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

Henry D. Pfister
EPFL- Swiss Federal Institute of Technology
School of Comp. and Comm. Science
Lausanne 1015, Switzerland
Email: henry.pfister@epfl.ch

Igal Sason
Technion – Israel Institute of Technology
Department of Electrical Engineering
Haifa 32000, Israel
Email: sason@ee.technion.ac.il

***Abstract*— The paper introduces ensembles of accumulate-repeat-accumulate (ARA) codes which asymptotically achieve capacity on the binary erasure channel (BEC) with *bounded complexity* (per information bit). It also introduces symmetry properties which play a central role in the construction of various capacity-achieving ensembles for the BEC. The results improve on the tradeoff between performance and complexity provided by the first capacity-achieving ensembles of irregular repeat-accumulate (IRA) codes with bounded complexity (constructed by Pfister, Sason and Urbanke). The superiority of ARA codes with moderate to large block lengths is exemplified by computer simulations comparing their performance with those of previously reported capacity-achieving ensembles of LDPC and IRA codes. ARA codes also have the advantage of being systematic.**

I. Introduction

The study of capacity-achieving (c.a.) sequences of LDPC ensembles for the binary erasure channel (BEC) was initiated by Luby et al. [1] and Shokrollahi [2]. 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 [1], [2], [3]. Jin et al. introduced irregular repeat-accumulate (IRA) codes and presented a c.a. sequence of systematic IRA (SIRA) ensembles [4]. All of the aforementioned codes have one major drawback; their decoding complexity scales like the log of the inverse of the gap (in rate) to capacity, which becomes unbounded as the gap to capacity vanishes (see [5], [6]).

In a previous paper [7], Pfister, Sason and Urbanke presented for the first time two sequences of ensembles of non-systematic IRA (NSIRA) codes which asymptotically (as their block length goes to infinity) achieve capacity on the BEC with bounded complexity per information bit. The new bounded complexity result in [7] is achieved by puncturing bits and allowing in this way a sufficient number of state nodes in the Tanner graph representing the codes.

In this paper, we are interested in constructing c.a. codes for the BEC with bounded complexity per information bit which also perform well at moderate block lengths and are systematic. To this end, we make use of a new channel coding scheme, called "Accumulate-Repeat-Accumulate" (ARA) codes, which was recently introduced by Abbasfar, Divsalar, and Yao [8]. These codes are systematic and have both outstanding performance, as exemplified in [8], [9], and a simple linear-time encoding. After defining 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.

Along the way, we study the symmetry of c.a. degree distributions 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 constructed easily based on the results of [7, Theorems 1 and 2]. This fact and structure was also proposed independently by Hsu and Anastasopoulos in [10]. The interested reader is referred to the full paper version of our work [11].

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 equation for the fixed points of the iterative message-passing decoder (this equation will be called the "DE fixed point equation"). 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 treat them as interleaved serially concatenated codes (see Fig. 1). Since their decoding algorithm

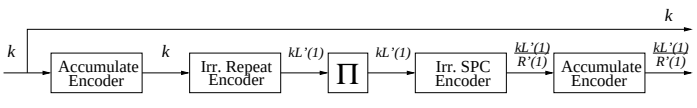

Fig. 1. Block diagram for the systematic ARA ensemble ("Irr." and "SPC" stand for "irregular" and "single-parity check").

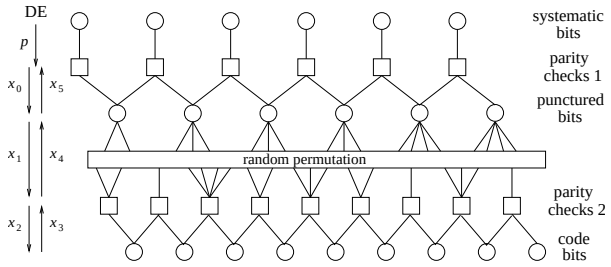

Fig. 2. Tanner graph for the ARA ensemble.

is simply belief propagation on the appropriate Tanner graph (see Fig. 2), one can also view them as sparse-graph codes.

In this work, we consider the ensemble of irregular ARA codes which is the natural generalization of irregular IRA codes [4], [6], [7]. This ensemble differs slightly from those proposed in [8]. 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.

