Spectra and Minimum Distances of Repeat Multiple–Accumulate Codes

Chiara Ravazzi and Fabio Fagnani

***Abstract*—In this paper, the ensembles of repeat multiple–accumulate codes (RA^m), which are obtained by interconnecting a repeater with a cascade of m accumulate codes through uniform random interleavers, are analyzed. It is proved that the average spectral shapes of these code ensembles are equal to 0 below a threshold distance ϵ_m and, moreover, they form a nonincreasing sequence in m converging uniformly to the maximum between the average spectral shape of the linear random ensemble and 0. Consequently the sequence ϵ_m converges to the Gilbert–Varshamov (GV) distance. A further analysis allows to conclude that if $m \geq 2$ the RA^m are asymptotically good and that ϵ_m is the typical normalized minimum distance when the interleaver length goes to infinity. Combining the two results it is possible to conclude that the typical distance of the ensembles RA^m converges to the Gilbert–Varshamov bound.**

***Index Terms*—Asymptotic spectral shape, Gilbert–Varshamov distance, input–output weight distribution, multiple serially concatenated codes, uniform random interleavers.**

I. Introduction

REPEAT–accumulate codes have made their appearance in the context of turbo concatenated schemes with the pioneering work [1]. They consist of a repetition code interconnected with a cascade of m simple accumulate codes and all interconnections are through random interleavers. Their particular structure makes them simpler to analyze than generic turbo-like codes. In the past literature (see [2]–[6]) they have received lots of attention and a lot of variations and extensions have been proposed with respect to the original concatenated scheme (see [7]–[13]).

The case with $m = 1$ represents a classical example of serial turbo scheme (just two convolutional codes interconnected by an interleaver [14]). The first theoretical analysis of these codes under maximum likelihood (ML) decoding has been done in [1] and [3]: their main result is an upper bound on the average error probability which vanishes to zero as a negative power of the interleaver length N (if the noise of the channel is not too large). The exponent of N, called interleaver gain exponent, depends only on the repetition parameter. On the other hand, as it happens for all classical serial turbo codes [15], such codes are not asymptotically good: typical minimum distances indeed grow only sublinearly in the length N of the interleaver.

Manuscript received November 12, 2008; revised June 19, 2009. Current version published October 21, 2009. The material in this paper was presented at the Information Theory and Applications Workshop San Diego, CA, January 2008. This work was supported by the NEWCOM++ project.

The authors are with the Department of Mathematics, Politecnico di Torino, Torino I-10129, Italy (e-mail: chiara.ravazzi@polito.it; fabio.fagnani@polito.it).

Communicated by I. Sason, Associate Editor for Coding Theory.
Digital Object Identifier 10.1109/TIT.2009.2030459

However, it was remarked in [5] that when $m \geq 2$ interleaver gain exponent improves when m grows and for the case with $m = 2$ minimum distances present a linear growth in N (see [16]). In [5] it was conjectured that distance performance should increase as m grows and should achieve the Gilbert–Varshamov bound when $m \to +\infty$. At the moment, repeat-accumulate codes for $m \geq 2$ are of limited practical interest because of the fact that iterative decoding shows in these cases very slow convergence [17]. These codes are though of theoretical importance since the understanding of the ingredients which lead to linear growth of the minimum distance may in principle open the possibility to find some simple and powerful variations which possess low encoding complexity and, hopefully, better performance under iterative decoding.

Our work is largely motivated by [18], which contains a first analysis of the distance properties of general multiple serial concatenations encompassing Repeat Multiple–Accumulate (RA^m) codes: an outer code interconnected with a cascade of m rate-1 convolutional codes. The authors show that, for fixed length N and letting $m \to +\infty$, the output weight enumerating functions converge to the weight enumerating function of the random linear coding ensemble. This implies that there exists a sequence of codes whose minimum distance converges to the Gilbert–Varshamov bound but it does not guarantee that the typical distance of all repeat multiple-accumulate codes converges to the Gilbert–Varshamov bound. This difficulty is mathematically due to the fact that the two limits, for $m \to +\infty$ and $N \to +\infty$, cannot be automatically interchanged.

In this paper we undertake a theoretical analysis of the average spectral shapes of repeat multiple–accumulate codes for any m. In particular, we prove that (Theorem 4) for each m, the average spectral shape is symmetric with respect to $1/2$, it is 0 in the intervals $[0, \epsilon_m] \cup [1 - \epsilon_m, 1]$, where $\epsilon_m > 0$ for every $m \geq 2$, and it is strictly positive otherwise. This property was noted in [5] but never proved.

Moreover we show that (Theorem 5), as m grows, the average spectral shapes form a nonincreasing uniformly convergent sequence of functions. Their limit is the maximum between 0 and the average spectral shape of the random linear coding ensemble. As a consequence, the threshold sequence ϵ_m converges to the Gilbert–Varshamov (GV) distance. Since the spectral shapes are not negative but only equal to 0 before ϵ_m, this is not sufficient to conclude that their typical relative minimum distances also reach ϵ_m. However by estimating weight enumerators and using techniques proposed by Jin and McEliece [19], we conclude (Theorem 6) that indeed for such ensembles, minimum distances grow linearly in N and the typical linear growth rate, for a specific m, is exactly given by ϵ_m.

Our theoretical work extends and completes the analysis in [5], [16], [18], and gives a deeper insight into the problem of the distance spectra of RA^m. It also corrects some wrong statements made in [20]–[22] and partially revised in [23].

