RA Codes Achieve AWGN Channel Capacity*

Hui Jin and Robert J. McEliece
California Institute of Technology
Pasadena, California USA
E-mail: (hui, rjm)@systems.caltech.edu

Abstract.

In ref. [3] we introduced a simplified ensemble of serially concatenated "turbo-like" codes which we called repeat-accumulate, *or RA codes. These codes are very easy to decode using an iterative decoding algorithm derived from belief propagation on the appropriate Tanner graph, yet their performance is scarcely inferior to that of full-fledged turbo codes. In this paper, we prove that on the AWGN channel, RA codes have the potential for achieving channel capacity. That is, as the rate of the RA code approaches zero, the average required bit E_b/N_0 for arbitrarily small error probability with maximum-likelihood decoding approaches* $\log 2$*, which is the Shannon limit. In view of the extreme simplicity of RA codes, this result is both surprising and suggestive.*

1. Introduction.

In ref. [3] we introduced a class of turbo-like codes, the *repeat and accumulate* (RA) codes, which are simple enough to allow a fairly complete theoretical analysis, yet powerful enough to perform nearly as well as full-fledged turbo codes. The general idea is shown in Figure 1. An information block of length k is repeated q times, scrambled by a pseudorandom permutation (interleaver) of size qk, and then encoded by a rate 1 *accumulator*. The accumulator can be viewed as a truncated rate-1 recursive convolutional encoder with transfer function $1/(1+D)$, but we prefer to think of it as a block encoder whose input block $[x_1, \ldots, x_n]$ and output block $[y_1, \ldots, y_n]$ are related by the formula

$$
\begin{aligned}
y_1 &= x_1 \\
y_2 &= x_1 + x_2 \\
y_3 &= x_1 + x_2 + x_3 \\
&\vdots \\
y_n &= x_1 + x_2 + x_3 + \cdots + x_n.
\end{aligned}
\tag{1.1}
$$

In other words, the outputs of the accumulator are the mod-2 partial sums of the inputs. An RA code as just described is therefore a (qk, k) linear block code, but because of the unspecified interleaver, there are a large number, $(qk)!$ to be precise, RA codes with these parameters.

For the theorist, the most important thing about RA codes is that their combinatorial properties are reasonably well understood [3],[4]. For the practitioner, the most important thing is the experimentally verified fact [3] that if an iterative decoding algorithm derived from belief propagation on the appropriate Tanner graph is applied

* This work was supported by NSF grant no. CCR-9804793, and grants from Sony and Qualcomm.

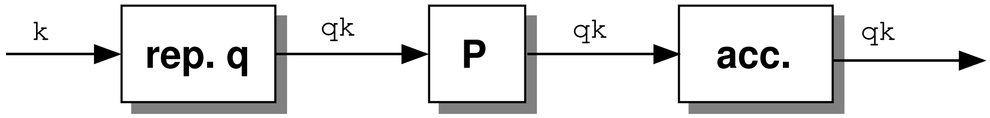

Figure 1. Encoder for a (qk, k) RA code. The "rep. q" component repeats its k-bit input block q times; the "P" block represents an arbitrary permutation of its qk-bit input block; and the "acc." is an accumulator, whose outputs are the mod-2 partial sums of its inputs.

to them, their AWGN channel performance is scarcely inferior to that of full-fledged turbo codes. In this paper, we will give a partial explanation of the latter property, by using the former property to show that on the AWGN channel, RA codes have the *potential* for achieving channel capacity. That is, as the rate of the RA code approaches zero, the average required bit E_b/N_0 for arbitrarily small error probability with maximum-likelihood decoding approaches $\log 2$, which is the Shannon limit. The remaining problem, of course, is to explain why the low-complexity but suboptimal iterative decoding algorithm performs so well.

2. An Ensemble AWGN Coding Theorem for RA Codes.

Our main theorem deals not with any particular RA code, but with the average, or ensemble performance, as the block length approaches infinity. Therefore we begin with a brief description of what we mean by an ensemble of codes.

