^2}{R'(1)} \frac{R'(x)}{\left(1 - pR(x)\right)^2}$$

and by integrating both sides of this equation, we get

$$\int_0^x \widetilde{\rho}(t)\, dt = \frac{(1-p)^2}{R'(1)} \frac{R(x)}{1 - pR(x)}. \tag{29}$$

Since $R(1) = 1$, substituting $x = 1$ in the last equality and solving for $R'(1)$ gives

$$R'(1) = \frac{1 - p}{\int_0^1 \widetilde{\rho}(t)\, dt}.$$

By substituting the last equality in (29), we get

$$Q(x) = \frac{(1-p)R(x)}{1 - pR(x)} \tag{30}$$

where

$$Q(x) \triangleq \frac{\int_0^x \widetilde{\rho}(t)\, dt}{\int_0^1 \widetilde{\rho}(t)\, dt}. \tag{31}$$

Using the fact that $y = \frac{z}{1 - pz} \implies z = \frac{y}{1 + py}$, we solve (30) for $R(x)$ and get

$$R(x) = \frac{Q(x)}{1 - p + pQ(x)}. \tag{32}$$

Combining (3), (31) and (32) gives

$$\rho(x) = \frac{R'(x)}{R'(1)}$$

$$= \frac{\widetilde{\rho}(x)}{\left(1 - p + pQ(x)\right)^2}. \tag{33}$$

As long as we have $\widetilde{\rho}(1) = 1$, then evaluating (33) at $x = 1$ gives $\rho(1) = 1$. Therefore, there is no need to truncate the power series of ρ. As we noted above, a very similar approach can be applied to solve for $L(x)$ in terms of $R(x)$; due to the symmetry property, one can simply apply the above procedure to a parity-check d.d. $R(x)$ with an erasure probability of $1 - p$.

B. Bit and Check Regular Capacity-Achieving ARA Ensembles

The symmetry between bit-regular and check-regular c.a. ensembles of ARA codes follows from the symmetry properties presented in Section III-B. So, without loss of generality, we focus on a bit-regular ARA ensemble. Let $\lambda(x) = x^2$, so $L(x) = x^3$, and from (11)

$$\widetilde{\lambda}(x) = \frac{p^2 x^2}{\left(1 - (1-p)x^3\right)^2}.$$

Based on (13), we get

$$
\begin{aligned}
\widetilde{\rho}^{-1}(x) &= 1 - \widetilde{\lambda}(1-x) \\
&= 1 - \frac{p^2(1-x)^2}{\left(1-(1-p)(1-x)^3\right)^2}.
\end{aligned}
\tag{34}
$$

This is exactly [10, Eq. (39)] with p replaced by $1-p$ and ρ switched with λ. Therefore, we obtain from [10, Theorem 2] that the tilted d.d. $\widetilde{\rho}$ has the form

$$
\widetilde{\rho}(x) = 1 + \frac{2(1-x)^2 \sin\left(\frac{1}{3}\sin^{-1}\left(\sqrt{-\frac{27(1-p)(1-x)^{\frac{3}{2}}}{4p^3}}\right)\right)}{\frac{\sqrt{3}\,p^4}{1-p}\left(-\frac{(1-p)(1-x)^{\frac{3}{2}}}{p^3}\right)^{\frac{3}{2}}}.
\tag{35}
$$

Following the procedure of Section V-A, starting from (31), gives (after some calculus)

$$
Q(x) = \frac{3(x-1)\widetilde{\rho}(x)}{p} + \frac{1-\left(1-\widetilde{\rho}(x)\right)^3}{1-(1-p)\left(1-\widetilde{\rho}(x)\right)^3}
\tag{36}
$$

where the calculations leading to the expression for Q are detailed in Appendix A.1. Substituting Q in (36) into (32) gives the following expression of the d.d. R:

$$
\begin{aligned}
R(x) &= \frac{1}{p}\left(1 - \frac{1-p}{1-p+pQ(x)}\right) \\
&= \frac{Q(x)}{1-p+pQ(x)} \\
&= \frac{\frac{3(x-1)\widetilde{\rho}(x)}{p} + \frac{1-\left(1-\widetilde{\rho}(x)\right)^3}{1-(1-p)\left(1-\widetilde{\rho}(x)\right)^3}}{1-p+3(x-1)\widetilde{\rho}(x) + \frac{p-p\left(1-\widetilde{\rho}(x)\right)^3}{1-(1-p)\left(1-\widetilde{\rho}(x)\right)^3}}
\end{aligned}
\tag{37}
$$

where the function $\widetilde{\rho}$ is given in (35). It was verified numerically that for $p \leq 0.384$, the first 300 coefficients of the power series expansion of the d.d. R are non-negative; in Appendix C.3.2, we prove that if $p \leq 0.26$, then R has indeed a power series expansion about $x = 0$ whose all coefficients are non-negative. According to Lemma 1 (see p. 7), it also holds in general that any d.d. pair satisfying (8) has a design rate equal to the capacity of the BEC. It therefore appears that the d.d. pair above characterizes a c.a. ensemble of bit-regular ARA codes over the BEC; the capacity of the BEC is achieved with bounded complexity for rates greater than 0.616. We note that the convergence speed of the degree distribution for the parity-check nodes is relatively fast. As an example, for $p = 0.3$, the fraction of check nodes with degree less than 32 is equal to 0.968.

Using the symmetry between $\widetilde{\lambda}$ and $\widetilde{\rho}$ (see Section III), this also implies that for rates less than 0.384, the ensemble of check-regular ARA codes with $R(x) = x^3$ achieves capacity over the BEC with bounded complexity. Based on the symmetry property for c.a. ensembles of ARA codes, the d.d. L for the check-regular ARA ensemble is obtained from the d.d. R for the bit-regular ARA ensemble when p is replaced by $1-p$. From (37) and the symmetry property, the d.d. L

for the check-regular ARA ensemble which corresponds to $R(x) = x^3$ has the form

$$
L(x) = \frac{\frac{3(x-1)\widetilde{\lambda}(x)}{1-p} + \frac{1-\left(1-\widetilde{\lambda}(x)\right)^3}{1-p\left(1-\widetilde{\lambda}(x)\right)^3}}{p+3(x-1)\widetilde{\lambda}(x) + \frac{(1-p)-(1-p)\left(1-\widetilde{\lambda}(x)\right)^3}{1-p\left(1-\widetilde{\lambda}(x)\right)^3}}
\tag{38}
$$

where for this ensemble

$$
\widetilde{\lambda}(x) = 1 + \frac{2(1-x)^2 \sin\left(\frac{1}{3}\sin^{-1}\left(\sqrt{-\frac{27p(1-x)^{\frac{3}{2}}}{4(1-p)^3}}\right)\right)}{\frac{\sqrt{3}(1-p)^4}{p}\left(-\frac{p(1-x)^{\frac{3}{2}}}{(1-p)^3}\right)^{\frac{3}{2}}}.
\tag{39}
$$

Note that the d.d. $\widetilde{\lambda}$ in (39) is obtained by replacing p by $1-p$ in the RHS of (35), which finally gives the d.d. function introduced in [10, Eq. (15)].

C. Capacity-Achieving ALDPC Ensembles

Using the symmetry relationship between NSIRA and ALDPC ensembles from Section III-C, we find that we already have from [10, Theorems 1 and 2] two c.a. ensembles of ALDPC codes. These ensembles are based on the bit-regular and check-regular NSIRA ensembles of [10]. This was also observed independently by Hsu and Anastasopoulos [14].

Using symmetry, the check-regular NSIRA ensemble gives a bit-regular ALDPC ensemble which provably achieves capacity with bounded complexity for $p \in (0, 1)$. Since d.d. for small p has long tails, one can also use random puncturing to increase the effective erasure rate of the channel, and therefore simplify code design. Similarly, the bit-regular NSIRA ensemble gives a check-regular ALDPC ensemble which provably achieves capacity with bounded complexity for $p \in \left[\frac{12}{13}, 1\right)$. In this case, random puncturing can be used to extend the valid range to $(0, 1)$.

The replacement of p by $1-p$ in [10, Eq. (15)] which corresponds to the right d.d. of the check-regular NSIRA ensemble, gives the following check d.d. for the bit-regular ALDPC ensemble with left degree of 3

$$
\rho(x) = 1 + \frac{2(1-x)^2 \sin\left(\frac{1}{3}\sin^{-1}\left(\sqrt{-\frac{27(1-p)(1-x)^{\frac{3}{2}}}{4p^3}}\right)\right)}{\frac{\sqrt{3}\,p^4}{1-p}\left(-\frac{(1-p)(1-x)^{\frac{3}{2}}}{p^3}\right)^{\frac{3}{2}}}.
\tag{40}
$$

Likewise, the replacement of p by $1-p$ in [10, Eq. (10)] which corresponds to the left d.d. of the bit-regular NSIRA ensemble, and the substitution $q = 3$ gives the following bit d.d. for the check-regular ALDPC ensemble with right degree of 3

$$
\lambda(x) = \frac{1-(1-x)^{\frac{1}{2}}}{\left[1-(1-p)\left(1-3x+2\left[1-(1-x)^{\frac{3}{2}}\right]\right)\right]^2}.
\tag{41}
$$

It is worth noting that the bit-regular ALDPC ensemble has minimum bit degree of 3. Therefore, truncating the check d.d.

to finite maximum check degree makes the ensemble unconditionally stable. Recall that the typical minimum distance of regular LDPC ensembles, with bit degree 3, grows linearly with the block length [23]. This suggests that the minimum distance of this bit-regular ALDPC ensemble might also grow linearly with the block length. To prove this rigorously, however, one must also consider the effect of the accumulate structure on the minimum distance.

VI. CAPACITY-ACHIEVING ENSEMBLES WITH BOUNDED COMPLEXITY FROM SELF-MATCHED LDPC ENSEMBLES

In this section, we introduce another way of constructing c.a. ensembles of ARA codes for the BEC. Rather then solving for the function R in terms of the function L (as in Section V-A) or doing the inverse via the symmetry property, we consider another natural way of searching for c.a. degree distributions. We start by choosing a candidate d.d. pair $(\widetilde{\lambda}, \widetilde{\rho})$ which satisfies equation (13) and test if it can be used to construct an ensemble of c.a. ARA codes. The testing process starts by mapping the tilted pair $(\widetilde{\lambda}, \widetilde{\rho})$ back to (λ, ρ) via (11) and (12), and then testing the non-negativity of the resulting power series expansions of λ and ρ.

Following the notation in Section III-A, it enables one to rewrite (13) as $\widetilde{\rho} = \mathcal{T}\widetilde{\lambda}$ (so the tilted degree distributions $\widetilde{\lambda}$ and $\widetilde{\rho}$ are matched), and gives a compact description of capacity-achieving d.d. pairs of LDPC codes. Since $\mathcal{T}^2 f = f$ for every invertible function f, we note that $f \in \mathcal{A}$ if and only if $\mathcal{T}f \in \mathcal{A}$. Based on (13), we obtain that we need to choose the tilted d.d. so that $\widetilde{\lambda} \in \mathcal{P}$ and also $\mathcal{T}\widetilde{\lambda} \in \mathcal{P}$, i.e., we need that the d.d. $\widetilde{\lambda}$ (or $\widetilde{\rho}$) both belong to the set $\mathcal{A}$. The reader is referred to [4, Lemma 1] which considers basic properties of the set $\mathcal{A}$ and the transformation $\mathcal{T}$.

So far, by choosing $\widetilde{\lambda} \in \mathcal{A}$ (or $\widetilde{\rho} \in \mathcal{A}$), we only know that both tilted d.d. have non-negative power series expansions. This property does not ensure that both of the original (i.e., non-tilted) d.d. λ and ρ also have non-negative power series expansions. Calculation of λ and ρ from the tilted d.d. $\widetilde{\lambda}$ and $\widetilde{\rho}$ is not straightforward since both equations involve the d.d. L and R which are the normalized integrals of the unknown λ and ρ. In order to overcome this difficulty in solving the two integral equations, we suggest calculating the tilted d.d. pair w.r.t. the nodes of the graph using

$$\widetilde{L}(x) = \frac{\int_0^x \widetilde{\lambda}(t)\, dt}{\int_0^1 \widetilde{\lambda}(t)\, dt} \;, \quad \widetilde{R}(x) = \frac{\int_0^x \widetilde{\rho}(t)\, dt}{\int_0^1 \widetilde{\rho}(t)\, dt}. \tag{42}$$

The original d.d. pair w.r.t. the nodes (i.e., the original d.d. pair before the graph reduction) can be calculated from (9) and (10). We obtain that

$$L(x) = \frac{\widetilde{L}(x)}{p + (1-p)\widetilde{L}(x)} \;, \quad R(x) = \frac{\widetilde{R}(x)}{1 - p + p\widetilde{R}(x)} \tag{43}$$

and then use the equations in (3) to find (λ, ρ). The critical issue here is to verify whether the functions L and R have non-negative power series expansions.