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 and 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 the edges in the graph:

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

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

$$L(x) = \frac{\int_0^x \lambda(t)\,dt}{\int_0^1 \lambda(t)\,dt}\,, \qquad R(x) = \frac{\int_0^x \rho(t)\,dt}{\int_0^1 \rho(t)\,dt}\,. \tag{2}$$

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

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

B. Density Evolution of Systematic ARA Ensembles

In the following, we present two different approaches for the DE analysis of ARA codes for the BEC which, as expected, provide equivalent results. A random code is chosen from the ensemble and a random codeword is transmitted over a BEC with erasure probability p. 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: 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 l^{th} iteration. The technique was introduced in [12] and is known as density evolution (DE). The main assumption of DE 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. 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 graph in Fig. 2, we obtain the following DE equations of the iterative message-passing decoder (for more details, the reader is referred to [11, Section 2.3.1]):

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

Convergence to a fixed point is implied by $\lim_{l\to\infty} x_i^{(l)} \triangleq x_i$ for $i = 0, 1, \ldots, 5$. Eliminating all the variables except $x_1 \triangleq x$ gives the equation:

$$\frac{p^2\, \lambda\left(1 - \left(\frac{1-p}{1-pR(1-x)}\right)^2 \rho(1-x)\right)}{\left[1 - (1-p)\, L\left(1 - \left(\frac{1-p}{1-pR(1-x)}\right)^2 \rho(1-x)\right)\right]^2} = x. \tag{4}$$

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

2) Density Evolution via Graph Reduction: The DE fixed point equation (4) can be also derived using a *graph reduction* approach for the BEC. This approach introduces two new operations on the Tanner graph in Fig. 2 which remove nodes and edges while preserving the information in the graph.

We start by noting that any "code bit" node whose value is not erased by the BEC can be removed from the graph by

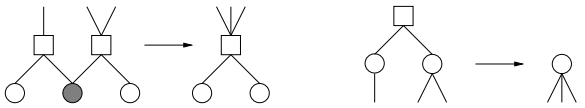

Fig. 3. Graph reduction operation applied to parity-check nodes (left plot) and bit nodes (right plot).

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

$$\widetilde{R}(x) = \sum_{k=0}^{\infty} p^k(1-p)R(x)^{k+1} = \frac{(1-p)R(x)}{1-pR(x)}. \qquad (5)$$

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. Using the symmetry between graph reduction on the information bits and the parity checks, we see that the new d.d. of the "punctured bit" nodes after graph reduction is given by

$$\widetilde{L}(x) = \sum_{k=0}^{\infty}(1-p)^k pL(x)^{k+1} = \frac{pL(x)}{1-(1-p)L(x)}. \qquad (6)$$

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

$$\widetilde{\lambda}(x) = \frac{\widetilde{L}'(x)}{\widetilde{L}'(1)} = \frac{p^2\lambda(x)}{\big(1-(1-p)L(x)\big)^2} \qquad (7)$$

$$\widetilde{\rho}(x) = \frac{\widetilde{R}'(x)}{\widetilde{R}'(1)} = \frac{(1-p)^2\rho(x)}{\big(1-pR(x)\big)^2}. \qquad (8)$$

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

$$\widetilde{\lambda}\big(1-\widetilde{\rho}(1-x)\big) = x. \qquad (9)$$

Based on (7) and (8), the last equation is equivalent to (4).

Remark 1: The tilted degree distributions $\widetilde{\lambda}$ and $\widetilde{\rho}$, which are given in (7) and (8), depend on the erasure probability p of the BEC. For simplicity of notation, we do not include this dependency explicitly in our notation. In Section III, however, the erasure probability is explicitly noted when discussing symmetry properties to distinguish between p and $1-p$.

III. Symmetry Properties of Capacity-Achieving Codes

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 Properties of Capacity-Achieving LDPC Codes

Starting with the DE fixed point equation

$$p\lambda\big(1-\rho(1-x)\big) = x \qquad (10)$$

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.

Following the notation in [3], let $\mathcal{P}$ be the set of d.d. functions, i.e.,

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

Finding a d.d. pair (λ, ρ) which satisfies (10) is typically the first step towards proving that the pair is capacity-achieving. Truncation and normalization issues which depend on the erasure probability of the BEC must also be considered. When $p = 1$, many of these issues disappear, so we denote the set of d.d. pairs which satisfy (10) by

$$\mathcal{C}_{\text{LDPC}} \triangleq \left\{ (\lambda,\rho) \in \mathcal{P}\times\mathcal{P} \mid \lambda\big(1-\rho(1-x)\big) = x \right\}.$$

The *symmetry property* of c.a. LDPC codes (with rate 0) asserts that

$$(\lambda,\rho) \in \mathcal{C}_{\text{LDPC}} \xleftrightarrow{\text{symmetry}} (\rho,\lambda) \in \mathcal{C}_{\text{LDPC}}. \qquad (11)$$

One can prove this result by transforming (10) when $p = 1$. First, we let $x = 1-\rho^{-1}(1-y)$, which gives $\lambda(y) = 1-\rho^{-1}(1-y)$, then we rewrite this expression as $\rho\big(1-\lambda(y)\big) = 1-y$ and finally, let $y = 1-z$ to get $\rho\big(1-\lambda(1-z)\big) = z$. Comparing this with the DE fixed point equation (10) when $p = 1$ shows the symmetry between λ and ρ.

