New Results on the Minimum Distance of Repeat Multiple Accumulate Codes

Jörg Kliewer
Department of Electrical and Computer Engineering,
New Mexico State University, Las Cruces, NM 88003, USA
Email: jkliewer@nmsu.edu

Kamil S. Zigangirov, Daniel J. Costello, Jr.
Department of Electrical Engineering,
University of Notre Dame, Notre Dame, IN 46556, USA
Email: {kzigangi, costello.2}@nd.edu

***Abstract*—In this paper we consider the ensemble of codes formed by a serial concatenation of a repetition code with multiple accumulators through uniform random interleavers. Based on finite length weight enumerators for these codes, asymptotic expressions for the minimum distance and an arbitrary number of accumulators larger than one are derived. In accordance with earlier results in the literature, we first show that the minimum distance of RA codes can grow, at best, sublinearly with the block length. Then, for RAA codes and rates of $1/3$ or smaller, it is proved that these codes exhibit linear distance growth with block length, where the gap to the Gilbert-Varshamov bound can be made arbitrarily small by increasing the number of accumulators beyond two. In order to address rates larger than $1/3$, random puncturing of a low-rate mother code is introduced. We show that in this case the resulting ensemble of RAA codes asymptotically achieves linear distance growth close to the Gilbert-Varshamov bound. This holds even for very high rate codes.**

I. Introduction

Since the invention of turbo codes. iterative decoding schemes with even better performance than turbo codes have been designed. Among these are serially concatenated codes, where the simplest code is the repeat-accumulate (RA) code [1] which consists only of a repetition code, an interleaver, and an accumulator. Such a code has the advantage of very low decoding complexity compared to serially concatenated code constructions with convolutional constituent codes. Another benefit of RA codes, compared to powerful code constructions such as LDPC codes, is the extremely low encoding complexity of $O(1)$, whereas LDPC codes have an encoding complexity of $O(g)$, with g much smaller than the block length [2]. This makes RA codes well suited in power-limited environments, for example, for physical layer error correction in battery powered sensor network nodes or in space communications.

In this paper we address multiple serially concatenated RA codes, where design guidelines for more general double serially concatenated codes have been given in [3]. In particular, we focus on the serial concatenation of an outer repetition code with two or more accumulators connected through random interleavers. The resulting code ensemble is then analyzed using the uniform interleaver approach [4] by averaging over all possible interleavers. Our work is mainly motivated by [5], where a similar setup was considered. There, the authors show that for an asymptotically large number of accumulators and uniform interleavers there are codes in the ensemble whose minimum distance grows linearly with block length and which achieve the Gilbert-Varshamov-Bound (GVB). They also provide numerical calculations of the minimum distance for different numbers of accumulators and finite block lengths, but they do not make any statement regarding the asymptotic minimum distance growth rate for the practically more relevant case of a finite (small) number of accumulators. For the case of two accumulators, in [6] it is shown that there exist RAA ensembles with linear distance growth, but besides the existence proof no growth rate is given. In the following, we try to fill this gap and present an analysis of the asymptotic minimum distance growth rate for RA and repeat multiple accumulate codes. The main result of the paper is that for RAA codes and rates equal to $1/3$ or smaller, we show that these codes exhibit linear distance growth with block length, where the gap to the GVB can be made arbitrarily small by increasing the number of accumulators beyond two. In addition, we consider random puncturing at the output of the last accumulator. We show that in this case the resulting ensemble of RAA codes achieves linear distance growth close to the GVB for any code rate smaller than one.

The paper is organized as follows. In Section II we consider the ensemble average weight spectrum of RA codes and show that the minimum distance cannot grow linearly with block length. Section III addresses the case of repeat and multiple accumulate codes, and random puncturing and its effect on the minimum distance is discussed in Section IV.

II. Ensemble average weight spectrum for RA codes

In this section we briefly address the minimum distance of RA codes, and we show that these codes cannot achieve linear distance growth with block length, i.e., they are asymptotically bad. Related results have already been established in [6], [7], and [8], where lower and upper bounds on minimum distance for more general serially concatenated codes have been derived. In order to introduce the notation and for tutorial reasons we restate these results, where our derivation is based on the uniform interleaver approach.

The RA encoder is depicted in Fig. 1. The binary input sequence $\boldsymbol{u}$ has length K and weight w, $\mathcal{R}$ denotes the repetition code of rate $R = 1/q$, which leads to a code

This work was supported in part by NSF Grant CCF05-15012, NASA Grant NNX07AK53G, and the University of Notre Dame Faculty Research Program.