A. Capacity-Achieving ARA Ensembles

It is easy to verify that the function

$$f(x) = \frac{(1-b)x}{1-bx} \;, \quad 0 < b < 1 \tag{44}$$

belongs to the set $\mathcal{A}$ and also $\mathcal{T}f = f$; in the case where $\mathcal{T}f = f$, the function f is said to be self-matched. Therefore, based on (13), we examine here whether the choice $\widetilde{\lambda}(x) = \widetilde{\rho}(x) = \frac{(1-b)x}{1-bx}$ can be transformed into an ensemble of ARA codes whose degree distributions have non-negative power series expansions. From (42) and (43), we get

$$\widetilde{L}(x) = \widetilde{R}(x) = \frac{bx + \ln(1-bx)}{b + \ln(1-b)} \tag{45}$$

and from (43), we obtain that

$$L(x) = \frac{bx + \ln(1-bx)}{p\left[b + \ln(1-b)\right] + (1-p)\left[bx + \ln(1-bx)\right]} \tag{46}$$

$$R(x) = \frac{bx + \ln(1-bx)}{(1-p)\left[b + \ln(1-b)\right] + p\left[bx + \ln(1-bx)\right]}. \tag{47}$$

Since we started with the function f in (44) which is self-matched, the resulting functions L and R in this approach are exactly the same, except that p and $1-p$ are switched. In Appendix C.3, it is proved that the degree distributions L and R in (46) and (47), respectively, have non-negative power series expansion if and only if

$$\left| p - \frac{1}{2} \right| \le \frac{1}{2} - \frac{1}{1 - \frac{13-\sqrt{61}}{9} \cdot \left(b + \ln(1-b)\right)}. \tag{48}$$

Fortunately, there exists a region of (b, p) where this condition is satisfied. For the specification of this region, we use the Lambert W-function $W(x)$ which is defined implicitly via $W(x)e^{W(x)} = x$; this function is real for $x > -\frac{1}{e}$. In the following we introduce and prove the following theorem:

Theorem 1 (Ensembles of Self-Matched ARA Codes): The ensemble of self-matched ARA codes, defined by the pair of degree distributions (L, R) in (46) and (47), achieves the capacity of the BEC for any erasure probability $p \in (0, 1)$. This result is achieved under iterative message-passing decoding with *bounded complexity*.

The tails of the d.d. (i.e., the partial sums $\sum_{i=k}^{\infty} L_i$ and $\sum_{i=k}^{\infty} R_i$) *decay exponentially* like $O(b^k)$ where the parameter b is given by

$$b = W\left(-e^{-\frac{13+\sqrt{61}}{12} \frac{1+|1-2p|}{1-|1-2p|} - 1} \right) + 1. \tag{49}$$

The complexity, per information bit, of encoding and decoding is given by

$$\chi_{\mathrm{E}} = \chi_{\mathrm{D}} = \frac{3-p}{1-p} - \frac{b^2 p}{(1-b)\left[b + \ln(1-b)\right]} . \tag{50}$$

Proof: Referring to the pair of degree distributions L and R in (46) and (47), respectively, we need to obtain the necessary and sufficient conditions which ensure that these two function have non-negative power series expansion about $x = 0$. For a given value of b in these degree distributions, it is proved in Appendix C.3.3 that this property is satisfied if and only if the inequality in (48) holds.

The encoding and decoding complexities of c.a. ensembles of ARA codes for the BEC are discussed in Section IV-C. Since our ensemble is c.a., then $R = 1 - p$ where p designates the erasure probability of the BEC, and from (43) and (45)

$$L'(1) = p\,\widetilde{L}'(1) = -\frac{b^2 p}{(1-b)[b + \ln(1-b)]}\ .$$

Combining (26) with the last equality provides the expression in (50) for the complexity, per information bit, of encoding and decoding.

For fixed $p \in (0,1)$, the complexity in (50) forms a monotonic increasing function of b (which becomes unbounded as $b \to 1^-$). In order to minimize the encoding/decoding complexity, we wish to find the smallest value of b in the interval $(0,1)$ so that the power series expansions about zero of the degree distributions L and R are both non-negative. For a fixed value of p, it is equivalent to solving for the minimal value of $b \in (0,1)$ which satisfies the condition in (48). This gives the equation

$$\frac{1}{1 - \frac{13-\sqrt{61}}{9}\cdot\left(b + \ln(1-b)\right)} = \min(p, 1-p)$$

which can be rewritten as

$$-b - \ln(1-b) = \frac{13 + \sqrt{61}}{12}\,\frac{1 + |1 - 2p|}{1 - |1 - 2p|} \qquad (51)$$

by using the equality

$$\min(p, 1-p) = \frac{1}{2} - \left|\frac{1}{2} - p\right|.$$

For $a \in \mathbb{R}$, the solution of the equation $-b - \ln(1-b) = a$ is given by $b = W(-e^{-1-a}) + 1$. To verify this, one needs to write the equation in the form $(b-1)e^{b-1} = -e^{-a-1}$, and rely on the definition of the Lambert W-function. Hence, the solution of equation (51) is given by the expression for b in (49). For $p \in (0,1)$, the expression for b in (49) achieves its global minimum at $p = \frac{1}{2}$, and its value is

$$b^* = W(-e^{-\frac{25+\sqrt{61}}{12}}) + 1 \approx 0.9304.$$

Eq. (49) therefore implies that for $0 < p < 1$, the parameter b ranges in the interval $[b^*, 1)$; it achieves the value $b = b^*$ at $p = \frac{1}{2}$, and tends to 1 when p approaches zero or unity.

The asymptotic behavior of the two d.d. pairs w.r.t. the nodes and the edges is derived in Appendix B.3, and is given by

$$L_k, R_k = O\left(\frac{b^k}{k\ln^2(k)}\right), \qquad \lambda_k, \rho_k = O\left(\frac{b^k}{\ln^2(k)}\right) \quad (52)$$

so the tails of the d.d. pair (L, R) decay exponentially with k. ∎

The region (b, p) characterized by the inequality in (48) is depicted in Fig. 4, and we point out that its width grows as b gets closer to 1. Note that it follows from (48) that in the limit where $b \to 1^-$, the d.d. pair L and R have non-negative power series expansions for $0 < p < 1$. However, from (50), the complexity in the limit where $b \to 1^-$ becomes unbounded.
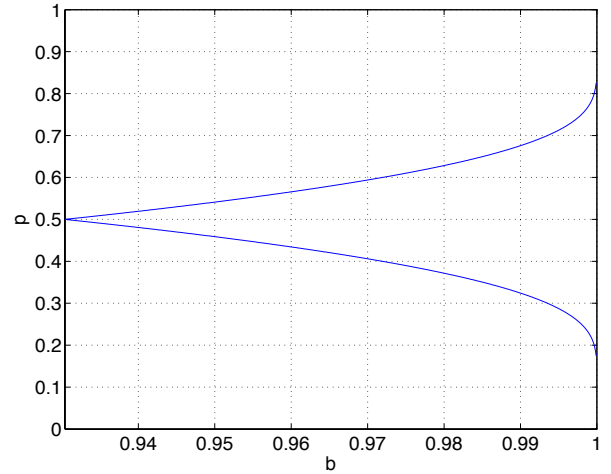
Fig. 4. Valid values of p as a function of the parameter b in (46) and (47), so that the coefficients in the power series expansions around zero of $L(\cdot)$, $R(\cdot)$, $\lambda(\cdot)$ and $\rho(\cdot)$ are all non-negative. This region is determined by the inequalities in (48), and it is therefore bounded by the upper and lower curves in this figure.

An efficient algorithm for the calculation of the d.d. pair in (46) and (47) w.r.t. the nodes of the graph, and the d.d. pair (λ, ρ) w.r.t. the edges is given in Appendix B.2.

We believe the performance advantage of this ensemble over other c.a. ensembles is mainly due to the *exponential* decay of the d.d. coefficients, as given in (52). In contrast, most other ensembles in Table VII-C have d.d. tails which decay polynomially.

1) Numerical Results: The encoding and decoding complexity of the self-matched ARA ensemble introduced in this section is shown in Fig. 5. The minimal value of k such that the partial sums, $\sum_{i=2}^{k}\lambda_i$ and $\sum_{i=2}^{k}\rho_i$, exceed 0.95 is shown in Fig. 5.
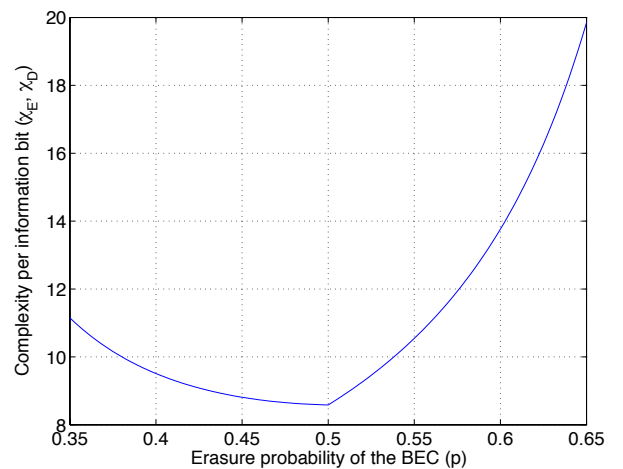
Fig. 5. Encoding/decoding complexity per information bit of the considered ensemble of ARA codes. The complexity is given in (50) as a function of the erasure probability (p) of the BEC, and to this end, we choose the parameter b according to equation (49) for minimizing the complexity.

Both of these sums converge to 1 as k goes to infinity, and the convergence time is measured by the minimal value of k

Fig. 6. A plot of the minimal value of k as a function of the erasure probability (p) so that the partial sums $\sum_{i=2}^{k} \lambda_i$ and $\sum_{i=2}^{k} \rho_i$ both exceed 0.95. We choose the parameter b according to equation (49) which also minimizes k for a fixed value of p.

Fig. 8. The degree distributions, (46) and (47), of the ensemble of self-matched ARA codes for $p = 0.6$ and $b = 0.972$. Also shown are the partial sums $\sum_{i=2}^{k} L_i$, $\sum_{i=2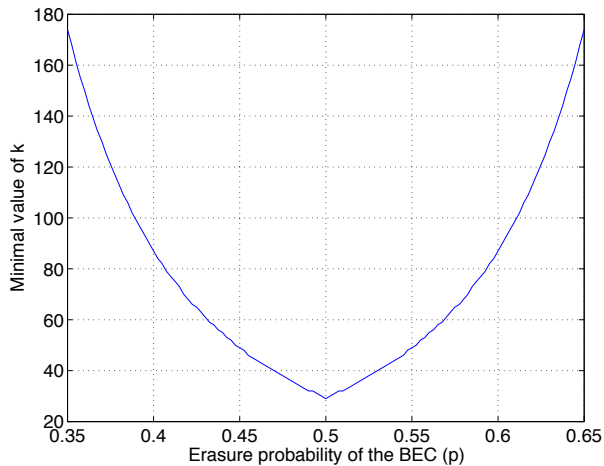}^{k} R_i$, $\sum_{i=2}^{k} \lambda_i$, and $\sum_{i=2}^{k} \rho_i$ as a function of the integer k. The (encoding and decoding) complexity per information bit is 13.78.

where these partial sums exceed a threshold which is close to 1 (e.g., 0.95). We note that, like the complexity, the convergence time of these partial sums is increasing with b. Therefore, the choice we made according to equation (49) minimizes both quantities simultaneously. Notice that the complexity and convergence time are rather small for $0.35 \leq p \leq 0.65$. Both of these quantities achieve their minimal value at $p = \frac{1}{2}$ since the value of b required by (49) is also minimized. We also note that the minimal value of k is symmetric around $p = \frac{1}{2}$ (see right plot) while the complexity is asymmetric (see left plot). The reason for the symmetry property around $p = \frac{1}{2}$ in Fig. 6 is because the replacement of p by $1-p$ yields the same value of b in (49); actually, the replacement of p by $1-p$ yields the same d.d. pair, except that $\{\lambda_i\}_{i \geq 2}$ and $\{\rho_i\}_{i \geq 2}$ are switched (this follows directly from (46) and (47)).
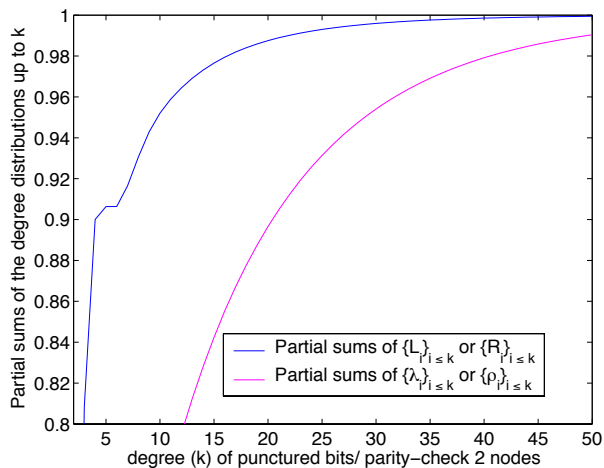

Fig. 7. The degree distributions, (46) and (47), of the ensemble of self-matched ARA codes for $p = 0.5$ and $b = 0.9304$. Also shown are the partial sums $\sum_{i=2}^{k} L_i$, $\sum_{i=2}^{k} R_i$, $\sum_{i=2}^{k} \lambda_i$, and $\sum_{i=2}^{k} \rho_i$ as a function of the integer k. The complexity per information bit is 8.59.