An *ensemble* of linear codes is a sequence $\mathcal{C}_{n_1}, \mathcal{C}_{n_2}, \ldots$ of sets of linear codes of a common rate R, where $\mathcal{C}_{n_i}$ is a set of (n_i, k_i) codes with $k_i/n_i = R$. We assume that the sequence $n_1, n_2, \ldots$ approaches infinity. If C is an (n, k) code in the ensemble, the *weight enumerator* of C is the list

$$A_0(C), A_1(C), \ldots, A_n(C),$$

where $A_h(C)$ is the number of words of weight h in C. The *average weight enumerator* for the set $\mathcal{C}_n$ is defined as the list

$$\overline{A}_0^{(n)}(C), \overline{A}_1^{(n)}(C), \ldots, \overline{A}_n^{(n)}(C),$$

where

$$\overline{A}_h^{(n)} \stackrel{\text{def}}{=} \frac{1}{|\mathcal{C}_k|} \sum_{C \in \mathcal{C}_k} A_h(C) \qquad \text{for } h = 0, 1, \ldots, n. \tag{2.1}$$

Also, we define the *spectral shape* of $\mathcal{C}_n$,

$$r_n(\delta) \stackrel{\text{def}}{=} \frac{1}{n} \log \overline{A}_{\lfloor \delta n \rfloor}^{(n)}, \qquad \text{for } 0 < \delta < 1, \tag{2.2}$$

and the *ensemble spectral shape* :

$$r(\delta) \stackrel{\text{def}}{=} \lim_{n \to \infty} r_n(\delta) \qquad \text{for } 0 < \delta < 1, \tag{2.3}$$

where we are implicitly assuming that the limit exists.

As a special case, consider the ensemble of rate $R = 1/q$ RA codes. Here the sequence $n_1, n_2, n_3 \ldots$ is $q, 2q, 3q, \ldots$, and the set $\mathcal{C}_{qk}$ consists of $(qk)!$ linear (n, k) block codes, all with $n = qk$. In [3][4] it was shown that the spectral shape for this ensemble of codes is given by the formula

$$r_q(\delta) = \max_{0 \le u \le \min(2\delta, 2-2\delta)} \left\{ f(u,\delta) + \frac{H(u)}{q} \right\}, \tag{2.4}$$

where

$$f(u,\delta) \stackrel{\text{def}}{=} -H(u) + (1-\delta)H\left(\frac{u}{2(1-\delta)}\right) + \delta H\left(\frac{u}{2\delta}\right), \tag{2.5}$$

and $H(u) = -(u \log u + (1-u)\log(1-u))$ is the (natural) entropy function. Figure 2 shows the function $r_q(\delta)$ for $q = 4$.

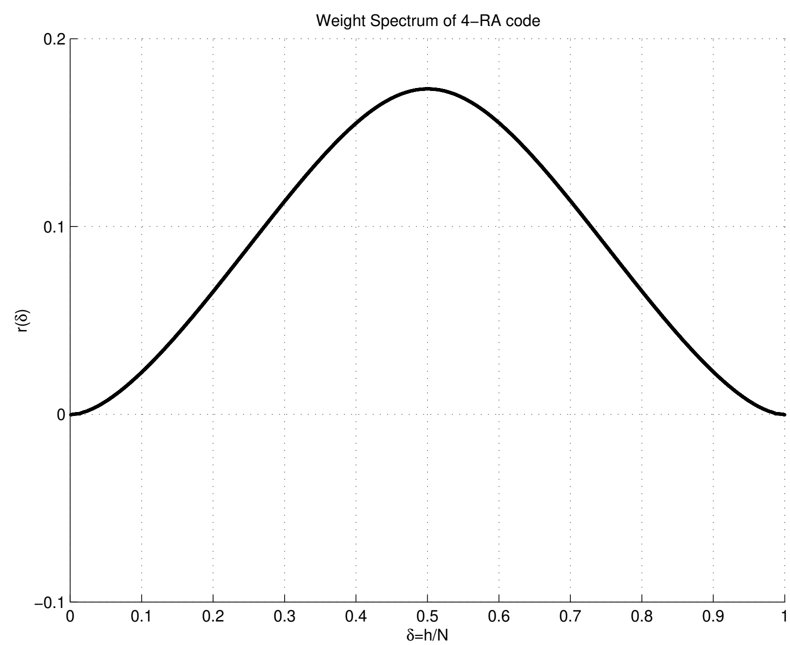

Figure 2. The function $r_4(\delta)$

In [4], using a bound derived in [2], we proved that if we define, for each integer $q \ge 2$,