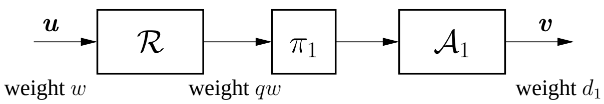

Fig. 1. Repeat-accumulate encoder.

sequence of weight qw and length $N = qK$. The subsequent interleaver π_1 permutes the code sequence. We consider the ensemble of all interleavers by using the uniform interleaver approach [4], where each possible interleaver realization has the same probability $1/2^N$. The permuted output sequence is applied to the recursive convolutional code $\mathcal{A}_1$ with generator polynomial $g(D) = 1/(1+D)$ (accumulator), leading to an output sequence $\boldsymbol{v}$ of weight d_1.

Theorem 1. *The average distance d_1 for an ensemble of RA codes with repetition factor $q \geq 3$ is smaller than $N^{\frac{q-2}{q}}$ with vanishing probability if the block length tends to infinity, i.e.,*

$$\lim_{N\to\infty} \Pr(d_1 \leq N^{\frac{q-2}{q}-\epsilon}) = 0,$$

where ϵ can be made arbitrarily small.

Proof: The conditional probability that a weight d_1 codeword is obtained at the output of the accumulator for a given input weight w can be expressed using the uniform interleaver approach as [1], [4]

$$\Pr(d_1|w) = \frac{\binom{d_1-1}{\lceil\frac{qw}{2}\rceil-1}\binom{qK-d_1}{\lfloor\frac{qw}{2}\rfloor}}{\binom{qK}{qw}}. \quad (1)$$

The total number of input sequences having weight w is $\binom{K}{w}$. Then, the ensemble average input-output weight enumerator function (IOWEF) $\bar{A}_{d_1,w}$, specifying the total number of weight-d_1 output sequences for a fixed input weight w, is given as

$$\bar{A}_{d_1,w} = \binom{K}{w} \Pr(d_1|w). \quad (2)$$

By using Stirling's approximation and the fact that

$$\prod_{\lambda=0}^{\ell}(N-\lambda) \geq \frac{N^\ell}{\varphi(\ell)}, \quad \text{with} \quad \varphi(\ell) := \exp\left(\frac{\ell(\ell+1)}{2N}\right),$$

we can upper bound (2) as

$$\bar{A}_{d_1,w} \leq \frac{N^w\, d_1^{\lceil qw/2\rceil-1}\, N^{\lfloor qw/2\rfloor}}{q^w\, N^{qw}}\, 2^{qw} \left\lceil\frac{qw}{2}\right\rceil \varphi(w). \quad (3)$$

The ensemble average IOWEF giving the total number of output sequences with weight $d \leq \delta$ can now be obtained from $\bar{A}_{d_1\leq\delta,w} = \sum_{d_1=1}^{\delta} \bar{A}_{d_1,w}$. Using (3) then yields

$$\bar{A}_{d_1\leq\delta,w} \leq \frac{N^w\, \delta^{\lceil qw/2\rceil}\, N^{\lfloor qw/2\rfloor}}{N^{qw}} \underbrace{\left(\frac{2^q}{q}\right)^w \left\lceil\frac{qw}{2}\right\rceil \varphi(w)}_{=:f(w,N)}. \quad (4)$$

We now assume that the weight δ can be expressed in terms of the block length N as $\delta = N^\nu$, with $0 < \nu < 1$, where ν is chosen such that

$$\lim_{N\to\infty} \bar{A}_{d_1\leq\delta,w} = 0 \quad (5)$$

is satisfied for all w. Since $\lim_{N\to\infty} f(w,N) = f(w)$ in (4), we only need to satisfy

$$\nu\left\lceil\frac{qw}{2}\right\rceil - (q-1)w + \left\lfloor\frac{qw}{2}\right\rfloor < 0 \quad (6)$$

in order to achieve (5). Two cases can be considered: For even qw, the evaluation of (6) directly leads to $\nu < 1 - 2/q$. For odd qw, we obtain $\nu < \frac{1+w(q-2)}{qw+1}$. We thus have $\nu < \min\left(1-2/q, \frac{1+w(q-2)}{qw+1}\right) = 1 - 2/q$ for all w. This leads to expressing the weight δ as $\delta = N^{1-2/q-\epsilon}$, for any $\epsilon > 0$. Since (5) holds we also have $\lim_{N\to\infty} \bar{A}_{d_1\leq\delta} = 0$, where

$$\bar{A}_{d_1\leq\delta} = \sum_{d_1=1}^{\delta}\sum_{w=1}^{K} \bar{A}_{d_1,w}$$

represents the ensemble average output weight enumerator function (WEF) specifying the overall number of output sequences of weight $d_1 \leq \delta$. Since (see [5], [9])

$$\Pr(d \leq \delta) \leq (\bar{A}_0 - 1) + \bar{A}_{d_1\leq\delta},$$

which is a consequence of the Markov inequality, the theorem is proved. ∎

The following corollary follows immediately.