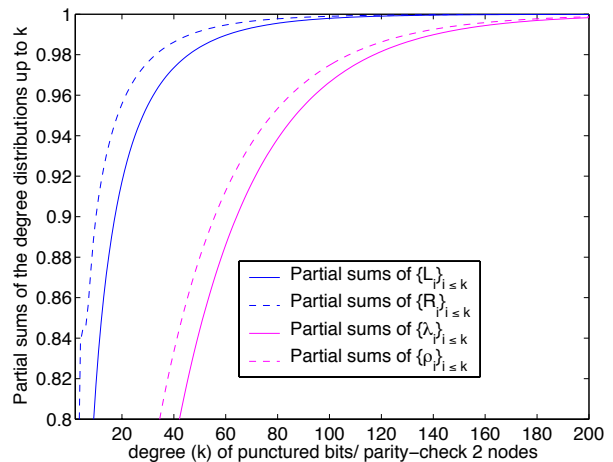For $p = 0.5$ and $p = 0.6$, the convergence rates of the

degree distributions w.r.t. the nodes and edges of the graph are shown in Figs. 7 and 7. The value of the parameter b is determined by (49), and the corresponding complexity (per information bit) is equal to 8.59 and 13.78, respectively. The results for a BEC with erasure probability $p = 0.5$ (i.e., a design rate of $\frac{1}{2}$) are particularly encouraging. In this case, the encoding/decoding complexity is equal to 6.59, and the partial sum $\sum_{i=2}^{k} \lambda_i$ (or equivalently, $\sum_{i=2}^{k} \rho_i$) exceeds 0.95 for $k \geq 29$. For comparison, consider the check-regular NSIRA ensemble in [10, Theorem 2] which requires more than 300 terms so that the partial sum $\sum_{i=2}^{k} \lambda_i$ exceeds 0.95. This significant improvement in the convergence rate of the degree distributions yields ensembles whose performance for moderate block lengths is superior to previous constructions. The considered ensemble of self-matched ARA codes has the property that for a design rate of one-half, $L = R$ and $\lambda = \rho$, so the d.d. pairs of the punctured bits and the parity-checks coincide.

In Fig. 9, we compare the asymptotic expressions of the degree distributions $\{L_k\}$, $\{R_k\}$, $\{\lambda_k\}$, $\{\rho_k\}$ to their exact values. There is a good match between the asymptotic and exact values for moderate to large values of k. The best match between the two expressions is obtained when $p = 0.5$ because this affords the minimal value of b. To see this phenomenon exactly, one can look at the error terms of the asymptotic expressions given in Appendix B.3.

To conclude, we note that the ensemble of self-matched ARA codes without puncturing, as considered in this section is well suited for moderate rates while, on the other hand, the ensembles of bit-regular and check-regular ARA codes are well suited for high and low rates, respectively. In order to make the ensemble of self-matched ARA codes suitable for high code rates, we use random puncturing (as will be exemplified later, the performance of these ensembles with puncturing is good also for moderate block lengths).
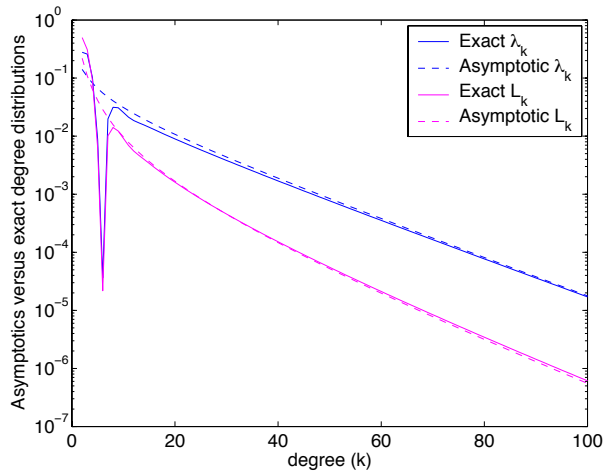
Fig. 9. A plot comparing the asymptotic versus exact expressions of the degree distributions $\{L_k\}$, $\{R_k\}$, $\{\lambda_k\}$ and $\{\rho_k\}$ of the self-matched ARA ensemble, designed for a BEC with erasure probability of $p = \frac{1}{2}$. The parameter b in the left and right degree distributions (46) and (47), respectively, is determined by Eq. (49), so its value is $b = 0.9304$.

2) Not All Self-Matched LDPC Ensembles Give Valid Results: The starting point in Section VI-A was the choice of the function f in (44) which belongs to the set $\mathcal{P}$ and which also satisfies the property $f = \mathcal{T}f$. This choice simplifies the analysis in Section VI-A by setting $\widetilde{\lambda}(x) = \widetilde{\rho}(x) = f(x)$. From the symmetry property stated in Section III, we see that L and R (and also λ and ρ) have the same form except that p is replaced by $1 - p$. The function f in (44) is however not the only function in the set $\mathcal{P}$ which satisfies the property $f = \mathcal{T}f$. In [4, Appendix V], there is a discussion on the fixed points of the operator $\mathcal{T}$. We cite here a necessary condition for the satisfaction of this property.

Proposition ([4, Appendix V]): If $f = \mathcal{T}f$ for some $f \in \mathcal{P}$, then $f(x) = t(x) + 1 - x$ where $t(x)$ satisfies the identity

$$t(x) = 2x - 1 - h(t(x)), \quad x \in [0, 1] \tag{53}$$

and the function h satisfies the properties

$$0 \le h(t) < 1, \quad h(t) = h(-t), \, t \in [0, 1], \quad h(1) = h(-1) = 0.$$

Choosing appropriate functions in (53) may lead to new families of fixed points of the operator $\mathcal{T}$. For example, the choice $h(t) = \alpha(1 - t^2)$ gives the function

$$f(x) = 1 + \frac{1}{2\alpha} - x - \sqrt{\left(1 + \frac{1}{2\alpha}\right)^2 - \frac{2x}{\alpha}} \tag{54}$$

which indeed belongs to the set $\mathcal{P}$ and satisfies the property $f = \mathcal{T}f$ for $0 < \alpha \le \frac{1}{2}$. Using an approach similar to the one in Section VI-A, we set $\widetilde{\lambda} = \widetilde{\rho} = f$, and rely on (42) to calculate $\widetilde{L}$ and $\widetilde{R}$. These functions are identical because f is a fixed point of the operator $\mathcal{T}$. The choice of the function in (54) gives $\widetilde{L}(x) = \widetilde{R}(x)$ equal to

$$\frac{\left(1 + \frac{1}{2\alpha}\right)x - \frac{x^2}{2} + \frac{\alpha}{3}\left\{\left[\left(1 + \frac{1}{2\alpha}\right)^2 - \frac{2x}{\alpha}\right]^{\frac{3}{2}} - \left(1 + \frac{1}{2\alpha}\right)^3\right\}}{\frac{1}{2} - \frac{2\alpha}{3}} \tag{55}$$

and the candidate d.d. pair (L, R) is then calculated by using (43). From our numerical experiments, L is likely to have a non-negative power series expansion for $p \in [0.85, 1)$ (with $\alpha \in [0.38, 0.5]$). As we noted above, the expressions for L and R are the same except that p is replaced by $1 - p$. Since the intervals $[0.85, 1)$ and $(0, 0.15]$ do not overlap, we find that both $L(x)$ and $R(x)$ cannot simultaneously have non-negative power series expansions about $x = 0$. This shows that not all self-matched functions give rise to valid degree distributions of ARA ensembles.

B. Capacity-Achieving NSIRA Ensembles

This section is focused on the construction of NSIRA ensembles from LDPC ensembles whose degree distributions from the edge perspective are matched. We apply here the concept of DE via graph reduction to ensembles of NSIRA codes. In this case, the graph reduction only applies to the "parity-check 2" nodes (see Fig. 2). This is because the upper part of Fig. 2 does not exist in the Tanner graph of NSIRA codes (i.e., the "punctured bit" nodes in this figure are the "information bit" nodes in the graph of NSIRA codes). Based on graph reduction, we obtain that $L = \widetilde{L}$ for ensembles of NSIRA codes, while the functions R and $\widetilde{R}$ satisfy the equality in (43). In a similar manner, the equality $\lambda = \widetilde{\lambda}$ holds for NSIRA ensembles while equality (12) is satisfied for the degree distributions of the parity-checks from the edge perspective. We note that from (12) and (13), the fixed point of the DE equations for NSIRA ensembles is given by

$$\lambda\left(1 - \frac{(1-p)^2 \rho(1-x)}{\left(1 - pR(1-x)\right)^2}\right) = x.$$

Of course, this equation coincides with the DE fixed point equation [10, Eq. (6)] (with x_0 replaced by x) derived previously for NSIRA codes.

For the construction of ensembles of NSIRA codes using LDPC codes whose degree distributions from the edge perspective are both matched to themselves, we rely as a starting point on the function f in (44) which forms a d.d. which is matched to itself, and set $\widetilde{\lambda}(x) = \widetilde{\rho}(x) = \frac{(1-b)x}{1-bx}$ for $0 < b < 1$, similarly to Section VI-A. For the considered ensemble of NSIRA codes, the d.d. $L(x)$ is then equal to $\widetilde{L}(x)$ in (45), i.e.,

$$L(x) = \frac{bx + \ln(1 - bx)}{b + \ln(1 - b)}. \tag{56}$$

From this, we see that there are no degree-1 "information bit" nodes, and that the fraction of "information bit" nodes with degree-i is given by

$$L_i = -\frac{b^i}{i} \frac{1}{b + \ln(1 - b)}, \quad i = 2, 3, \ldots.$$

The non-negativity of the sequence $\{L_i\}$ holds when $0 < b < 1$ (so $b + \ln(1 - b) < 0$). Therefore, the power series expansion of the d.d. L is always non-negative and there is no requirement on the erasure probability p in this regard. The condition for the d.d. R to be non-negative is identical to that

of the self-matched ARA codes derived in Appendix C.3.3 and is given by

$$p \leq 1 - \frac{1}{1 - \frac{13 - \sqrt{61}}{9} [b + \ln(1 - b)]} \ . \tag{57}$$

By comparing it to the parallel requirement for the ARA ensemble, as given in (48), one observes that (57) requires a weaker condition on p which is only the upper bound on p in (48). As mentioned above, the d.d. R is the same as for the ARA ensemble in Section VI-A. The encoding and decoding complexities of this ensemble are equal and have the form

$$\chi_{\mathrm{E}} = \chi_{\mathrm{D}} = \frac{2}{1 - p} - \frac{b^2}{(1 - b)[b + \ln(1 - b)]} \ . \tag{58}$$

This gives an explicit construction of NSIRA ensembles from LDPC codes whose degree distributions from the edge perspective are matched to themselves. In general, we find by computer simulations for finite-length codes over the BEC that ARA codes have the best performance.

C. Capacity-Achieving ALDPC Ensembles

We now construct ensembles of ALDPC codes using LDPC codes whose degree distributions from the edge perspective are matched. Using the symmetry property between NSIRA and ALDPC codes, this construction follows almost trivially from the results of Section VI-B. The symmetry transformation acts by switching λ with ρ and p with $1 - p$, which gives

$$R(x) = \frac{bx + \ln(1 - bx)}{b + \ln(1 - b)} \ . \tag{59}$$

From this, we see that there are no degree-1 "parity-check 2" nodes and that the fraction of "parity-check 2" nodes with degree-i is given by

$$R_i = -\frac{b^i}{i} \frac{1}{b + \ln(1 - b)} \ , \quad i = 2, 3, \dots \ .$$

The non-negativity of the sequence $\{R_i\}$ follows directly because $0 < b < 1$ (so $b + \ln(1 - b) < 0$). Since the R_i are non-negative for $p \in [0, 1]$, the valid range of p for this construction is determined by the non-negativity of $L(x)$. Similar to the NSIRA ensemble in Section VI-B, the d.d. L is equal to the one in (46) for the ARA ensemble. From Appendix C.3.3, we see that $L(x)$ will have a non-negative power series expansion if

$$p \geq \frac{1}{1 - \frac{13 - \sqrt{61}}{9} [b + \ln(1 - b)]} \ . \tag{60}$$

Using (28) and (46), the decoding complexity of this ensemble is given by

$$\chi_{\mathrm{D}} = \frac{3}{1 - p} - \frac{b^2 p}{(1 - p)(1 - b)[b + \ln(1 - b)]} \ . \tag{61}$$

This gives an explicit construction of ALDPC ensembles from ensembles of LDPC codes with self-matched degree distributions. In Section VII, we compare the performance of this ensemble with the ensembles of NSIRA and ARA codes in Sections VI-A and VI-B, respectively. In general, we find that the ARA codes have the best performance.

VII. COMPUTER SIMULATIONS

The details of our computer simulations are described in this section and the results are discussed. In particular, our main result is that the ARA ensemble of Section VI shows a distinct advantage over all the other c.a. ensembles we consider. We believe this advantage is largely due to the exponential decay of its d.d. tails.

A. Encoding Details

We consider here the explicit construction of finite-length ARA codes from infinite degree distributions. Let k be the number of "systematic bit" nodes and $n - k$ be the number of "parity-check 2" nodes (see Fig. 2 in p. 3). In the first step, we scale and quantize the degree distributions into a list of integers corresponding to the number of "punctured bit" nodes and "parity-check 2" nodes of each degree. All "parity-check 2" nodes with degree greater than d_{R} are reduced to degree-d_{R}. This truncation has the effect of reducing the number of edges between "parity-check 2" nodes and "punctured bit" nodes. All "punctured bit" nodes with degree greater than d_{L} are converted into pilot bits. Therefore, the values of these bits are forced to a known value (e.g., zero) at the encoder. This is done by choosing the value of one of the "systematic bit" nodes carefully as described in Section II-B. We note that each pilot bit has the effect of reducing the dimension of the code by 1.

The main reason for introducing pilot bits is that decoding cannot get started without them (see the DE equations in Section II-C.1). A slight modification of the encoding process is required for pilot bits, and we describe it now using the notation from Section II-B. Our goal is to set the punctured bit v_j to zero by picking a systematic bit appropriately. We find that the punctured bit v_j can be set to zero by choosing the systematic bit u_j to be $v_{j-1} = \sum_{i=1}^{j-1} u_i$. Since each pilot bit determines one of the systematic bits, this process reduces the actual number of information bits by the number of pilot bits.