B. Symmetry Properties of ARA Codes

The decoding of an ARA code can be broken into two stages. The first stage transforms the ARA code into an equivalent LDPC code via graph reduction, and the second one decodes the LDPC code. This allows to describe the symmetry property of c.a. ARA codes in terms of the symmetry property of c.a. LDPC codes. For $f \in \mathcal{P}$, let us define

$$\widetilde{f}_p(x) \triangleq \frac{(1-p)^2\ f(x)}{\Big(1-\frac{p\int_0^x f(t)dt}{\int_0^1 f(t)dt}\Big)^2}. \tag{12}$$

One can write the d.d. pair $(\widetilde{\lambda},\widetilde{\rho})$ after graph reduction by combining (2), (7) and (8) which gives $\widetilde{\lambda}=\widetilde{\lambda}_{1-p}$ and $\widetilde{\rho}=\widetilde{\rho}_p$. This allows graph reduction to be interpreted as a mapping $\mathcal{G}_{\text{ARA}}$ from an ARA d.d. pair to an LDPC d.d. pair which can be expressed as

$$(\lambda,\rho) \xrightarrow{\ \mathcal{G}_{\text{ARA}}\ } (\widetilde{\lambda}_{1-p},\widetilde{\rho}_p).$$

The inverse of the graph reduction mapping is represented by a dashed arrow because this inverse mapping, while always well-defined, does not necessarily preserve the non-negativity of d.d. functions.

Referring to ensembles of ARA codes, the set of d.d. pairs which satisfy the DE fixed point equation (4) is given by

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

where the equivalence to (4) follows from (7), (8) and (12).

The symmetry between the bit and check degree distributions of a c.a. ARA ensemble follows from the symmetry relationship in (11), and the equivalence between a d.d. pair (λ,ρ) for ARA codes and the d.d. pair $(\widetilde{\lambda}_{1-p},\widetilde{\rho}_p)$ for LDPC codes. The complete symmetry relationship is therefore given in the following diagram:

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

This symmetry relationship is very useful in order to generate new d.d. pairs which satisfy the DE equality in (9).

C. Symmetry Properties of NSIRA Codes

Now, we consider the graph reduction process and symmetry properties of non-systematic irregular repeat-accumulate (NSIRA) codes (for preliminary material on NSIRA codes, the reader is referred to [7, Section 2]). In this respect, we introduce a new ensemble of codes which we call "Accumulate-LDPC" (ALDPC) codes. These codes are the natural image of NSIRA codes under the symmetry transformation. In fact, this ensemble was discovered by applying the symmetry transformation to previously known c.a. code ensembles. Their decoding graph can be constructed from the ARA decoding graph (see Fig. 2) by removing 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

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

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

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

The symmetry between NSIRA and ALDPC ensembles follows from the symmetry relationship in (11), 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 relationship between a d.d. pair (λ,ρ) for ALDPC codes and the d.d. pair $(\widetilde{\lambda}_{1-p},\rho)$. The symmetry relationship is therefore given in the following diagram.

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

IV. Capacity-Achieving Ensembles with Bounded Complexity: Constructions Based on LDPC Codes

In this section, we introduce a way of constructing c.a. ensembles of ARA codes for the BEC. To this end, we start by choosing a candidate d.d. pair $(\widetilde{\lambda},\widetilde{\rho})$ which satisfies equation (9). Then, it is examined 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 (7) and (8), and testing the non-negativity of the power series of λ and ρ.

We assume that the tilted d.d. $\widetilde{\lambda}$ and $\widetilde{\rho}$ have non-negative power series expansions. Unfortunately, this property does not ensure that 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 calculate the tilted d.d. pair w.r.t. the nodes of the graph using

$$\widetilde{L}(x)=\frac{\int_0^x \widetilde{\lambda}(t)\ dt}{\int_0^1 \widetilde{\lambda}(t)\ dt}\ ,\quad \widetilde{R}(x)=\frac{\int_0^x \widetilde{\rho}(t)\ dt}{\int_0^1 \widetilde{\rho}(t)\ dt}. \tag{13}$$

The original d.d. pair w.r.t. the nodes (i.e., before graph reduction) is calculated from Eqs. (5) and (6), and

$$L(x) = \frac{\widetilde{L}(x)}{p + (1-p)\widetilde{L}(x)}\ , \quad R(x) = \frac{\widetilde{R}(x)}{1 - p + p\widetilde{R}(x)}. \quad (14)$$

Then, we use equation (1) to find (λ, ρ). The critical issue here is to verify whether the functions L and R have non-negative power series expansions.