Corollary 1. *In the ensemble of length-N RA codes with repetition factor $q \geq 3$ and $N \to \infty$ almost all codes have minimum distance lower bounded by*

$$d_{min}^{RA} > N^{\frac{q-2}{q}-\epsilon},$$

where ϵ can be made arbitrarily small.

An upper bound for the minimum distance is given by the following theorem.

Theorem 2. *The minimum distance of the ensemble of RA codes is upper bounded for $N \to \infty$ as*

$$d_{min}^{RA} < O(N^{\frac{q-1}{q}}),$$

where $q \geq 2$.

A proof is given in [8]. An alternative proof can be obtained by employing sphere-packing arguments analogous to the proof of the sphere-packing bound for multiple turbo codes in [10]. From Corollary 1 and Theorem 2 we conclude that the minimum distance of almost all RA codes grows with the block length N as $O(N^\nu)$ where $(q-2)/q - \epsilon \leq \nu < (q-1)/q$.

III. Ensemble Average Weight Spectrum for Repeat Multiple Accumulate Codes

In this section we generalize the encoder from Fig. 1 and consider a serial concatenation of M accumulators $\mathcal{A}_\ell$ with generator polynomials $1/(1+D)$ separated by interleavers π_ℓ, $1 \leq \ell \leq M$. This setup is shown in Fig. 2. In particular, based

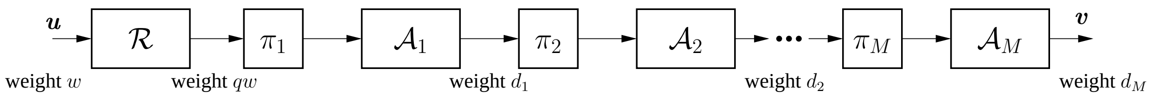

Fig. 2. Repeat multiple accumulate encoder.

on an average weight enumerator analysis, we show that the repeat double accumulate (RAA) ensemble ($M = 2$) for $q \geq 3$ is asymptotically good, i.e., its average minimum distance asymptotically exhibits linear growth with block length. This analysis is then extended to $M \geq 3$, where it is shown that for $q = 2$ asymptotically linear minimum distance growth can be obtained for the RAAA code ensemble.

Definition 1. *The asymptotic spectral shape function [9], [11] is denoted as*

$$r(\rho) := \lim_{N\to\infty} \frac{1}{N} \ln \bar{A}_{\rho N}, \tag{7}$$

where $\rho = d/N$ is the normalized codeword weight and $\bar{A}_d$ is the ensemble average WEF.

From (7), the WEF can be expressed as $\bar{A}_{\rho N} \sim e^{N\,r(\rho)}$. Thus, the important property of $r(\rho)$ in terms of minimum distance can be stated as follows: if the function is negative for some ρ, $\rho_0 > \rho > 0$, then crosses zero and is positive for $\rho > \rho_0$, it follows that, for almost all codes in the ensemble, the minimum distance is lower bounded by $d_{\text{min}} \geq \rho_0 N$ as the block length N tends to infinity.

A. Repeat double accumulate codes

In the following we derive an expression for the spectral shape function for the ensemble of RAA codes. By expurgating this ensemble, we are then able to find a normalized codeword weight ρ_0 such that $r(\rho_0) = 0$ and $dr/d\rho\big|_{\rho=\rho_0} > 0$. For notational convenience, we set $d := d_2$ in the following.

Analogous to (3), the conditional probability that a weight d_1 codeword is obtained at the output of the first accumulator and a weight d codeword at the output of the second accumulator for a given input weight w is

$$\Pr(d_1, d|w) = \frac{\binom{d_1-1}{\lceil \frac{qw}{2}\rceil - 1}\binom{qK-d_1}{\lfloor \frac{qw}{2}\rfloor}\binom{d-1}{\lceil \frac{d_1}{2}\rceil - 1}\binom{qK-d}{\lfloor \frac{d_1}{2}\rfloor}}{\binom{qK}{qw}\binom{qK}{d_1}},$$

which can be rewritten as

$$\Pr(d_1, d|w) = \frac{\binom{qK-qw}{d_1-\lceil \frac{qw}{2}\rceil}\binom{qw}{\lceil \frac{qw}{2}\rceil}\binom{d_1}{\lceil \frac{d_1}{2}\rceil}\binom{qK-d_1}{d-\lceil \frac{d_1}{2}\rceil}}{\binom{qK}{d_1}\binom{qK}{d}} \cdot \frac{\lceil \frac{qw}{2}\rceil\ \lceil \frac{d_1}{2}\rceil}{d_1\, d}. \tag{8}$$