While long random codes chosen from the ensemble of Section VI-A tend to have vanishing bit erasure probability as the block length goes to infinity, the block erasure probability does not vanish. This phenomenon is caused by small weaknesses in the graph, and can be mitigated by using a high-rate outer code as a second layer of protection. In the case of ARA codes, we believe that it is most effective to apply this code to the punctured bits. In particular, we force the last m punctured bits to equal random linear combinations of the first $k - m$ punctured bits. This operation reduces the information content of the code by m bits.

Again, we must slightly modify the encoding process to accommodate the outer code. Using the notation of Section II-B, we find that this involves forcing the sequence $\underline{v}$ to satisfy $H\underline{v} = \underline{0}$ for some $m \times k$ parity-check matrix H. For our purposes, it suffices to consider matrices of the form $H = [P \ I]$, where I is the $m \times m$ identity and P is an $m \times (k - m)$ matrix with entries $P_{j,i}$. We can require the $\underline{v}$ sequence to satisfy $H\underline{v} = \underline{0}$ by choosing each u_j, for $j \in \{k - m +$

Fig. 10. Simulations for the ensembles of ARA and NSIRA codes constructed from LDPC codes with self-matched degree distributions (see Sections VI-A and VI-B), and the ensemble of right-regular LDPC codes in [3]. The plots refer to block lengths of 8192 and 65536 bits (see upper and lower plots, respectively) and a design rate of 0.5 bits per channel use. No high-rate outer code is assumed.
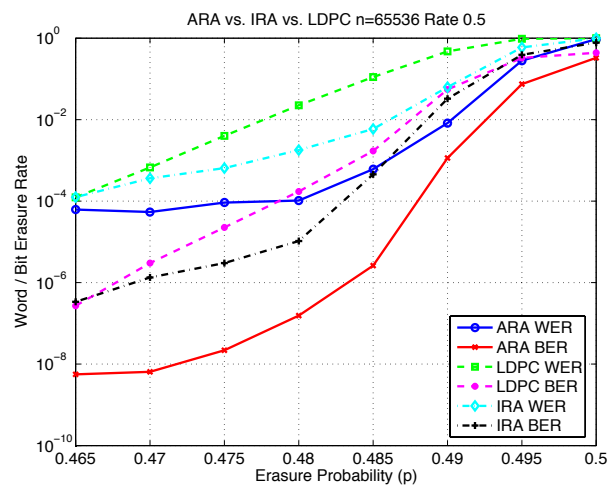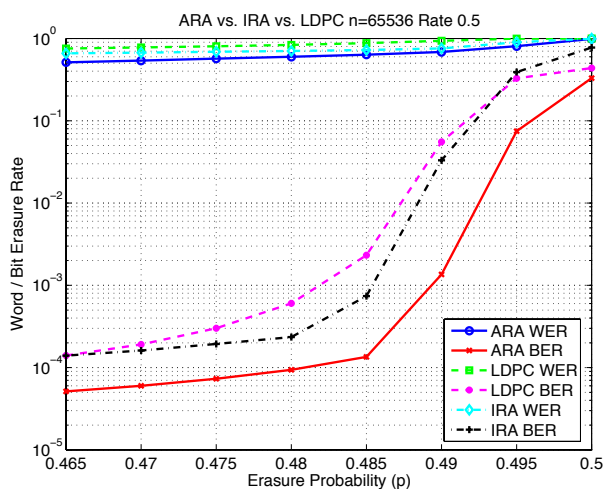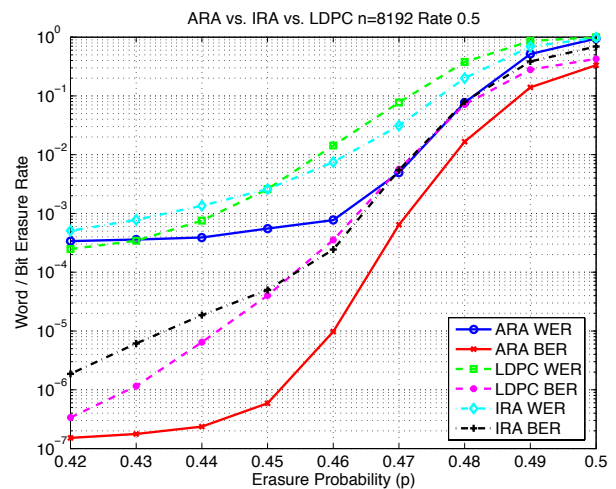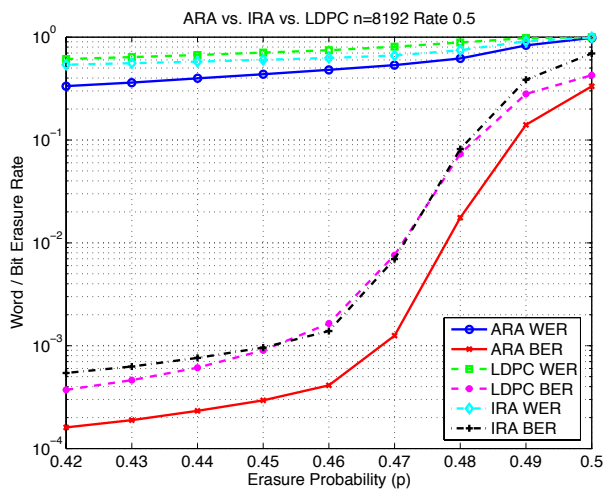
Fig. 11. Simulations for the ensembles of ARA and NSIRA codes constructed from LDPC codes (see Sections VI-A and VI-B), and right-regular LDPC codes [3]. The plots refer to block lengths of 8192 and 65536 bits (see upper and lower plots, respectively) and a design rate of 0.5 bits per channel use. Since the ensemble averaged performance is simulated, high-rate outer codes (rates $\frac{8179}{8192}$ and $\frac{65520}{65536}$, respectively) are used to lower the error floor due to small stopping sets. These outer codes are chosen uniformly at random from the ensemble of the binary linear block codes and their rate loss is neglected.

$1, \ldots, k\}$, as follows. We choose

$$u_j = v_{j-1} + \sum_{i=1}^{k-m} P_{j,i} v_i$$

because combining this with (1) shows that

$$v_j = \sum_{i=1}^{k-m} P_{j,i} v_i \qquad \Rightarrow \qquad H\underline{v} = \underline{0}.$$

B. Decoding Details

The simulations essentially use Luby's decoding algorithm [2], which starts with the full graph and deletes edges as they become known. This process starts by choosing a degree-1 parity-check node and finding the bit node to which it attaches. This bit node is declared known, and all of the edges which are attached to it are removed from the graph. Without loss of generality, one can assume that the all-zero codeword was

transmitted, and the received bits are erased with probability p.

One advantage of Luby's decoding algorithm for LDPC codes is that it averages over all random graphs without explicitly constructing each graph. It does this simply by tracking the number of nodes of each degree throughout the decoding process. For ARA codes, we first use graph reduction (from Section II-C.2) to convert the ARA graph into an LDPC graph. This reduction is not done in the average sense, but instead by explicitly placing erasures and combining nodes. This step implicitly averages over all orderings of the systematic and parity bits. Next, the resulting LDPC code is decoded using Luby's algorithm. This step implicitly averages over all random graphs. Therefore, this simulation technique averages over the entire ARA ensemble. Careful design of the graph can only improve performance.

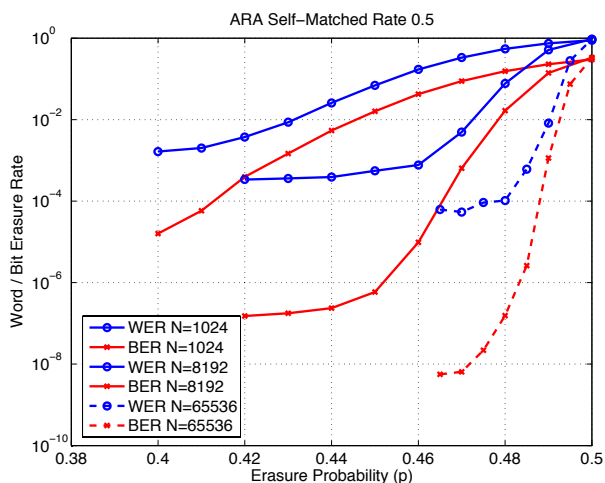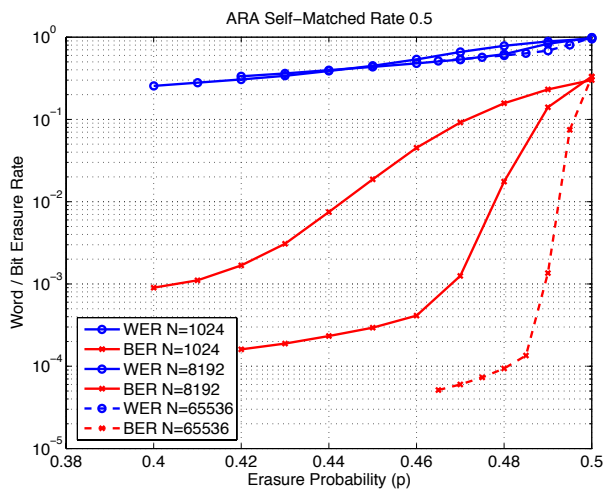If decoding terminates before all bits are known, then the

Fig. 12. Simulations for the ensemble of ARA codes in Section VI-A, constructed from the ensemble of LDPC codes whose degree distributions are matched to themselves. The design rate of the ensemble is 0.5 bits per channel use. The upper plot refers to the case where there is no outer high-rate code, and the lower plot refers to the case where there is such a code.

Fig. 13. Simulations for the ensemble of self-matched ARA codes whose rate is 0.7 bits per channel use, having high-rate outer codes (the rate of the outer code is $\frac{1014}{1024}$, $\frac{8179}{8192}$ and $\frac{65520}{65536}$, for a block length of 1024, 8192 and 65536 bits, respectively.) The upper plot refers to the case where the ensemble of self-matched ARA codes is directly designed for a rate of 0.7 (without puncturing), and the lower plot refers to the design of the self-matched ARA ensemble for a rate of 0.5, and then increasing the rate to 0.7 by random puncturing of the code bits.

high-rate outer code is decoded. Assume that l bits remain unknown after iterative decoding finishes. Decoding the outer code is equivalent to solving a set of m linear equations with l unknowns. Of course, the parity checks leftover from iterative decoding can be used to increase the number of equations (or reduce the number of unknowns). In particular, it is easy to use one of the leftover degree-2 parity-check nodes to reduce the number of unknowns by one. Such a parity check implies the equality of the two unknowns connected to it. This equality allows one to add two columns of the matrix together and reduce the number of unknowns by one. This process can be continued, but one must be wary of linear dependencies among the degree-2 parity-check nodes. Therefore, if t is the number of linearly independent degree-2 parity-check nodes remaining at the end of iterative decoding, then decoding is successful if and only if the rank of the new $m \times (l-t)$ matrix is $l-t$. We note that each entry in the P matrix is chosen randomly from $\{0, 1\}$ with equal probability. Choosing the P matrix more carefully can only improve performance.

C. Discussion of Results

The results of the computer simulations are discussed in this section. Although all of the ensembles introduced have been simulated, results are presented mainly for the ARA and IRA ensembles which are constructed from self-matched LDPC codes. These codes seem to have the best performance in the waterfall region (probably because their degree distributions decay exponentially fast). The results are compared with simulations of Shokrollahi's check-regular ensemble [3].

Figures 10 and 11 compare the self-matched ARA, the self-matched IRA, and the check-regular LDPC ensembles at rate $\frac{1}{2}$. Figures 10 shows the raw error rates without using a high-rate outer code. Therefore, the error floor is rather severe because the results are averaged over the entire ensemble and no attempt was made to avoid small stopping sets. Figure 11 shows the results when a high-rate outer code is used to

Fig. 14. Simulations for the ensemble of punctured ARA codes in Section VI-A where we compare the case where the self-matched ensemble is designed directly to a rate of 0.7 bits per channel use, versus the case of designing the ensemble for a rate of 0.5 and increasing the rate by puncturing. The upper and lower plots refer to block lengths of 8192 and 65536 bits, respectively. P and NP stand for 'punctured' and 'non-punctured'.

Fig. 15. Simulations for the ALDPC-BR3 ensemble (i.e., bit-regular accumulate-LDPC ensemble whose bit nodes have degree-3); the design rate of this ensemble is set to $\frac{1}{4}$ bits per channel use. The upper plot refers to the case where there is no high-rate outer code, and the lower plot refers to the case where there is such an outer code. The plots refer to block lengths of 1024, 8192 and 65536 bits; for the lower plot and these block lengths, the rate of the outer random code is equal to $\frac{1014}{1024}$, $\frac{8179}{8192}$ and $\frac{65520}{65536}$, respectively.
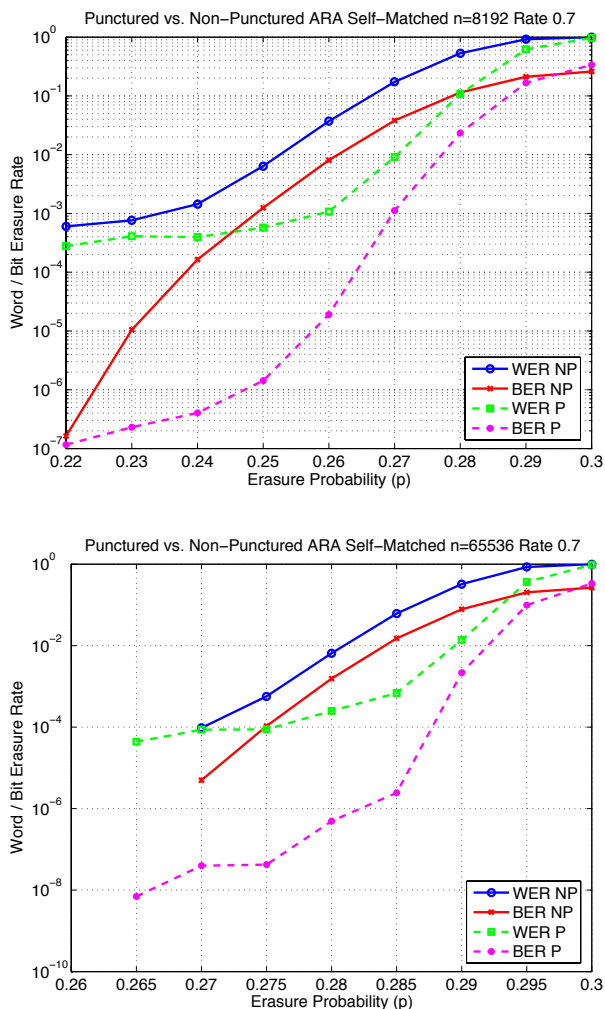
mitigate small stopping sets. The results show that the self-matched ARA ensemble can handle an erasure rate roughly 0.005 larger than the other ensembles while maintaining the same performance. This gain is present throughout the waterfall region and similar for block lengths of 8192 and 65536 bits.