$$c_o(q) \stackrel{\text{def}}{=} \sup_{0<\delta<1} \frac{1-\delta}{\delta} \frac{1-e^{-2r_q(\delta)}}{2} \tag{2.6}$$

and

$$\gamma_q \stackrel{\text{def}}{=} q c_0(q), \tag{2.7}$$

then if $q \geq 3$, and $E_b/N_0 > \gamma_q$, as $n \to \infty$, the average maximum-likelihood word error probability for the ensemble of rate $1/q$ RA codes approaches 0. A short table of these thresholds, together with the corresponding AWGN Shannon limit, is given below.

q	R	γ_q (dB)	Shannon (dB)
2	1/2	3.384	0.184
3	1/3	0.792	−0.495
4	1/4	−0.052	−0.794
5	1/5	−0.480	−0.963
6	1/6	−0.734	−1.071
7	1/7	−0.900	−1.15
8	1/8	−1.015	−1.210
$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	0	(−1.592)	−1.592

For example, the $q = 3$ line of this table tells us that if $E_b/N_0 > 0.792$ dB, then as $n \to \infty$, the word error probability for the ensemble of rate 1/3 RA codes approaches zero.[1] On the other hand, for $E_b/N_0 < -0.495$ dB, (the Shannon limit for binary codes of rate 1/3) as $n \to \infty$, the word error probability for any sequence of codes of rate 1/3 must approach 1.

In the table, and more clearly in Figure 3, we see that as the rate R approaches zero, the Shannon limit approaches $\log 2 = -1.592$ dB. This is of course well known, and in fact the value -1.592 dB is usually referred to as the Shannon limit for the AWGN channel. The interesting thing for us, however, is that the RA code thresholds also seem to approach -1.592 dB. This empirical observation is in fact true, and it is the object of this paper to prove the following theorem.

2.1 Theorem. *We have* $\lim_{q\to\infty} \gamma_q = \log 2$, *i.e., RA codes achieve the Shannon limit for the AWGN channel.*

We will give a proof of Theorem 2.1 in Section 3. But here we note that it is easy to prove that the limit cannot be smaller than $\log 2$:[2]

2.2 Theorem. *With the threshold* γ_q *defined as above,* $\liminf_{q\to\infty} \gamma_q \geq \log 2$.

Proof: First note that in (2.5), we have $f(u, 1/2) = 0$, so that

$$r_q(1/2) = \max_{0\le u\le 1} H(u)/q = H(1/2)/q = \frac{\log 2}{q}.$$

---

[1] We have included the threshold for $q = 2$ in the table because it can be shown that for rate 1/2 RA codes, if $E_b/N_0 > \gamma_2$, the ensemble *bit* error probability approaches zero, although the word error probability does not.

[2] Of course, Theorem 2.2 follows from the fact that $\log 2$ is the Shannon limit for the channel, but it is interesting to have an "elementary" proof.