The ensemble average IOWEF $\bar{A}_{d_1,d,w}$ can now be obtained as

$$\bar{A}_{d_1,d,w} = \binom{K}{w}\Pr(d_1, d|w). \tag{9}$$

By using Stirling's approximation, combining (8) and (9) yields

$$\bar{A}_{d_1,d,w} = \exp\left(f(\alpha,\beta,\rho)\,N + O(\ln N)\right), \tag{10}$$

where $\alpha = w/K = qw/N$ and $\beta = d_1/N$ are normalized weights, and the function

$$\begin{aligned} f(\alpha,\beta,\rho) = \frac{H(\alpha)}{q} - H(\beta) - H(\rho) + H\left(\frac{\beta-\alpha/2}{1-\alpha}\right)(1-\alpha) \\ + \alpha\ln 2 + H\left(\frac{\rho-\beta/2}{1-\beta}\right)(1-\beta) + \beta\ln 2, \end{aligned} \tag{11}$$

where $H(\cdot)$ is the binary entropy function. Note that, due to the serial concatenation of accumulators, the quantities α, β, and ρ can be expressed as functions of each other such that $f(\cdot,\cdot,\cdot)$ can be regarded as a one-dimensional function with two of the arguments being fixed.

Furthermore, the ensemble average WEF specifying the number of codewords with weight d, $1 \leq d \leq N/2$, can be upper bounded as

$$\bar{A}_d = \sum_{w=1}^{K}\sum_{d_1=1}^{N} \bar{A}_{d_1,d,w} \leq K\,N \max_{1\leq w<K}\max_{1\leq d_1<N} \bar{A}_{d_1,d,w}. \tag{12}$$

Combining (12), (10), and (7) leads to an upper bound on the asymptotic spectral shape function given by

$$r(\rho) \leq r'(\rho) = \max_{0<\alpha<1}\max_{0<\beta<1} f(\alpha,\beta,\rho). \tag{13}$$

In the following we address the maximization of the function $f(\alpha,\beta,\rho)$ over α and β. A maximum of $f(\cdot)$ can occur on the boundary, i.e., for $\alpha = 0$ and $\beta = 0$ or in the region $\{0 < \alpha < 1, 0 < \beta < 1\}$. In the latter case, the necessary condition is that both the partial derivatives $\partial f/\partial\alpha$ and $\partial f/\partial\beta$ equal zero at the point corresponding to the maximum. The derivative $\partial f/\partial\alpha$ is given as

$$\begin{aligned} \frac{\partial f}{\partial \alpha} = -\frac{1}{q}\ln\left(\frac{\alpha}{1-\alpha}\right) + \frac{1}{2}\ln\left(\frac{\beta-\alpha/2}{1-\alpha}\right) + \\ \frac{1}{2}\ln\left(\frac{1-\beta-\alpha/2}{1-\alpha}\right) + \ln 2, \end{aligned}$$

and $\partial f/\partial\alpha = 0$ is obtained if

$$4\left(\frac{\beta-\alpha/2}{1-\alpha}\right)\left(1-\frac{\beta-\alpha/2}{1-\alpha}\right) = \left(\frac{\alpha}{1-\alpha}\right)^{\frac{2}{q}},$$

which leads to

$$\beta = \frac{1}{2} \pm \frac{1-\alpha}{2}\sqrt{1-\left(\frac{\alpha}{1-\alpha}\right)^{\frac{2}{q}}}. \tag{14}$$

From (14) it can be observed that the extremal points of $f(\cdot)$ occur when $\alpha \leq 1/2$. Also, because of symmetry, we restrict ourselves to $\beta \leq 1/2$. Then, from (14) we have

$$\beta = \frac{1}{2} - \frac{1-\alpha}{2}\sqrt{1-\left(\frac{\alpha}{1-\alpha}\right)^{\frac{2}{q}}}. \tag{15}$$