Fig. 12 shows how the performance of the self-matched ARA ensemble varies with block length. The upper plot shows the results without a high-rate outer code and one notices that the word erasure probability never goes below 10%. Even with the high-rate outer code, this ensemble has a word erasure floor due to the fraction of degree-2 bit nodes. As with stable ensembles of irregular LDPC codes, this floor can be made arbitrarily low by expurgating low weight stopping sets and/or adding a stronger outer code.

For a rate $R > \frac{1}{2}$, one can either design self-matched ARA codes directly for this rate or, alternatively, by first designing rate-$\frac{1}{2}$ self-matched ARA codes and puncturing

the code bits up to rate R. The problem with designing the code directly for rate R is that the parameter b in (46) and (47) must be increased in this case and becomes very close to 1. This increases the encoding and decoding complexities and the required maximum degree (see Fig. 5). For example, the rate-$\frac{1}{2}$ ensemble requires only about the 30 first terms of the degree distributions in order to achieve 99% of the design rate while the rate-$\frac{7}{10}$ ensemble requires about the 160 first terms of these degree distributions. Figure 13 shows the performance of these two design methods for rate $\frac{7}{10}$. The results are compared directly in Fig. 14, showing the advantage of the methodology where the ensemble is designed for rate one-half and then punctured to obtain the higher rate. This advantage over the approach of designing self-matched ARA codes without puncturing is exemplified in Fig. 13 either if the ARA code is combined with a high-rate outer code or not.

It can be observed from Figs. 6 and 5 that the ensemble of self-matched ARA codes with the pair of degree distributions

Code Ensemble	Type	Range of p	Bit d.d.	Check d.d.	Gap to capacity (ε)	Complexity as $\varepsilon \to 0$
LDPC	Tornado	$(0,1)$	$\lambda(x) = -\frac{1}{\alpha}\ln(1-x)$	$\rho(x) = e^{\alpha(x-1)}$	$O\left(\frac{1}{M}\right)$	$\chi_D = O\left(\log\frac{1}{\varepsilon}\right)$
LDPC	CR-k	$(0,1)$	$\lambda(x) = 1-(1-x)^{\frac{1}{k-1}}$	$\rho(x) = x^{k-1}$	$O\left(\frac{1}{M}\right)$	$\chi_D = O\left(\log\frac{1}{\varepsilon}\right)$
SIRA	[6, Th. 2]	$(0,1)$	[6, Eq. (33)]	$\rho(x) = e^{\alpha(x-1)}$	$O\left(\frac{1}{M}\right)$	$\chi_E = \chi_D = O\left(\log\frac{1}{\varepsilon}\right)$
NSIRA	CR-3	$(0,0.95)$	[10, Eq. (15)]	$\rho(x) = x^2$	$O\left(\frac{1}{M^{\frac{3}{2}}}\right)$	$\chi_E = \chi_D = \frac{5}{1-p}$
NSIRA	BR-3	$\left(0,\frac{1}{13}\right)$	$\lambda(x) = x^2$	[10, Eq. (10)]	$O\left(\frac{1}{\sqrt{M}}\right)$	$\chi_E = \chi_D = 3+\frac{2}{1-p}$
NSIRA	SM	Eq. (57)	Eq. (56)	Eq. (47)	$O\left(\frac{b^M}{\ln^2(M)}\right)$	Eq. (58)
ARA	CR-3	$(0.616,1)^\dagger$	Eq. (38)	$R(x) = x^3$	$O\left(\frac{1}{M^{\frac{3}{2}}}\right)$	$\chi_E = \chi_D = 3+\frac{5p}{1-p}$
ARA	BR-3	$(0,0.348)$	$L(x) = x^3$	Eq. (37)	$O\left(\frac{1}{M^{\frac{3}{2}}}\right)$	$\chi_E = \chi_D = 6+\frac{2p}{1-p}$
ARA	SM	Eq. (48)	Eq. (46)	Eq. (47)	$O\left(\frac{b^M}{\ln^2(M)}\right)$	Eq. (50)
ALDPC	CR-3	$\left(\frac{12}{13},1\right)^\dagger$	Eq. (41)	$\rho(x) = x^2$	$O\left(\frac{1}{\sqrt{M}}\right)$	$\chi_D = \frac{3(1+p)}{1-p}$
ALDPC	BR-3	$(0.05,1)$	$\lambda(x) = x^2$	Eq. (40)	$O\left(\frac{1}{M^{\frac{3}{2}}}\right)$	$\chi_D = \frac{6}{1-p}$
ALDPC	SM	Eq. (60)†	Eq. (46)	Eq. (59)	$O\left(\frac{b^M}{\ln^2(M)}\right)$	Eq. (61)

TABLE I

LIST OF CAPACITY-ACHIEVING (C.A.) ENSEMBLES FOR THE BEC. IN THIS TABLE, 'BR', 'CR' AND 'SM' STAND FOR 'BIT-REGULAR', 'CHECK-REGULAR' AND 'SELF-MATCHED' DEGREE DISTRIBUTIONS, RESPECTIVELY. THE VALID RANGE OF THE ERASURE PROBABILITY p IS MARKED BY $\dagger$ IF IT CAN BE EXTENDED TO $(0,1)$ VIA PUNCTURING. THE PARAMETER b IN THE D.D. PAIRS OF THE SM ENSEMBLES IS IN THE INTERVAL $(0.9304,1)$.

in (46) and (47) is not suitable for designing codes of low rates; the lower the design rate becomes below $\frac{1}{2}$ bits per channel use, the complexity of this ensemble increases significantly. To this end, we propose the bit-regular accumulate-LDPC (ALDPC) codes (see Section VI-C) as a preferable alternative for designing codes of low rates. The performance of the bit-regular ALDPC ensemble, where the degree of the bit nodes is set to 3, is shown in Fig. 15. In this figure, the performance of these codes is exemplified for moderate to large block lengths, showing the significance of a high-rate outer code in reducing the erasure floor.

VIII. SUMMARY AND CONCLUSIONS

In this section, we provide a table of capacity-achieving (c.a.) code constructions for the binary erasure channel (BEC). With the exception of low-density parity-check (LDPC) codes and systematic irregular repeat-accumulate (SIRA) codes, all of these codes achieve capacity on the BEC with bounded complexity per information bit. Since each of the actual c.a. degree distributions has infinite support, we let M be the truncation depth of the degree distribution (d.d.) and give the gap to capacity as a function of M.

Simulation results show that among all these ensembles, the self-matched ARA ensemble, constructed in Section VI, has the best performance for moderate to large block lengths (considering rates which are at least $\frac{1}{2}$ bit per channel use). We believe the performance advantage of the self-matched ARA ensemble is mainly due to the *exponential* decay of the d.d. coefficients. In contrast, most other ensembles in Table VII-C have d.d. tails which decay polynomially. For low-rate codes, we propose the bit-regular accumulate-LDPC (ALDPC) ensemble in Section VI-C as the preferred alternative (see Fig. 15).

APPENDIX A
DERIVATION OF Q IN (36)

Following the procedure of Section V-A, and starting from (31) gives

$$Q(x) = \frac{\int_0^x \widetilde{\rho}(t)\, dt}{\int_0^1 \widetilde{\rho}(t)\, dt}$$

$$= \frac{3}{p} \int_0^x \widetilde{\rho}(t)\, dt \qquad (A.1)$$

where the last equality follows from (22) and (23), i.e.,

$$\int_0^1 \widetilde{\rho}(t)\, dt = \int_0^1 \widetilde{\lambda}(t)\, dt = \frac{p}{L'(1)} = \frac{p}{3}.$$

The substitution $t = \widetilde{\rho}^{-1}(u)$ and integration by parts gives

$$\int_0^x \widetilde{\rho}(t)\, dt$$

$$= \int_0^{\widetilde{\rho}(x)} u \frac{d}{du}\left(\widetilde{\rho}^{-1}(u)\right) du$$

$$= u\widetilde{\rho}^{-1}(u)\Big|_{u=0}^{\widetilde{\rho}(x)} - \int_0^{\widetilde{\rho}(x)} \widetilde{\rho}^{-1}(u)\, du$$

$$= x\widetilde{\rho}(x) - \int_0^{\widetilde{\rho}(x)} \left[1 - \frac{p^2(1-u)^2}{\left(1-(1-p)(1-u)^3\right)^2}\right] du$$

$$= (x-1)\widetilde{\rho}(x) - \frac{p^2}{3(1-p)} \frac{1}{1-(1-p)(1-u)^3}\Big|_{u=0}^{\widetilde{\rho}(x)}$$

$$= (x-1)\widetilde{\rho}(x) + \frac{p}{3} \frac{1-\left(1-\widetilde{\rho}(x)\right)^3}{1-(1-p)\left(1-\widetilde{\rho}(x)\right)}$$

and the substitution of the last equality in the RHS of (A.1) gives (36).

APPENDIX B
ON THE POWER SERIES EXPANSION OF (46) AND (47)

From (46), we get

$$L(x) = \frac{1}{1-p}\left(1 - \frac{1}{1 + \frac{(1-p)\left[bx+\ln(1-bx)\right]}{p\left[b+\ln(1-b)\right]}}\right)$$

$$= \frac{1}{1-p}\left(1 - \frac{1}{1 - \alpha(p,b)\left[bx+\ln(1-bx)\right]}\right)$$

$$= -\frac{1}{1-p} \sum_{m=1}^{\infty} \alpha^m(p,b) \left[bx+\ln(1-bx)\right]^m$$

$$= \frac{1}{1-p} \sum_{m=1}^{\infty} \left[(-1)^{m-1} \alpha^m(p,b) \left(\sum_{i=2}^{\infty} \frac{b^i x^i}{i}\right)^m\right] \qquad (B.1)$$

where

$$\alpha(p,b) \triangleq -\frac{1-p}{p\left[b+\ln(1-b)\right]}. \qquad (B.2)$$

We note that $\alpha(p,b)$ is positive for $0 < b < 1$ and $0 < p < 1$. From equation (B.1), it is easily verified that the minimal

degree in the power series expansion of $\left[bx+\ln(1-bx)\right]^m$ is x^{2m}, so if $k < 2m$, then the coefficient of x^k is equal to zero. Since m and k are integers, then if follows that the coefficient of x^k in the power series expansion of $\left[bx+\ln(1-bx)\right]^m$ vanishes for $m \geq \lfloor \frac{k}{2} \rfloor + 1$. This gives

$$L_k \triangleq [x^k]L(x)$$

$$= -\frac{1}{1-p} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \alpha^m(p,b) \cdot [x^k]\left[\left(bx+\ln(1-bx)\right)^m\right]$$

where the infinite sum we had before turns to be a finite sum in the last equality. Therefore, the power series expansion of $bx+\ln(1-bx)$ yields that

$$L_k = \frac{1}{1-p} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \left\{(-1)^{m-1} \alpha^m(p,b)\left[[x^k]\left(\sum_{i=2}^{\infty}\frac{b^i x^i}{i}\right)^m\right]\right\}$$

$$= \frac{\alpha(p,b)}{1-p} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \left\{\left(-\alpha(p,b)\right)^{m-1}\left[[x^k]\left(\sum_{i=2}^{\infty}\frac{b^i x^i}{i}\right)^m\right]\right\}.$$

From the last equality, we obtain that

$$L_k = \frac{\alpha(p,b)\, b^k}{1-p} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \left[(-1)^{m-1}\, c_{m,k}\alpha^{m-1}(p,b)\right] \qquad (B.3)$$

where, for $m = 1, 2, \ldots, \lfloor \frac{k}{2} \rfloor$,

$$c_{m,k} \triangleq \sum_{\substack{i_1+i_2+\ldots+i_m = k \\ i_1,i_2,\ldots,i_m \geq 2}} \frac{N_{i_1,i_2,\ldots,i_m}}{i_1\, i_2\, \ldots\, i_m} \qquad (B.4)$$

and $N_{i_1,i_2,\ldots,i_m}$ designates the number of different permutations of the sequence $\{i_1,i_2,\ldots,i_m\}$. From the duality between R and L in our example (see (46) and (47)), then for the calculation of the coefficient R_k in the power series expansion of R we only need to replace p in (B.3) by $1-p$, so

$$R_k = \frac{\alpha(1-p,b)\, b^k}{p} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \left\{(-1)^{m-1}\, c_{m,k}\alpha^{m-1}(1-p,b)\right\} \qquad (B.5)$$

Since $\lambda(x) = \frac{L'(x)}{L'(1)}$, then $\lambda_k = \frac{kL_k}{L'(1)}$. From (B.2) and (50), we obtain the equality $L'(1) = \frac{b^2 p^2 \alpha(p,b)}{(1-p)(1-b)}$, and then it follows from (B.3) and the last two equalities that

$$\lambda_k = \frac{(1-b)\, k\, b^{k-2}}{p^2} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \left\{(-1)^{m-1}\, c_{m,k}\cdot\alpha^{m-1}(p,b)\right\} \qquad (B.6)$$

From the duality between (46) and (47), then by switching p in (B.6) by $1-p$, we obtain that

$$\rho_k = \frac{(1-b)\, k\, b^{k-2}}{(1-p)^2} \sum_{m=1}^{\lfloor \frac{k}{2} \rfloor} \left\{(-1)^{m-1}\, c_{m,k}\cdot\alpha^{m-1}(1-p,b)\right\} \qquad (B.7)$$

so (B.6) and (B.7) provide explicit expressions for the coefficients of the d.d. pair from the edge perspective.