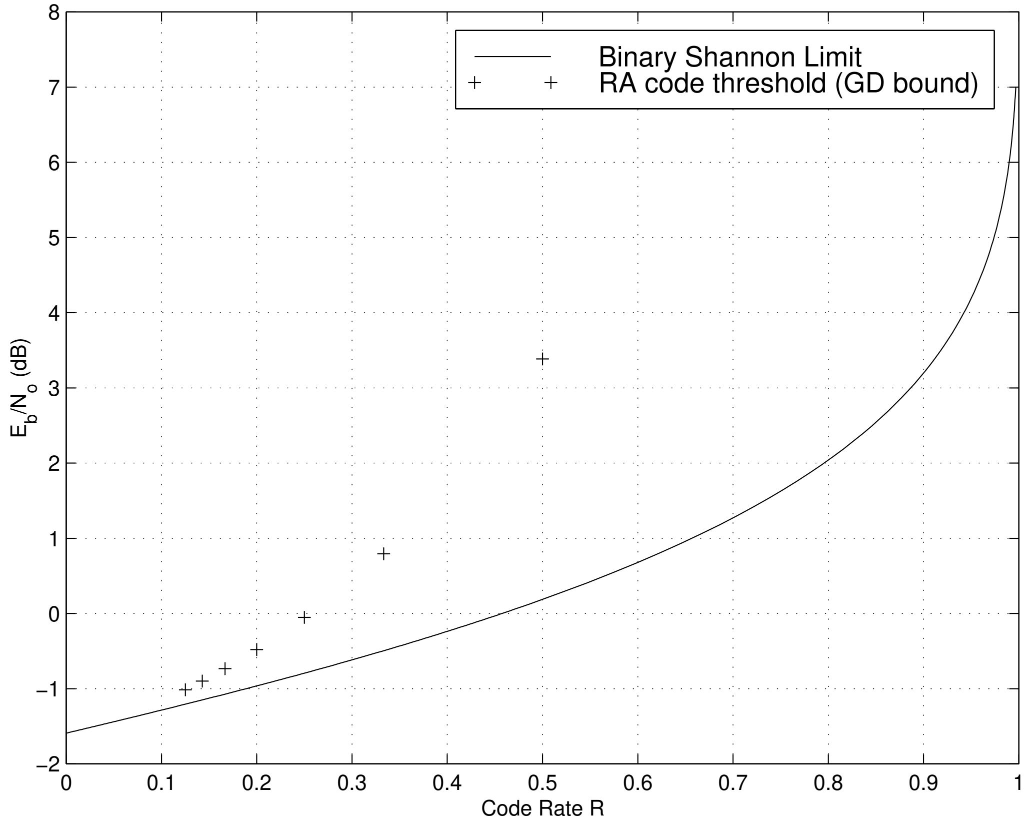

Figure 3. The RA thresholds

Thus by taking $\delta = 1/2$ in the definition (2.6) of $c_0(q)$, we have the lower bound

$$\begin{aligned} c_0(q) &\geq \frac{1-(1/2)}{(1/2)} \frac{1-e^{-2r_q(1/2)}}{2} \\ &= \frac{1-e^{-2\log 2/q}}{2}. \end{aligned}$$

It therefore follows that

$$\gamma_q = qc_0(q) \geq \frac{1-e^{-(2/q)\log 2}}{(2/q)}. \tag{2.8}$$

The desired result now follows by observing that the limit, as $q \to \infty$, of the right side of (2.8) is $\log 2$. ▪

3. Proof of Main Theorem.

In this section, we present the proof of Theorem 2.1. We begin with some preliminary technical results which concern the spectral shape functions $r_q(\delta)$.

3.1 Lemma. *The function $f(u, \delta)$ defined in (2.5) is nonpositive, i.e., for any $(u, \delta) \in [0, 1] \times [0, 1]$, $f(u, \delta) \leq 0$. Furthermore, $f(u, \delta) = 0$ if and only if $\delta = 1/2$ or $u = 0$.*

Proof: Jensen's inequality, and the fact that $H(u)$ is strictly convex $\cap$, implies that for any $\delta \in [0,1]$, and any two numbers $u_1, u_2 \in [0,1]$, we have

$$H(\delta u_1 + (1-\delta)u_2) \geq \delta H(u_1) + (1-\delta)H(u_2),$$

with equality if and only if $u_1 = u_2$ and/or $\delta = 0$ or 1. Letting $u_1 = u/(2\delta)$ and $u_2 = u/(2(1-\delta))$, we obtain the desired result. ▪

3.2 Corollary. *We have, for each $q \geq 2$ and $\delta \in [0,1]$,*

$$r_q(\delta) \leq \frac{\log 2}{q},$$

with equality if and only if $\delta = 1/2$.