Evaluating the second derivative $\partial^2 f/\partial\alpha^2 = 0$ at the value of β given in (15) leads to a negative value, and thus $f(\alpha, \beta, \rho)$ exhibits a maximum at this point for any value of ρ. Likewise, the derivative $\partial f/\partial\beta$ can be obtained as

$$\frac{\partial f}{\partial \beta} = \ln\left(\frac{\beta}{1-\beta}\right) - \ln\left(\frac{\beta-\alpha/2}{1-\beta-\alpha/2}\right) + \frac{1}{2}\ln\left(\frac{\rho-\beta/2}{1-\beta}\right) + \frac{1}{2}\ln\left(\frac{1-\rho-\beta/2}{1-\beta}\right) + \ln 2 \tag{16}$$

where $\partial f/\partial\beta = 0$ holds if

$$4\left(\frac{\rho-\beta/2}{1-\beta}\right)\left(1-\frac{\rho-\beta/2}{1-\beta}\right) = \left(\frac{1-\beta}{\beta}\cdot\frac{\beta-\alpha/2}{1-\beta-\alpha/2}\right)^2. \tag{17}$$

Finally, the solution of the quadratic equation (17) in the variable $x = (\rho - \beta/2)/(1-\beta)$ leads to

$$\rho = \frac{1}{2} - \frac{1-\beta}{2}\sqrt{1-\left(\frac{1-\beta}{\beta}\cdot\frac{\beta-\alpha/2}{1-\beta-\alpha/2}\right)^2}, \tag{18}$$

where β satisfies (15). For the solution of (17) we have again invoked symmetry and assumed that $\rho \leq 1/2$. Hence, a triple α, β, ρ maximizing $f(\alpha, \beta, \rho)$ over α and β can be computed via (15) and (18) for a given ρ[1]. An upper bound for $r(\rho)$ is then obtained via (13) and (11).

In Fig. 3 the upper bound $r'(\rho)$ is plotted for rates $R = 1/q$ with $q = 3, 4, 5, 6$ in the interesting region. We observe that for all considered rates the function $r'(\rho)$ exhibits a zero crossing at $\rho = \rho_0$, and therefore almost all codes in the ensemble do not have any codewords with normalized weight $\rho < \rho_0$. In other words, the minimum distance of almost all codes in the ensemble is lower bounded by $d_{\min} \geq \rho_0 N$ and thus grows linearly with block length. In Fig. 3 the values of the GVB for the different rates are shown as vertical dashed lines. The exact values of the zero crossings for $\rho = \rho_0$ are listed in Table I along with the corresponding values of the GVB. We can see that, especially for smaller code rates, the growth rate of the minimum distance of the RAA ensemble is close to the GVB. Note that for $q = 2$ we are not able to show linear growth with block length. This is due to the fact that for the RA code, according to Theorem 1, $q = 2$ does not even guarantee a sublinear distance growth rate. However, by introducing random puncturing of a low rate RAA mother code, the resulting code ensemble exhibits linear distance growth, as will be shown in Section IV.

[1]The structure of equations (15) and (18) suggests a slightly different but equivalent approach, where α is used as the free parameter and β and ρ are determined via (15) and (18).

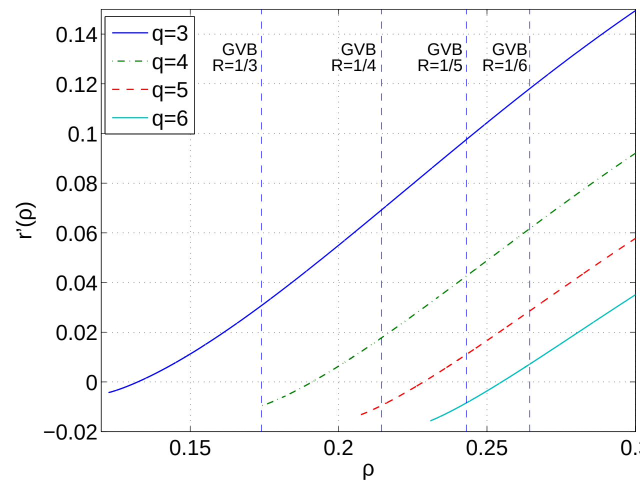

Fig. 3. Upper bound $r'(\rho)$ for the RAA code ensemble with rates $R = 1/3, 1/4, 1/5, 1/6$ and the corresponding values of the GVB.