A. A Recursion for the Sequence $\{c_{m,k}\}$

We intend to give now an efficient way to calculate the coefficients $c_{m,k}$ where $1 \leq m \leq \lfloor \frac{k}{2} \rfloor$ (as otherwise $c_{m,k}$ is equal to zero). To this end, we can rewrite $c_{m,k}$ in (B.4) as

$$c_{m,k} \triangleq \sum_{\substack{i_1 + i_2 + \ldots + i_m = k \\ i_1, i_2, \ldots, i_m \geq 2}} \frac{1}{i_1 \, i_2 \, \ldots \, i_m} \qquad (B.8)$$

for $m = 1, 2, \ldots, \lfloor \frac{k}{2} \rfloor$; the difference from (B.4) is that now the order of the terms in the sequence $\{i_1, i_2, \ldots, i_m\}$ is relevant, so effectively every term of the form $\frac{1}{i_1 \, i_2 \, \ldots \, i_m}$ is counted $N_{i_1, i_2, \ldots, i_m}$ times in the sum above. We will show that this sequence satisfies a simple recursive equation. First, it is trivial that

$$c_{1,k} = \frac{1}{k}, \quad k = 2, 3, \ldots \qquad (B.9)$$

Since we impose the constraint $i_1 + i_2 + \ldots i_m = k$ on the calculation of $c_{m,k}$, then we can rewrite $c_{m,k}$ in (B.8) as

$$c_{m,k} = \frac{1}{k} \sum_{\substack{i_1 + i_2 + \ldots + i_m = k \\ i_1, i_2, \ldots, i_m \geq 2}} \frac{i_1 + i_2 + \ldots + i_m}{i_1 \, i_2 \, \ldots \, i_m}$$

$$= \frac{1}{k} \sum_{j=1}^{m} \sum_{\substack{i_1 + i_2 + \ldots + i_m = k \\ i_1, i_2, \ldots, i_m \geq 2}} \frac{i_j}{i_1 \, i_2 \, i_3 \, \ldots i_m}$$

Symmetry considerations imply that each term in the sum over j is equal, so the expression of $c_{m,k}$ can be simplified to

$$c_{m,k} = \frac{m}{k} \sum_{\substack{i_1 + i_2 + \ldots + i_m = k \\ i_1, i_2, i_3 \ldots, i_m \geq 2}} \frac{1}{i_2 \, i_3 \, \ldots \, i_m} \qquad (B.10)$$

Comparing to (B.8), the last summation is over all the possible vectors $\{i_2, i_3, \ldots, i_m\}$ (i.e., i_1 does not appear in the term $\frac{1}{i_2 \, i_3 \, \ldots i_m}$, but only i_2, i_3, $\ldots$, i_m appear there). Since $i_j \geq 2$ for $1 \leq j \leq m$, then the sum $i_2 + i_3 + \ldots i_m$ gets all possible values between $2(m-1)$ and $k-2$ (where the inequality $2(m-1) \leq k-2$ is automatically satisfied because of the assumption that $k \geq 2m$, as otherwise $c_{m,k} = 0$). Since the summations in (B.8) and (B.10) are taken w.r.t. all possible combinations of $\{i_1, i_2, \ldots, i_m\}$ when the two constraints in (B.10) are satisfied, then we obtain the following recursive equation:

$$c_{m,k} = \frac{m}{k} \sum_{j=2(m-1)}^{k-2} c_{m-1,j}. \qquad (B.11)$$

The combination of the recursive equation in (B.11) with the initial values in (B.9) gives an efficient way to calculate the terms of the sequence $\{c_{m,k}\}$. We implemented this algorithm in software.

B. Calculating the Degree Distributions in Section VI-A

We provide here an algorithm to calculate the coefficients in the d.d. pairs (L, R) and (λ, ρ). This algorithm was implemented in software.

1. Calculate the sequence $\{c_{m,k}\}$ with the recursive equation (B.11) and the initial values in (B.9).
2. Calculate $\alpha(p, b)$ from (B.2).
3. Calculate $\{L_k\}$ and $\{\lambda_k\}$ from (B.3) and (B.6), respectively.
4. Calculate $\alpha(1-p, b)$ from (B.2), and $\{R_k\}$ and $\{\rho_k\}$ from (B.5) and (B.7), respectively.

C. Asymptotic Behavior of these Degree Distributions

We wish to find here asymptotic expressions for the sequences $\{L_k\}$, $\{R_k\}$, $\{\lambda_k\}$ and $\{\rho_k\}$ where k is sufficiently large. First we see that it is enough to solve the problem for the sequence $\{L_k\}$, since if there exists a sequence of functions $\{g_k(\cdot, \cdot)\}$ so that $L_k \approx g_k(p, b)$ for $k \gg 1$, then we obtain from (B.3)–(B.7) that the following asymptotic expressions are valid for the other degree distributions when $k \gg 1$:

$$R_k \approx g_k(1-p, b)$$

$$\lambda_k \approx \frac{(1-b)k}{b^2 p^2} \cdot g_k(p, b) \qquad (B.12)$$

$$\rho_k \approx \frac{(1-b)k}{b^2 (1-p)^2} \cdot g_k(1-p, b).$$

We shall therefore focus on the derivation of the asymptotic behavior of the sequence $\{L_k\}$ in the power series expansion of the function $L(\cdot)$ in (46). To this end, we rewrite $L(\cdot)$ in the form

$$L(x) = \frac{bx + \ln(1 - bx)}{p[b + \ln(1-b)] + (1-p)[bx + \ln(1-bx)]}$$

$$= \frac{1}{1-p} \left(\frac{(1-p) \cdot \frac{bx + \ln(1-bx)}{p[b + \ln(1-b)]}}{1 + (1-p) \frac{bx + \ln(1-bx)}{p[b + \ln(1-b)]}} \right)$$

$$= \frac{1}{1-p} \left(\frac{-\alpha(p,b)[bx + \ln(1-bx)]}{1 - \alpha(p,b)[bx + \ln(1-bx)]} \right)$$

$$= \frac{1}{1-p} \left(1 - \frac{1}{1 - \alpha(p,b)[bx + \ln(1-bx)]} \right)$$

where the function α is introduced in (B.2), and is positive for $0 < p < 1$ and $0 < b < 1$. Let us define the function

$$h(z; c) = \frac{1}{1 - c[z + \ln(1-z)]} \qquad (B.13)$$

so that

$$L(x) = \frac{1}{1-p} \cdot \left(1 - h\big(bx; \alpha(p,b)\big) \right). \qquad (B.14)$$

Translating the restriction on α for $L_6 \geq 0$ gives for $c = \alpha(p, b)$ the new restriction

$$0 < c \leq \frac{13 - \sqrt{61}}{9} \approx 0.5766. \qquad (B.15)$$

The following equality therefore holds for $k \geq 2$:

$$L_k = -\frac{b^k}{1-p} \cdot h_k(c), \quad c \triangleq \alpha(p, b) \qquad (B.16)$$

where $\{h_k(c)\}_{k\geq 0}$ is the sequence the coefficients in the power series expansion $h(z;c) = \sum_{k=0}^{\infty} h_k(c)z^k$. In the continuation, we will find the asymptotic behavior of the power series expansion of the function $h(z;c)$ in (B.13), and then use (B.16) to derive the asymptotic behavior of the sequence $\{L_k\}$, and use (B.12) for the derivation of the asymptotic behavior of the other degree distributions.

In [24], a class of methods is presented which enables one to translate, on a term-by-term basis, an asymptotic expression of a function around a dominant singularity into a corresponding asymptotic expansion for the Taylor coefficients of the function. In the continuation of the asymptotic analysis, we rely on [24]. The function $h(z;c)$ has singularities at the points where the following equation is satisfied:

$$1 - c(z + \ln(1-z)) = 0.$$

The closed form solution of the above equation is $z = z_{1,2}$ where

$$z_1 = 1 + W\left(-e^{\frac{1}{c}-1}\right), \quad z_2 = z_1^*$$

and W denotes the Lambert W-function. Since we require that $0 < c \leq \frac{13-\sqrt{61}}{9} \approx 0.5766$, then z_1 and z_2 have an absolute value which is at least equal to 2.074 (we note that the absolute value of $z_{1,2}$ is a monotonic decreasing function of c, so it is achieved when c gets its maximal value within this interval). The dominant singularity of the function $h(z;c)$ is therefore at the point $z = 1$ where the logarithm in (B.13) becomes singular. In the region around the dominant singularity at $z = 1$, the function $h(z;c)$ behaves like

$$h(z;c) = O\left(\ln^{-1}\left(\frac{1}{1-z}\right)\right)$$

so from [24], the asymptotic behavior of the coefficients in the power series expansion of this function is

$$h_k = O\left(\frac{1}{k\ln^2(k)}\right).$$

A more careful analysis which is based on the singularity analysis in [24] shows that

$$h_k = \frac{1}{1-c}\frac{1}{k\left(1+d\ln(k)\right)^2}\left[-1 + \frac{2\gamma}{1+d\ln(k)}\right.$$
$$\left. + \frac{\pi^2 - 6\gamma^2}{2\left(1+d\ln(k)\right)^2}\right] + O\left(\frac{1}{k\ln^5(k)}\right) \quad \text{(B.17)}$$

where $d \triangleq \frac{c}{1-c}$, and $\gamma \approx 0.5772$ designates Euler's constant. From (B.16) and (B.17), we obtain that

$$L_k = \frac{b^k}{k\left(1-\alpha(p,b)\right)(1-p)} \cdot \frac{1}{\left(1+d(p,b)\ln(k)\right)^2} \cdot$$
$$+ \left(1 - \frac{2\gamma}{1+d(p,b)\ln(k)}\right) + O\left(\frac{b^k}{k\ln^4(k)}\right) \quad \text{(B.18)}$$

where the transition from (B.17) to (B.18) follows by substituting $c = \alpha(p,b)$ where $\alpha(\cdot,\cdot)$ is introduced in (B.2), and then based on (B.2) and (B.17)

$$d(p,b) = \frac{\alpha(p,b)}{1-\alpha(p,b)} = -\frac{1-p}{1-p+p[b+\ln(1-b)]}.$$

The asymptotic expression for the other degree distributions (i.e., $\{R_k\}$, $\{\lambda_k\}$ and $\{\rho_k\}$) follow now immediately from the asymptotic behavior of $\{L_k\}$ in (B.18) and the transition to asymptotic behavior of the other d.d. in (B.12). The following degree distributions are therefore obtained:

$$R_k = \frac{b^k}{k\left(1-\alpha(1-p,b)\right)p} \cdot \frac{1}{\left(1+d(1-p,b)\ln(k)\right)^2} \cdot$$
$$\left(1 - \frac{2\gamma}{1+d(1-p,b)\ln(k)}\right) + O\left(\frac{b^k}{k\ln^4(k)}\right)$$

$$\lambda_k = \frac{(1-b)\,b^{k-2}}{\left(1-\alpha(p,b)\right)p^2(1-p)} \cdot \frac{1}{\left(1+d(p,b)\ln(k)\right)^2} \cdot$$
$$\left(1 - \frac{2\gamma}{1+d(p,b)\ln(k)}\right) + O\left(\frac{b^k}{\ln^4(k)}\right)$$

$$\rho_k = \frac{(1-b)\,b^k}{\left(1-\alpha(1-p,b)\right)(1-p)^2 p} \cdot \frac{1}{\left(1+d(1-p,b)\ln(k)\right)^2} \cdot$$
$$\left(1 - \frac{2\gamma}{1+d(1-p,b)\ln(k)}\right) + O\left(\frac{b^k}{\ln^4(k)}\right).$$

Therefore, the asymptotic behavior of the d.d. pairs w.r.t. the nodes and the edges is given by

$$L_k, R_k = O\left(\frac{b^k}{k\ln^2(k)}\right), \qquad \lambda_k, \rho_k = O\left(\frac{b^k}{\ln^2(k)}\right).$$

The parameter b above is chosen according to (49), so it is determined as a function of the erasure probability (p) of the BEC ($0.9304 \leq b < 1$). The closer is the value of p to one-half, then the smaller becomes the value of b, and this accelerates the convergence rate to zero of the d.d. pairs above. It is therefore clear from these two equations that the convergence rate of the d.d. pairs is improved when p becomes closer to one-half (as is also shown in Figs. 7 and 8).