Proof: Let $u(q,\delta)$ denote the optimizing value of u in the computation of $r_q(\delta)$, i.e.,

$$u(q,\delta) = \arg \max_{0 \leq u \leq \min(2\delta, 2-2\delta)} \left\{ f(u,q) + \frac{H(u)}{q} \right\}. \tag{3.1}$$

Then by the definition of $r_q(\delta)$, we have

$$r_q(\delta) = f(u(q,\delta),\delta) + H(u(q,\delta))/q.$$

But since by Lemma 3.1, $f(u,\delta) \leq 0$, we have

$$r_q(\delta) \leq H(u(q,\delta))/q \leq \frac{\log 2}{q}. \tag{3.2}$$

For equality to hold in (3.2), we must have, for $u = u(q,\delta)$, both $f(u,\delta) = 0$ and $u = 1/2$. But by Lemma 3.1, if $f(u,\delta) = 0$, then either $u = 0$ or $\delta = 1/2$. Thus equality can hold only when $\delta = 1/2$, as asserted. ▪

3.3 Lemma. *If $u(q,\delta)$ is defined as in (3.1), then*

$$u(q,\delta) \cdot (1 - u(q,\delta))^{q-1} \leq \left(2\sqrt{\delta(1-\delta)}\right)^q. \tag{3.3}$$

Proof: It is easy to check that the "max" in (2.4) does not occur at either endpoint (at $u = 0$ the function is zero and its derivative is infinite, and at $u = \min(2\delta, 2(1-\delta))$, the function is negative), so at the point $u = u(q,\delta)$, the partial derivative of the function in braces on the right side of (2.4), with respect to u, must be zero. This condition works out to be

$$-(1 - 1/q)\log(1-u) - (1/q)\log u + (1/2)\log(2 - 2\delta - u) + (1/2)\log(2\delta - u) = 0,$$

or in exponential form,

$$\frac{\sqrt{(2-2\delta-u)(2\delta-u)}}{(1-u)^{1-1/q}u^{1/q}} = 1. \tag{3.4}$$

But since $0 \leq u \leq \min(2\delta, 2(1-\delta))$, the numerator on the left side of (3.4) is $\leq 2\sqrt{\delta(1-\delta)}$. Therefore at the maximizing point $u = u(q,\delta)$, we must have

$$\frac{2\sqrt{\delta(1-\delta)}}{(1-u)^{1-1/q}u^{1/q}} \geq 1. \tag{3.5}$$

Rearranging this, we get (3.3). ▪

3.4 Corollary. *If* (δ_q) *is a sequence of real numbers in the range* $[0,1]$, *and* $\lim_{q\to\infty}\delta_q = \delta^* \neq 1/2$, *then* $\lim_{q\to\infty} u(q,\delta_q) = 0$.

Proof: If $\delta^* \neq 1/2$, then $2\sqrt{\delta^*(1-\delta^*)} < 1$, and so the right hand side of (3.3) approaches zero as $q \to \infty$. Thus by (3.3), $u(q,\delta)$ must approach either zero or one. But since $u \le \min(2\delta^*, 2-2\delta^*) < 1$, the only possibility is $u(q,\delta_q) \to 0$. ▪

3.5 Corollary. *There exists a* $q_0 \ge 2$ *and* $\delta_0 > 0$ *such that if* $q \ge q_0$ *and* $\delta \le \delta_0$, *then* $u(q,\delta) < \delta^2$.

Proof: Let $\delta < 1/2$. Then $u < \min(2\delta, 2-2\delta) = 2\delta$, so that the left side of (3.3) is lower bounded by the quantity $u(1-2\delta)^q$. Thus from (3.3), we obtain

$$u(q,\delta) \le \left(\frac{2\sqrt{\delta(1-\delta)}}{1-2\delta}\right)^q. \tag{3.6}$$

If the quantity in parenthesis in (3.6) is less than one, which is true for $\delta < 0.14$, then the right hand side is decreasing in q. For example, if $q \ge 6$, we have

$$u(q,\delta) \le \left(\frac{2\sqrt{\delta(1-\delta)}}{1-2\delta}\right)^6, \tag{3.7}$$

which, being of order δ^3, will plainly be $\le \delta^2$ for small enough δ. ▪

In the definition of $c_0(q)$ in (2.6), let $\delta(q)$ be the optimizing value of δ, i.e.,

$$\delta(q) \stackrel{\text{def}}{=} \arg \sup_{0<\delta<1} \frac{1-\delta}{\delta}\frac{1-e^{-2r_q(\delta)}}{2} \tag{3.8}$$

The following proposition is the key to the proof of Theorem 2.1.