TABLE I
NORMALIZED MINIMUM DISTANCES ρ_0 AND THE CORRESPONDING VALUES OF THE GVB FOR THE RAA ENSEMBLE WITH DIFFERENT CODE RATES.

q	ρ_0	GVB
3	0.1323	0.1740
4	0.1911	0.2145
5	0.2286	0.2430
6	0.2549	0.2644

B. Generalization to multiple accumulators

We now generalize the results of the previous subsection to repeat-multiple-accumulate codes with $M > 2$, i.e., with more than two accumulators (see Fig. 2). In this general case the conditional probability of the weight vector $\mathbf{d} = [d_1, d_2, \ldots, d_M]$ for a given input weight can be written as

$$\Pr(\mathbf{d}|w) = \Pr(d_1|w) \cdot \prod_{\ell=2}^{M} \frac{\binom{d_\ell - 1}{\lceil \frac{d_{\ell-1}}{2} \rceil - 1}\binom{qK - d_\ell}{\lfloor \frac{d_{\ell-1}}{2} \rfloor}}{\binom{qK}{d_{\ell-1}}},$$

where $\Pr(d_1|w)$ is defined in (3). The ensemble average IOWEF is then given by

$$\bar{A}_{\mathbf{d},w} = \binom{K}{w}\Pr(\mathbf{d}|w) = \exp\left(f(\boldsymbol{\gamma})N + O(\ln N)\right), \tag{19}$$

where in the last step Stirling's approximation is employed. The vector $\boldsymbol{\gamma}$ contains normalized weights and with $d := d_M$ is given by

$$\boldsymbol{\gamma} = [\gamma_0, \gamma_1, \ldots, \gamma_M] = \left[\alpha, \frac{d_1}{N}, \frac{d_2}{N}, \ldots, \rho = \frac{d}{N}\right].$$

The function $f(\boldsymbol{\gamma})$ in (19) can now be written as

$$f(\boldsymbol{\gamma}) = \frac{H(\alpha)}{q} - \sum_{\ell=1}^{M-1} H(\gamma_\ell) + \sum_{\ell=1}^{M} H\left(\frac{\gamma_\ell - \gamma_{\ell-1}/2}{1-\gamma_{\ell-1}}\right)(1-\gamma_{\ell-1}) + \ln(2)\sum_{\ell=0}^{M-1}\gamma_\ell.$$

Analogous to (13) for the RAA case, $f(\boldsymbol{\gamma})$ can be maximized over $\gamma_0, \gamma_1, \ldots, \gamma_{M-1}$, where by employing a similar derivation as in Section III-A the solution for a given $\rho = \gamma_M$ can be expressed by the following recursive set of equations:

$$\gamma_1 = \frac{1}{2} - \frac{1-\alpha}{2}\sqrt{1-\left(\frac{\alpha}{1-\alpha}\right)^{\frac{2}{q}}}, \tag{20}$$

$$\gamma_{\ell+1} = \frac{1}{2} - \frac{1-\gamma_\ell}{2}\sqrt{1-\left(\frac{1-\gamma_\ell}{\gamma_\ell}\cdot\frac{\gamma_\ell-\gamma_{\ell-1}/2}{1-\gamma_\ell-\gamma_{\ell-1}/2}\right)^2} \tag{21}$$

for $1 \leq \ell \leq M-1$.

As an example, for $M = 3$, the resulting values ρ_0 for the zero-crossings of the functions $r'(\rho)$ are shown for $q = 2, 3, 4, 5, 6$ and compared to the GVB in Fig. 4. The

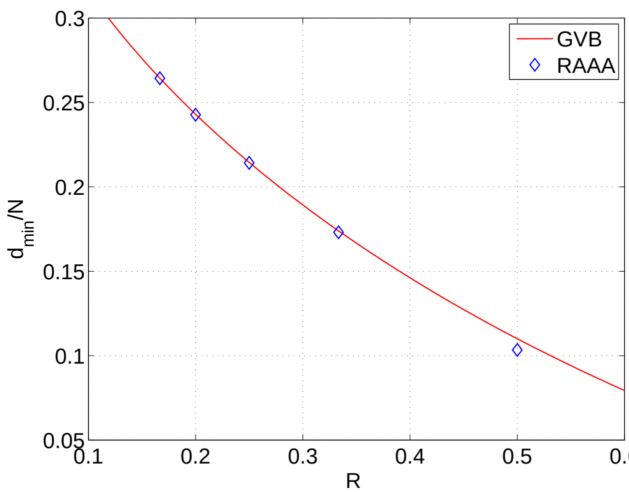

Fig. 4. GVB and corresponding normalized asymptotic minimum distances $\rho_0 = d_{\text{min}}/N$ for the RAAA code ensemble with rates $R = 1/2, 1/3, 1/4, 1/5, 1/6$.

exact values ρ_0 are listed in Table II. We observe that the resulting asymptotic minimum distance growth rates of the RAAA code ensemble for $q \geq 3$ essentially achieve the GVB, which verifies the results obtained in [5] for finite block length. Also, in contrast to the RAA case in Section III-A, the RAAA code ensemble achieves linear distance growth even for the $R = 1/2$ case due to the extra interleaver and accumulator.