APPENDIX C
A GENERALIZATION OF POLYA'S CRITERION TO DISCRETE DISTRIBUTIONS AND NON-NEGATIVITY PROOFS

A. Polya's Criterion

Polya's criterion [25, p. 509] is a simple condition which is sufficient to imply that a function $F(t)$ is the inverse Fourier transform (i.e., characteristic function) of a non-negative function $f(x)$. It requires that the function F is real, symmetric, non-negative, and convex non-increasing on $[0,\infty)$. This gives a simple method of proving that some functions are indeed characteristic functions.

Example C.1: Consider the function

$$F_\alpha(t) = e^{-|t|^\alpha}, \quad \alpha \in (0,1].$$

Since there is no closed-form expression for the (scaled) Fourier transform

$$f_\alpha(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_\alpha(t)e^{-itx}dt,$$

we cannot verify directly that $F_\alpha(t)$ is the characteristic function of a random variable (i.e., $f_\alpha(x) \geq 0$). It is easy, however, to verify that $F_\alpha(t)$ satisfies Polya's criterion for $\alpha \in (0, 1]$ and this proves that $f_\alpha(x)$ is indeed non-negative.

B. Generalization to Discrete Distributions

We now generalize Polya's criterion to discrete distributions and offer an elementary proof which is substantially different from the standard "tent function" proof outlined in [25, p. 505].

Definition C.1: A function $\gamma(x)$ has *cosine symmetry* if

$$\gamma(x) = \gamma(2\pi + x), \quad \forall x \in \mathbb{R},$$

$\gamma(x) = -\gamma(\pi - x) = -\gamma(\pi + x) = \gamma(2\pi - x)$ for $x \in [0, \pi/2]$, and $\gamma(x) \geq 0$ for $x \in [0, \pi/2]$.

Lemma C.1: Let $\gamma(x)$ be a function with cosine symmetry which is non-negative on $[0, \frac{\pi}{2}]$, and $f(x)$ be a real and convex function for $x \in [0, 2\pi)$. Then, the integral $\int_0^{2\pi} \gamma(x) f(x) dx$ is non-negative.

Proof: First, rewrite the integral as

$$\int_0^{2\pi} \gamma(x) f(x) dx = \int_0^{\pi/2} \big[\gamma(x) f(x) + \gamma(\pi - x) f(\pi - x)$$
$$+ \gamma(\pi + x) f(\pi + x) + \gamma(2\pi - x) f(2\pi - x) \big] dx.$$

Collecting terms on $\gamma(x)$ allows us to write this

$$\int_0^{\pi/2} \gamma(x) \left[f(x) - f(\pi - x) - f(\pi + x) + f(2\pi - x) \right] dx$$
$$= \int_0^{\pi/2} \gamma(x) g(x) dx$$

where $g(x) \triangleq f(x) - f(\pi - x) - f(\pi + x) + f(2\pi - x)$. Taking the derivative of the function g gives

$$g'(x) = f'(x) + f'(\pi - x) - f'(\pi + x) - f'(2\pi - x)$$
$$= [f'(x) - f'(\pi + x)] + [f'(\pi - x) - f'(2\pi - x)].$$

Notice that $g'(x) \leq 0$ for $x \in [0, \frac{\pi}{2}]$ because the convexity of f on $[0, 2\pi]$ implies that $f'(x) \leq f'(x + \pi)$ and $f'(\pi - x) \leq f'(2\pi - x)$ for $x \in [0, \frac{\pi}{2}]$. Finally, we note that the integral is non-negative because both $\gamma(x)$ and $g(x)$ are non-negative for $x \in [0, \frac{\pi}{2}]$. ∎

Lemma C.2: Let $\gamma(x)$ be a function with cosine symmetry which is non-negative on $[0, \frac{\pi}{2}]$, and let $F(t)$ be a real, symmetric, and convex function for $t \in [0, a]$. Then, for $x > 0$

$$\frac{1}{\pi} \int_0^a F(t) \gamma(tx) dt \geq \frac{1}{\pi} \int_{\frac{2\pi}{x} \lfloor \frac{xa}{2\pi} \rfloor}^a F(t) \gamma(tx) dt.$$

Proof: Using Lemma C.1, we find that the integral over each full cycle of $\gamma(x)$ is non-negative which leaves only the remaining partial cycle near a. We show this using, for $x > 0$, the decomposition

$$\frac{1}{\pi} \int_0^a F(t) \gamma(tx) dt = \frac{1}{\pi} \int_{\frac{2\pi}{x} \lfloor \frac{xa}{2\pi} \rfloor}^a F(t) \gamma(tx) dt$$
$$+ \frac{1}{\pi} \sum_{j=0}^{\lfloor \frac{xa}{2\pi} \rfloor - 1} \left[\int_{\frac{2\pi j}{x}}^{\frac{2\pi(j+1)}{x}} F(t) \gamma(tx) dt \right].$$

For each term in the sum, we can change the variable of integration to get

$$\frac{1}{\pi} \int_{\frac{2\pi j}{x}}^{\frac{2\pi(j+1)}{x}} F(t) \gamma(tx) dt = \frac{1}{\pi x} \int_0^{2\pi} F\left(\frac{2\pi j + u}{x} \right) \gamma(u) du.$$

The non-negativity of these terms follows from Lemma C.1 because $F\left(\frac{2\pi j + u}{x} \right)$ is convex on $u \in [0, 2\pi]$ for $0 \leq j \leq \lfloor \frac{xa}{2\pi} \rfloor - 1$. Since the entire sum is non-negative, the original integral must be larger than the remainder term. ∎

Definition C.2: A function $F(t) : \mathbb{R} \to \mathbb{R}$ satisfies Polya's criterion for $t \in [0, a]$, if and only if it satisfies the following conditions:

1. $F(t) = F(-t)$ (F is a symmetric function)
2. $F(t) \geq 0$ (F is non-negative on the interval)
3. $F'(t) \leq 0$ (F is non-increasing on the interval)
4. $F''(t) \geq 0$,   (F is convex on the interval)

Let $\mathcal{P}_a$ be the set of functions that satisfy Polya's criterion on $[0, a]$.

Theorem C.2: Any function $F \in \mathcal{P}_\infty$ is the characteristic function of a continuous probability distribution, and any function $F \in \mathcal{P}_\pi$ which is 2π-periodic (i.e., $F(x) = F(x + 2\pi)$) is the characteristic function of a discrete probability distribution.

Proof: For the $F \in \mathcal{P}_\infty$ case, we have

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^\infty F(t) e^{-itx} dt$$
$$= \frac{1}{\pi} \int_0^\infty F(t) \cos(tx) dt.$$

First, we note that $f(0) \geq 0$ because $F(t) \geq 0$. Next, we apply Lemma C.2 to this with $\gamma(x) = \cos(x)$, which gives

$$f(x) \geq \lim_{a \to \infty} \frac{1}{\pi} \int_{\frac{2\pi}{x} \lfloor \frac{xa}{2\pi} \rfloor}^a F(t) \cos(tx) dt$$
$$\geq \frac{1}{\pi} \lim_{a \to \infty} \left(a - \frac{2\pi}{x} \left\lfloor \frac{xa}{2\pi} \right\rfloor \right) \min_{\frac{2\pi}{x} \lfloor \frac{xa}{2\pi} \rfloor \leq t \leq a} F(t) \cos(tx)$$
$$\geq - \lim_{a \to \infty} \frac{2}{x} F\left(\frac{2\pi}{x} \left\lfloor \frac{xa}{2\pi} \right\rfloor \right).$$

The last step holds because, for any $t_0 > 0$,

$$\min_{t \in [t_0, t_1]} F(t) \cos(tx) \geq - \max_{t \in [t_0, t_1]} F(t) \geq -F(t_0)$$

since the function $F \in \mathcal{P}_\infty$ is non-negative and monotonic non-increasing. If $\lim_{a \to \infty} F(a) = 0$, then we have $f(x) \geq 0$ for $x \neq 0$. On the other hand, if $\lim_{a \to \infty} F(a) > 0$, then $f(x)$ has a Delta function at $x = 0$. Subtracting this Delta function returns us to the first case and shows that $f(x) \geq 0$ for $x \neq 0$.

For the 2π-periodic $F \in \mathcal{P}_\pi$ case, we have

$$h_k = \frac{1}{2\pi} \int_{-\pi}^\pi F(t) e^{-ikt} dt$$
$$= \frac{1}{\pi} \int_0^\pi F(t) \cos(kt) dt.$$

For $k = 0$, we see that $h_0 \geq 0$ by the non-negativity of $F(t)$ for $t \in [0, \pi]$. For $k > 0$, we apply Lemma C.2 with $\gamma(x) = \cos(x)$ to get

$$h_k \geq \frac{1}{\pi} \int_{\frac{2\pi}{k} \lfloor \frac{k}{2} \rfloor}^\pi F(t) \cos(kt) dt.$$

When k is even, we have $h_k \geq 0$ because the range of integration becomes zero. When k is odd, we have

$$h_k \geq \frac{1}{\pi} \int_{\frac{\pi(k-1)}{k}}^{\frac{\pi(k-1/2)}{k}} F(t) \cos(tk) dt + \frac{1}{\pi} \int_{\frac{\pi(k-1/2)}{k}}^{\pi} F(t) \cos(kt) dt$$

$$= \frac{1}{\pi} \int_0^{\frac{\pi}{2k}} \left[F\left(\frac{\pi(k-1)}{k} + u\right) - F(\pi - u) \right] \cos(ku) du.$$

This gives $h_k \geq 0$ because $\cos(tk) \geq 0$ for $t \in \left[0, \frac{\pi}{2k}\right]$ and $F(t)$ is non-increasing for $t \in [0, \pi]$. ∎

Corollary C.1: Let the function $g : \mathbb{C} \to \mathbb{C}$ be complex analytic on the open unit disc $D = \left\{x \in \mathbb{C} \big| |x| < 1\right\}$, well-defined and continuous on the boundary ∂D, and real on the sub-domain $D \cap \mathbb{R}$. If the function $h(x) \triangleq \mathrm{Re}\left\{2g\left(e^{ix}\right)\right\}$ is convex on $[0, \pi]$ and $g(0) \geq 0$, then the power series expansion of g about zero has non-negative coefficients. Furthermore, if $\lim_{x \to 0} x \, g(1-x) = 0$, then the result holds even if g has a singularity at $x = 1$.

Proof: Since g is real on $D \cap \mathbb{R}$, the power series expansion of g about zero has real coefficients, i.e., $g_n \in \mathbb{R}$ for all $n \geq 0$; therefore $g(x^*) = g(x)^*$ and $h(-x) = h(x)$. Next, we note that h is non-increasing on $[0, \pi]$ because it is convex (i.e., $h''(x) \geq 0$) on this interval and $h'(x) = \mathrm{Re}\left\{2ie^{ix}g'(e^{ix})\right\}$ implies that $h'(\pi) = \mathrm{Re}\left\{-2ig'(-1)\right\} = 0$. Since $g(0) \geq 0$ is the coefficient of x^0 in power series expansion, we can simply subtract it (without affecting the convexity of h) and assume that $g(0) = 0$ without loss of generality.

This allows us to apply Theorem C.2 to the 2π-periodic function h. One caveat is that h may not satisfy $h(x) \geq 0$ for $x \in [0, \pi]$. In this case, revisiting the proof of Theorem C.2 shows that this is needed only to prove $g_0 \geq 0$, and therefore $g(0) \geq 0$ can be used instead.

Now, we will relate the Fourier transform of h to the power series expansion of g. If g is analytic on D and continuous on the entire boundary ∂D, then we can apply Cauchy's theorem to get

$$g_n = \frac{1}{2\pi i} \oint_{\partial D} \frac{g(x)}{x^{n+1}} dx$$
$$= \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{g\left(e^{i\theta}\right)}{e^{i(n+1)\theta}} ie^{i\theta} d\theta$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} g\left(e^{i\theta}\right) e^{-in\theta} d\theta$$
$$\stackrel{(a)}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[g\left(e^{i\theta}\right) + g\left(e^{-i\theta}\right) \right] e^{-in\theta} d\theta$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\theta) e^{-in\theta} d\theta. \tag{C.1}$$

Step (a) follows from the fact that $g(e^{-i\theta})$ contributes nothing to the integral. This is because $g(0) = 0$ and analyticity together imply that $g(x)$ has only positive integer powers of x and $\int_{-\pi}^{\pi} e^{-in\theta} d\theta = 0$ for non-zero integer n.

On the other hand, if g is singular at $x = 1$, then we can modify the contour to avoid $x = 1$. If we choose the new contour to be the boundary of $D \cap \left\{x \in \mathbb{C} \big| |x-1| \leq \varepsilon\right\}$, then

it can be decomposed into two circular arc sections:

$$C_1 = \left\{x \in \mathbb{C} \big| |x| = 1, \ \phi \leq \arg(x) \leq 2\pi - \phi\right\}$$
$$C_2 = \left\{x \in \mathbb{C} \big| |1-x| = \varepsilon, \ -\psi \leq \arg(1-x) \leq \psi\right\}.$$

Therefore, we can write

$$g_n = \frac{1}{2\pi i} \oint_{C_1 \cup C_2} \frac{g(x)}{x^{n+1}} dx$$
$$= \frac{1}{2\pi i} \oint_{C_1} \frac{g(x)}{x^{n+1}} dx$$
$$+ \frac{1}{2\pi} \int_{-\psi}^{\psi} \frac{g(1 - \varepsilon e^{-i\theta}) \varepsilon e^{-i\theta}}{(1 - \varepsilon e^{-i\theta})^{n+1}} d\theta. \tag{C.2}$$

The magnitude of the second term can be upper bounded by

$$\left| \frac{1}{2\pi} \int_{-\psi}^{\psi} \frac{g(1 - \varepsilon e^{-i\theta}) \varepsilon e^{-i\theta}}{(1 - \varepsilon e^{-i\theta})^{n+1}} d\theta \right| \leq \max_{|x| \leq \varepsilon} \left| \frac{x g(1-x)}{(1-x)^{n+1}} \right|.$$

As $\varepsilon \to 0$, the second term vanishes because of the assumption that $\lim_{x \to 0} x g(1-x) = 0$. Furthermore, from (C.1), the first term on the RHS of (C.2) tends, in the limit where $\phi \to 0$, to the Fourier transform of h. Therefore, it follows from (C.1) and Theorem C.2 that $g_n \geq 0$ for any integer $n \geq 0$. ∎

C. Applications: Non-Negativity Proofs for the Degree Distributions of Some Capacity-Achieving Ensembles

In our study of codes on graphs, the validity of the ensembles constructed requires the verification that the degree distributions have non-negative positive power series expansions about $x = 0$.