Capacity-Achieving ARA Ensembles: It is easy to verify that for $0 < b < 1$, the function $f(x) = \frac{(1-b)x}{1-bx}$ belongs to the set $\mathcal{A}$ and also $f(x) = 1 - f^{-1}(1-x)$; in this case, the function f is said to be self-matched. Therefore, based on (9), we examine 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 (13) and (14), we get

$$L(x) = \frac{bx + \ln(1-bx)}{p\,[b + \ln(1-b)] + (1-p)\,[bx + \ln(1-bx)]} \quad (15)$$

$$R(x) = \frac{bx + \ln(1-bx)}{(1-p)\,[b + \ln(1-b)] + p\,[bx + \ln(1-bx)]}. \quad (16)$$

Since the function f we started with is self-matched, the resulting functions L and R in this approach are the same, except that p and $1-p$ are switched. In [11, Appendix C], we find a necessary and sufficient condition so that the degree distributions L and R in (15) and (16), respectively, have both non-negative power series expansions. Based on this proof, we state the following theorem (for a proof, the reader is referred to [11]). We rely here on the Lambert W-function $W(x)$ which is defined to be the w-solution of the equation $we^w = x$.

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 (15) and (16), 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 in terms of the Lambert W-function as

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

The complexity, per information bit, of encoding and decoding is given by $\chi_{\mathrm{E}} = \chi_{\mathrm{D}} = \frac{3-p}{1-p} - \frac{b^2 p}{(1-b)[b+\ln(1-b)]}$.

Acknowledgment: This work was supported by a grant from Intel Israel.

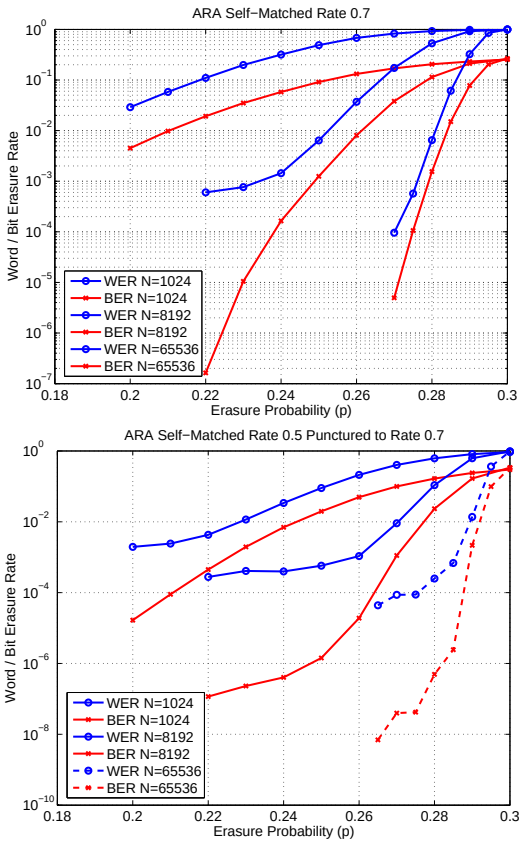

Fig. 4. Simulations for the ensemble of self-matched ARA codes whose rate is 0.7, having high-rate outer codes (the rate of the outer code is $\frac{1014}{1024}$, $\frac{8179}{8192}$ and $\frac{65520}{65536}$ for $N = 1024, 8192$ and 65536 bits, respectively). The upper plot refers to directly designing the rate to 0.7 (without puncturing). The lower plot refers to the design of the ensemble for a rate of 0.5, and increasing its rate to 0.7 by randomly puncturing code bits.

REFERENCES

[1] 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.

[2] 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.

[3] P. Oswald and A. Shokrollahi, "Capacity-achieving sequences for the erasure channel," *IEEE Trans. on Information Theory*, vol. 48, pp. 3017–3028, December 2002.

[4] 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.

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

[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] 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.

[8] A. Abbasfar, D. Divsalar, and Y. Kung, "Accumulate-repeat-accumulate codes," *IEEE 2004 Conference on Global Communications (GLOBECOM 2004)*, pp. 509–513, Dallas, TX, USA, December 2004.

[9] 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.

[10] 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*, Monticello, Illinois, USA, September 2005.

[11] H. Pfister and I. Sason, "Capacity-achieving ensembles of accumulate-repeat-accumulate codes for the erasure channel with bounded complexity," submitted to *IEEE Trans. on Information Theory*, December 2005. [Online]. Available: `http://arxiv.org/abs/cs.IT/0512006`.

[12] 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.