IV. Repeat Multiple Accumulate Codes with Random Puncturing

For the RAA code ensemble analyzed in Section III-A we were not able to show linear distance grow for $R = 1/q = 1/2$. This motivates us to employ random puncturing at the

TABLE II
Normalized minimum distances ρ_0 and the corresponding values of the GVB for the RAAA ensemble with different code rates.

q	ρ_0	GVB
2	0.1034	0.1100
3	0.1731	0.1740
4	0.2143	0.2145
5	0.2428	0.2430
6	0.2643	0.2644

output of the inner accumulator in connection with a lower-rate RAA mother code. In particular, random puncturing is useful for achieving rates $R > 1/3$ from the $R = 1/3$ RAA code ensemble, but for the sake of completeness the following analysis is given for the general case of punctured repeat multiple accumulate codes.

Let N' be the number of code symbols after puncturing, $d' = d'_M$ the corresponding codeword weight, and R' the code rate, respectively. We also define the ratios $r = N'/N$ and $\rho' = d'/N$ as the new normalized output weight after puncturing. The conditional probability of a weight-d' sequence after puncturing is now given by

$$\Pr(d'|d, N') = \frac{\binom{d}{d'}\binom{N-d}{N'-d'}}{\binom{N}{N'}}.$$

Again using Stirling's approximation, this can be expressed as

$$\Pr(d'|d, N') = \exp\left\{N\left[H\left(\frac{\rho'}{\rho}\right)\rho + H\left(\frac{r-\rho'}{1-\rho}\right)(1-\rho) - H(r)\right]\right\}.$$

Considering the general case of repeat multiple accumulate codes, the ensemble average IOWEF can now be obtained from (19) as follows:

$$\bar{A}_{\mathbf{d},d',w} = \binom{K}{w}\Pr(\mathbf{d}|w)\Pr(d'|d,N') = \exp\left(F(\boldsymbol{\gamma},\rho',r)\,N + O(\ln N)\right),$$

where $F(\boldsymbol{\gamma},\rho',r) = f(\boldsymbol{\gamma}) + \varphi(\rho',\rho,r)$ and

$$\varphi(\rho',\rho,r) = H\left(\frac{\rho'}{\rho}\right)\rho + H\left(\frac{r-\rho'}{1-\rho}\right)(1-\rho) - H(r). \tag{22}$$

Analogous to (12) and (13) for the RAA ensemble in Section III-A, an upper bound for the normalized spectral shape function $r(d/N') = r(\rho'/r)$ can be obtained as

$$r(\rho'/r) \leq r'(\rho'/r) = \max_{\boldsymbol{\gamma}} F(\boldsymbol{\gamma},\rho',r), \tag{23}$$

where the maximization must now be carried out over all elements of the vector $\boldsymbol{\gamma}$ including $\rho = \gamma_M$. Since $\varphi(\rho',\rho,r)$ in (22) does not depend on the variables $\gamma_0, \gamma_1, \ldots, \gamma_{M-1}$, we can simply employ (20) and (21) for $1 \leq \ell \leq M-1$. In addition, we need to compute the derivative $\partial F/\partial \rho$, which is given as

$$\frac{\partial F}{\partial \rho} = \ln\left(\frac{\rho^2\left(1-\rho-\gamma_{M-1}/2\right)}{(1-\rho)^2\left(\rho-\gamma_{M-1}/2\right)}\right)+\ln\left(\frac{1-\rho-r+\rho'}{\rho-\rho'}\right).$$

Rewriting the condition $\partial F/\partial \rho = 0$ for ρ' as a function of $\boldsymbol{\gamma}$ yields

$$\rho' = \frac{\rho\,(c+1)+r-1}{1+c}, \text{ where } c = \frac{(1-\rho)^2\left(\rho-\gamma_{M-1}/2\right)}{\rho^2\left(1-\rho-\gamma_{M-1}/2\right)}. \tag{24}$$

The upper bound for the asymptotic spectral shape function in (23) can now be computed using (20) and (21), for $1 \leq \ell \leq M-1$, and (24).

Fig. 5 considers the particularly interesting RAA case and shows the normalized minimum distances $\rho_0 = d_{\text{min}}/N$ corresponding to the zero crossings of $r'(\rho'/r)$ for different mother code rates R and punctured code rates R'. We observe that, compared to the unpunctured RAA code ensemble from Section III-A, linear distance growth is also obtained for the punctured ensemble for rates $R' > 1/3$. This behavior is due to the extra randomness added by puncturing the encoder output. We also see that the asymptotic normalized minimum distance tends to be closer to the GVB for higher rates, which is due to the fact that a larger number of puncturing patterns and thus a more "random-like" construction, is available for smaller N'. Table III gives some numerical values of ρ_0 for a rate $R = 1/3$ mother code.