1) Non-Negativity Proof for the Check-Regular NSIRA Ensemble: In [10, Theorem 2], a capacity-achieving ensemble of non-systematic irregular repeat-accumulate (NSIRA) codes was introduced, achieving the capacity of a BEC under iterative message-passing decoding with bounded complexity per information bit. The d.d. of interest is given by Eq. (39), and the construction is only valid for a particular erasure rate, p, if $g(x) \triangleq \lambda(x)$ has a non-negative power series expansion for that p. Using Corollary C.1, we can define $h(x) = \mathrm{Re}\left\{2g\left(e^{ix}\right)\right\}$ and verify numerically, for any p, that $h''(x) \geq 0$ for $x \in [0, \pi]$, and also $g(0) \geq 0$. This approach appears to work for all $p \in [0, 1)$, hence it verifies the non-negativity of the power series expansion of the left d.d. λ. In [10, Appendix C], the non-negativity of the power series expansion of λ was proved for $p \in [0, 0.95]$ where the proof there was quite involved. This approach extends the verification of the non-negativity of the power series expansion of λ for all $0 \leq p < 1$, thus proving Conjecture 2 in [10, Section 3].

2) Non-Negativity Proof for the Bit-Regular ARA Ensemble: Starting from (37), let

$$g(x) \triangleq \frac{R'(x)}{x}.$$

Since from (34), $\widetilde{\rho}^{-1}(0) = 0$, then $\widetilde{\rho}(0) = 0$, so the first non-negative coefficient of the power series expansion of $\widetilde{\rho}(x)$ about $x = 0$ is the one of x. Since $Q(x)$ in (36) is calculated by an appropriate scaling of the integral of $\widetilde{\rho}$ over the interval $[0, x]$ so that $Q(1) = 1$, then the first non-negative coefficient

of the power series expansion of $Q(x)$ is the one of x^2. Finally, from (32), it follows that also the first non-negative coefficient of the power series expansion of $R(x)$ about $x = 0$ is the one of x^2. Hence, since $R(0) = R'(0) = 0$, then $R(x)$ has a power series expansion about $x = 0$ whose all coefficients are non-negative if and only if the function $g(x)$ possesses this property.

Using Corollary C.1, let us define $h(x) \triangleq \mathrm{Re}\{2g(e^{ix})\}$. It was verified numerically that if $p \leq 0.26$, then $h''(x) \geq 0$ for $x \in [0, \pi]$, and also $g(0) \geq 0$. This therefore proves that if $p \leq 0.26$, then the function $g(x)$ (and hence, also the d.d. $R(x)$) has a power series expansion about $x = 0$ whose coefficients are non-negative.

3) Non-Negativity Proof for the Self-Matched ARA Ensemble: In Section VI-A, we construct capacity-achieving ARA ensembles from self-matched LDPC codes. From (46) and (47), the left and right degree distributions can be rewritten in the form

$$L(x) = \frac{K_1\big(-bx - \ln(1 - bx)\big)}{1 + c_1\big(-bx - \ln(1 - bx)\big)},$$

$$R(x) = \frac{K_2\big(-bx - \ln(1 - bx)\big)}{1 + c_2\big(-bx - \ln(1 - bx)\big)} \quad \text{(C.3)}$$

where

$$K_1 = \frac{-1}{p[b + \ln(1 - b)]}, \quad K_2 = \frac{-1}{(1 - p)[b + \ln(1 - b)]}$$

$$c_1 = \frac{-(1 - p)}{p[b + \ln(1 - b)]}, \quad c_2 = \frac{-p}{(1 - p)[b + \ln(1 - b)]}. \quad \text{(C.4)}$$

Note that $K_{1,2}$ and $c_{1,2}$ are positive for values of p and b in the range $0 < p < 1$ and $0 < b < 1$. Hence, in order to find the region of p and b where both degree distributions (i.e., $L(x)$ and $R(x)$) have non-negative power series expansions about $x = 0$, let us find conditions on the parameter c which ensure that

$$g(x) = \frac{-x - \ln(1 - x)}{1 + c\big(-x - \ln(1 - x)\big)} \quad \text{(C.5)}$$

has a non-negative power series expansion about $x = 0$. The first few terms of the expansion are

$$g(x) = \sum_{n=2}^{7} g_n(c)x^n = \frac{3x^2 + 2x^3}{6} + \frac{(1 - c)x^4}{4} + \frac{(3 - 5c)x^5}{15}$$

$$+ \frac{(12 - 26c + 9c^2)x^6}{70} + \frac{(120 - 308c + 210)x^7}{840} + O\left(x^8\right).$$

Applying Corollary C.1 directly to this function does not work because the function is neither non-increasing nor convex. Instead, we remove the first few terms of the power series and apply Corollary C.1 to

$$\widetilde{g}(x) = \frac{1}{x^8}\left(\frac{-x - \log(1 - x)}{1 + c\left[-x - \log(1 - x)\right]} - \sum_{n=2}^{7} g_n(c)x^n\right).$$

Following the notation in Corollary C.1, let the function h be

$$h(x) \triangleq \mathrm{Re}\left\{2\widetilde{g}(e^{ix})\right\}$$
$$= \left[\widetilde{g}(e^{ix}) + \widetilde{g}(e^{-ix})\right].$$

Numerically, we find that if $c \in [0, 0.6]$, then $h''(x) \geq 0$ for $x \in [0, \pi]$. Checking the first few terms by hand shows that $g_6(c)$ becomes negative first for $c > c^* \triangleq \frac{13 - \sqrt{61}}{9} \approx 0.5766$. Therefore, we find that the original power series expansion of the function g in (C.5) is non-negative if and only if $c \in [0, c^*]$. From (C.3), it follows that the left and right d.d. ($L(x)$ and $R(x)$, respectively) have non-negative power series expansion about $x = 0$ if and only if

$$c_1 \leq \frac{13 - \sqrt{61}}{9}, \quad c_2 \leq \frac{13 - \sqrt{61}}{9}$$

which based on (C.4), is translated to the condition in (48).

Finally, we consider the question of whether the numerical verification of convexity on $[0, \pi]$ is reasonable. While one could also compute any finite number of terms in the power series expansion and verify their non-negativity, we note that the next term could always be negative. This new approach is different because, in many cases, the convexity can be verified numerically in finite time using complex interval arithmetic. The basic approach is by subdividing the interval $[0, \pi]$ into a large number of small overlapping intervals. For each small interval, the function h'' is evaluated using interval arithmetic. If all the values (i.e., the resulting intervals) are non-negative, then this proves that h is convex on $[0, \pi]$. If h also depends on some parameter, then the parameter interval can also be subdivided into small overlapping subintervals. If the first test succeeds for each parameter subinterval, then we have shown that h is convex on $[0, \pi]$ for all parameter values as well.

REFERENCES

[1] D. J. Costello and G. D. Forney, Jr., "Channel coding: the road to channel capacity," submitted to *Proceedings of the IEEE*, November 2006. [Online]. Available: http://www.arxiv.org/abs/cs.IT/0611112.
[2] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Trans. on Information Theory*, vol. 47, pp. 569–584, February 2001.
[3] M. A. Shokrollahi, "New sequences of linear time erasure codes approaching the channel capacity," *Proceedings of the 13th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, Lectures Notes in Computer Science 1719, Springer Verlag, pp. 65–76, 1999.
[4] P. Oswald and A. Shokrollahi, "Capacity-achieving sequences for the erasure channel," *IEEE Trans. on Information Theory*, vol. 48, pp. 3017–3028, December 2002.
[5] H. Jin and R. J. McEliece, "Irregular repeat-accumlate codes," *Proceedings Second International Conference on Turbo Codes and Related Topics*, pp. 1–8, Brest, France, September 2000.
[6] I. Sason and R. Urbanke, "Complexity versus performance of capacity-achieving irregular repeat-accumulate codes on the binary erasure channel," *IEEE Trans. on Information Theory*, vol. 50, pp. 1247–1256, June 2004.
[7] A. Khandekar and R. J. McEliece, "On the complexity of reliable communication on the erasure channel," *Proceedings IEEE 2001 International Symposium on Information Theory*, p. 1, Washington, D.C., USA, June 2001.
[8] 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.
[9] A. Shokrollahi, "Raptor Codes," *IEEE Trans. on Information Theory*, vol. 52, no. 6, pp. 2551–2567, June 2006.
[10] H. D. Pfister, I. Sason and R. Urbanke, "Capacity-achieving ensembles for the binary erasure channel with bounded complexity," *IEEE Trans. on Information Theory*, vol. 51, no. 7, pp. 2352–2379, July 2005.
[11] A. Abbasfar, D. Divsalar, and Y. Kung, "Accumulate-repeat-accumulate codes," *Proceedings IEEE 2004 Conference on Global Communications (GLOBECOM 2004)*, pp. 509–513, Dallas, TX, USA, December 2004.

[12] D. Divsalar, S. Dolinar and J. Thorpe, "Accumulate-repeat-accumulate-accumulate codes," *2004 IEEE 60th Vehicular Technology Conference (VTC2004-Fall)*, vol. 2, pp. 1271–1275, Los Angeles, California, USA, September 2004.

[13] K. Andrews, S. Dolinar, D. Divsalar and J. Thorpe, "Design of low-density parity-check (LDPC) codes for deep-space applications," *IPN Progress Report 42–159*, November 15, 2004. [Online]. Available: http://tmo.jpl.nasa.gov/progress_report/42-159/159K.pdf.

[14] C. H. Hsu and A. Anastasopoulos, "Capacity-achieving codes with bounded graphical complexity on noisy channels," *Proceedings Forty-Third Annual Allerton Conference on Communication, Control and Computing*, pp. 1825–1834, Monticello, Illinois, USA, September 2005. [Online]. Available: http://www.arxiv.org/abs/cs.IT/0509062.

[15] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. on Information Theory*, vol. 27, no. 5, pp. 533–547, September 1981.

[16] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," *IPN Progress Report*, vol. 42, no. 154, August 2003. [Online]. Available: http://tmo.jpl.nasa.gov/progress_report/42-154/154C.pdf.

[17] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. on Information Theory*, vol. 50, no. 11, pp. 599–618, February 2001.

[18] T. J. Richardson and R. L. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 638–656, February 2001.

[19] S. Freundlich, D. Burshtein and S. Litsyn, "Approximately lower tri-angular ensembles of LDPC codes with linear encoding complexity," *Proceedings IEEE 2006 International Symposium on Information Theory*, pp. 821–825, Seattle, WA, USA, July 2006.

[20] G. D. Forney, "Codes on graphs: normal realizations," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 520–548, February 2001.

[21] D.J.C. MacKay and G. Mitchison, and P. L. McFadden, P.L., "Sparse-graph codes for quantum error correction," *IEEE Trans. on Information Theory*, vol. 50, no. 10, pp. 2315–2330, October 2004.

[22] A. Ashikhmin, G. Kramer and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Trans. on Information Theory*, vol. 50, no. 11, pp. 2657–2673, November 2004.

[23] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA, USA: The M.I.T. Press, 1963.

[24] P. Flajolet and A. Odlyzko, "Singularity analysis of generating functions," *SIAM Journal on Discrete Mathematics*, vol. 3, no. 2, pp. 216–240, 1990. [Online]. Available: http://algo.inria.fr/flajolet/Publications/Flod90b.ps.gz.

[25] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, John Wiley, 1971.

Henry D. Pfister (S'99-M'03) received the B.S. degree in physics in 1995, the M.S. degree in electrical engineering in 2000, and the Ph.D. degree in electrical engineering in 2003, all from the University of California, San Diego.

During 2003-2004, he worked as a Senior Engineer for Qualcomm, Inc. in San Diego. In 2005, he was a Postdoctoral Fellow in the School of Computer and Communication Sciences of the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. He joined the faculty of the School of Engineering at Texas A&M University, College Station in 2006, and he is currently an assistant professor in the Department of Electrical and Computer Engineering. His current research interests include information theory, channel coding, and iterative decoding with applications in wireless communications and data storage.

Igal Sason (S'98–M'02) was born in Israel in 1969. He received the B.Sc. and Ph.D. degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1992 and 2001, respectively.

During 1993–1998, he served in the army as an electrical engineer, and was involved in research and development. During 2001–2003, he was a Scientific Collaborator in the Faculty of Computer Science and Communications, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. Since 2003, he has been with the Department of Electrical Engineering at the Technion where he is currently a Senior Lecturer. His research interests include coding theory and information theory.

Dr. Sason is a co-recipient of the 2003 IEEE Communications Society and Information Theory Society Joint Paper Award. During 2003–2006 he was a recipient of the Alon Fellowship for young faculty members and the Career Development Chair.