3.6 Proposition.

$$\lim_{q\to\infty} \delta(q) = 1/2.$$

Proof: We will show that the set $\{\delta(q)\}$ can have no accumulation points in $[0,1]$ except $1/2$. We begin by proving an inequality, eq. (3.10), below. By definition,

$$\begin{aligned}\gamma_q &= qc_0(q) \\ &= q\frac{1-\delta(q)}{\delta(q)}\frac{1-e^{-2r_q(\delta(q))}}{2}.\end{aligned}$$

But since we have the elementary inequality

$$(1-e^{-2r})/2 \le r, \tag{3.9}$$

it follows that

$$\gamma_q \le \frac{1-\delta(q)}{\delta(q)} q r_q(\delta(q)).$$

But as noted in (3.2), $qr_q(\delta) \leq H(u(q,\delta))$, so that

$$\gamma_q \leq \frac{1-\delta(q)}{\delta(q)} \cdot H(u(q,\delta(q))). \tag{3.10}$$

Now we assume that there is a subsequence of q's for which the $\delta(q)$'s approach a limit $\delta^* \neq 1/2$. There are two cases to consider, $\delta^* \neq 0$, and $\delta^* = 0$. In both cases, we have from Corollary 3.4 that $u(q,\delta(q)) \to 0$. Thus if $\delta^* \neq 0$, it follows from (3.10) that with q restricted to the given subsequence,

$$\lim_{q\to\infty} \gamma_q \leq \frac{1-\delta^*}{\delta^*} H(0) = 0,$$

which contradicts Theorem 2.2.

On the other hand, if $\delta^* = 0$, we have from Corollary 3.5 that for q large enough $u(q,\delta(q)) < \delta(q)^2$. Thus from (3.10) again, for large enough q, we have

$$\gamma_q \leq (1-\delta(q)) \frac{H(\delta(q)^2)}{\delta(q)}.$$

But since $H(x^2)/x \to 0$ as $x \to 0$, with q restricted to the given subsequence,

$$\lim_{q\to\infty} \gamma_q \leq \lim_{q\to\infty} \frac{H(\delta(q)^2)}{\delta(q)} = 0,$$

again contradicting Theorem 2.2.

We have therefore shown that the only possible accumulation point of the set $\{\delta(q)\}$ is $1/2$, which proves that the limit of the $\delta(q)$'s exists and equals $1/2$. ▪

We can now prove the main theorem:

Proof of Theorem 2.1. We have, for every q,

$$\gamma_q = qc_0(q) = q\frac{1-\delta(q)}{\delta(q)} \frac{1-e^{-2r_q(\delta(q))}}{2}.$$

Applying the elementary inequality (3.9), we obtain

$$\gamma_q \leq q\frac{1-\delta(q)}{\delta(q)} r_q(\delta(q)).$$

Now from Corollary 3.2, we know that $r_q(\delta) \leq (\log 2)/q$, so that we have

$$\gamma_q \leq \frac{1-\delta(q)}{\delta(q)} \log 2.$$

But by Proposition 3.6, $\lim_{q\to\infty} \delta(q) = 1/2$, so

$$\limsup_{q\to\infty} \gamma_q \leq \log 2.$$

But we saw in Theorem 2.2 that $\liminf_q \gamma_q \geq \log 2$. Therefore

$$\lim_{q \to \infty} \gamma_q = \log 2. \qquad \blacksquare$$

References.

1. C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo codes," *Proc. 1993 IEEE International Conference on Communications*, Geneva, Switzerland (May 1993), pp. 1064–1070.

2. D. Divsalar, "A simple tight bound on error probability of block codes with applicatiuon to turbo codes," *in preparation.*

3. D. Divsalar, H. Jin, and R. McEliece. "Coding Theorems for 'Turbo-Like' Codes." Proc. 1998 Allerton Conf., pp. 201-210.

4. D. Divsalar, S. Dolinar, H. Jin, and R. McEliece, "AWGN Coding Theorems from Ensemble Weight Enumerators," *in preparation*.