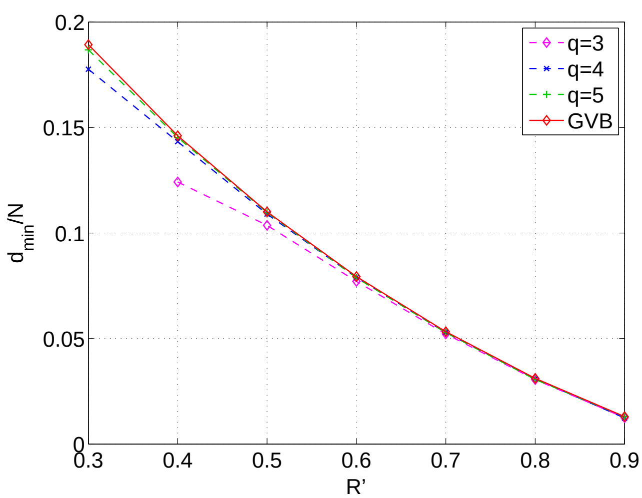

Fig. 5. GVB and corresponding normalized asymptotic minimum distances $\rho_0 = d_{\text{min}}/N$ for the randomly punctured RAA code ensemble with mother codes of rates $R = 1/q = 1/3, 1/4, 1/5$.

V. Concluding Remarks

We have shown that RAA codes for code rates equal to $1/3$ or smaller are asymptotically good in the sense that they achieve asymptotic linear distance growth with block length. Moreover, we have shown that the distance growth rates approach the GVB for small code rates. This extends the results of [5], where linear distance growth is only shown for an infinite number of accumulators. Our results also extend those of [6], where linear distance growth for the RAA ensemble is shown, but no growth rate is given. Further, by introducing random puncturing at the output of the inner accumulator, we demonstrate that the resulting high rate RAA ensembles exhibit linear distance growth, where the growth rate is close to the GVB if the mother code rate is sufficiently low. Finally, in the case of three accumulators, we obtain linear distance growth for code ensembles employing a rate $R = 1/2$ repetition code.

TABLE III
Normalized minimum distances ρ_0 and the corresponding values of the GVB for the punctured RAA ensemble with different code rates employing a mother code of rate $1/3$.

R	ρ_0	GVB
0.4	0.1242	0.1461
0.5	0.1036	0.1100
0.6	0.0771	0.0794
0.7	0.0522	0.0532
0.8	0.0306	0.0311
0.9	0.0125	0.0130

Despite the fact that the repeat multiple accumulate code ensembles considered in this paper proved to be asymptotically good, the convergence behavior of these codes may not be sufficient to provide an iterative decoding threshold close to capacity, as can be seen from the simulation results presented in [5]. However, the results obtained may be useful in constructing similar code ensembles based on simple component codes with low encoding complexity, asymptotically linear distance growth, and good convergence behavior.

References

[1] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," in *Proc. 36th Annual Allerton Conf. Commun., Control, Computing*, Monticello, IL, Sept. 1998, pp. 201–210.

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

[3] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Analysis, design, and iterative decoding of double serially concatenated codes with interleavers," *IEEE J. Sel. Areas in Commun.*, vol. 16, no. 2, pp. 231–244, Feb. 1998.

[4] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 909–926, May 1998.

[5] H. D. Pfister and P. H. Siegel, "The serial concatenation of rate-1 codes through uniform random interleavers," *IEEE Trans. Inf. Theory*, vol. 49, no. 6, pp. 1425–1438, June 2003.

[6] L. Bazzi, M. Mahdian, and D. A. Spielman, "The minimum distance of turbo-like codes," Submitted to *IEEE Trans. Inf. Theory*, May 2003.

[7] N. Kahale and R. Urbanke, "On the minimum distance of parallel and serially concatenated codes," in *Proc. IEEE Int. Symposium on Inform. Theory*, Cambridge, MA, Aug. 1998, p. 31.

[8] A. Perotti and S. Benedetto, "An upper bound on the minimum distance of serially concatenated convolutional codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5501–5509, Dec. 2006.

[9] R. G. Gallager, *Low-density parity-check codes*, MIT Press, Cambridge, MA, 1963.

[10] C. He, M. Lentmaier, D. J. Costello, Jr., and K. S. Zigangirov, "Joint permutor analysis and design for multiple turbo codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4068–4083, Sept. 2006.

[11] H. Jin and R. McEliece, "Coding theorems for turbo code ensembles," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1451–1461, June 2002.