We now present a brief outline of this paper. Section II is a preliminary section containing all notations and classical concepts needed from coding theory, e.g., weight enumerators and spectral shapes. Section III is devoted to the description of multiple serial concatenation of rate-1 codes and its weight structure; then the emphasis is put on the family of repeat multiple–accumulate codes. Section IV contains a summary of previous results: special attention is devoted to Pfister and Siegel's results collected in [18]. Section V presents, in a formal way, all the original theoretical results presented in this paper: Theorem 4 summarizes the main properties of the spectral shapes of the ensemble for fixed m; Theorem 5 contains the main results about the asymptotic analysis when m goes to infinity; Theorem 6 derives a probabilistic lower bound on the growth of minimum distances when the length N goes to infinity. Sections VI–VIII are technical sections whose results are proved in details. In Section IX, we present some numerical results, showing that our results are apparently true even if we replace the constituent encoders with more general convolutional encoders. Finally, an appendix, containing some combinatorial results and some of the more technical proofs, completes the paper.

A preliminary version of this paper has been presented at the ITA-2008 workshop [24].

II. Preliminaries

We begin our work by fixing notations and reviewing some concepts about block encoders. In particular, we give a general definition of an encoder ensemble and of its weight structure as done in [5], [19], [25]. Moreover, we recall some classical techniques for the estimation of the minimum distance distribution.

A. Notation

In this section, we introduce some notations that we will use in this paper.

- Given a set $\mathbb{S}$, the cardinality of a set $\mathbb{S}$ will be denoted by $|\mathbb{S}|$. Let $\mathbb{N}, \mathbb{R}$ be, respectively, the number sets of nonnegative integers and of real numbers. We let $\mathbb{Z}_2 = \{0, 1\}$ be the Galois field with two elements.
- Conventionally, we consider $\sup(\emptyset) = -\infty$ and $\max(\emptyset) = -\infty$.
- Given $a, b \in \mathbb{R}$ we denote with $a \wedge b$ and $a \vee b$ the minimum and the maximum between them.

B. Weight Enumerators and Spectral Shapes for Coding Ensemble

Let $\mathbb{Z}_2$ be the usual binary field. A $\mathbb{Z}_2$-linear block encoder with rate R and length N is a $\mathbb{Z}_2$-linear map $\mathcal{E} : \mathbb{Z}_2^{\lfloor RN \rfloor} \to \mathbb{Z}_2^N$. Let $\mathcal{C}_{\mathcal{E}} = \mathrm{Im}(\mathcal{E})$ be the associated block code. Given a sequence $\boldsymbol{u} \in \mathbb{Z}_2^N$, denote with $w_H(\boldsymbol{u})$ its Hamming weight, namely the number of its non zero elements. The *minimum distance* of $\mathcal{E}$ (or of $\mathcal{C}_{\mathcal{E}}$) is defined as

$$d_{\min}(\mathcal{E}) = d_{\min}(\mathcal{C}_{\mathcal{E}}) = \min \{ w_H(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{C}_{\mathcal{E}} \setminus \{\boldsymbol{0}\} \}.$$

For any block encoder $\mathcal{E}$ we denote

$$A_d(\mathcal{E}) = |\{\boldsymbol{u} \in \mathbb{Z}_2^{\lfloor RN \rfloor} : w_H(\mathcal{E}(\boldsymbol{u})) = d\}|$$
$$A_{w,d}(\mathcal{E}) = |\{\boldsymbol{u} \in \mathbb{Z}_2^{\lfloor RN \rfloor} : w_H(\boldsymbol{u}) = w, w_H(\mathcal{E}(\boldsymbol{u})) = d\}|.$$

Let now $\mathfrak{E}$ be a set of $\mathbb{Z}_2$-linear encoders with rate R and length N. We can introduce a probabilistic structure on $\mathfrak{E}$ by considering a random encoder chosen uniformly from this set. We then define the *average output* and average input–output weight enumerators of $\mathfrak{E}$, respectively, as follows:

$$\overline{A}_d(\mathfrak{E}) \doteq \frac{1}{|\mathfrak{E}|} \sum_{\mathcal{E} \in \mathfrak{E}} A_d(\mathcal{E})$$
$$\overline{A}_{w,d}(\mathfrak{E}) \doteq \frac{1}{|\mathfrak{E}|} \sum_{\mathcal{E} \in \mathfrak{E}} A_{w,d}(\mathcal{E}).$$

Consider now a sequence $\overline{\mathfrak{E}} = \{\mathfrak{E}_N\}_{N \in \mathbb{N}}$, where each $\mathfrak{E}_N$ is an ensemble of encoders of length N. For each ensemble $\mathfrak{E}_N$, $\overline{A}_d(\mathfrak{E}_N)$ and $\overline{A}_{w,d}(\mathfrak{E}_N)$ are well defined.

We define the Nth spectral shape of $\overline{\mathfrak{E}}$ as

$$r_N(\delta; \overline{\mathfrak{E}}) \doteq \frac{1}{N} \ln \overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{E}_N), \quad \text{for } \delta \in [0,1]$$

and the asymptotic spectral shape of $\overline{\mathfrak{E}}$ as

$$\widehat{r}(\delta; \overline{\mathfrak{E}}) \doteq \limsup_{N \to \infty} r_N(\delta; \overline{\mathfrak{E}}), \quad \text{for } \delta \in [0,1]. \tag{1}$$

Whenever $\overline{\mathfrak{E}}$ is clear from the context, spectral shapes will simply be denoted by $r_N(\delta)$ and $\widehat{r}(\delta)$, respectively.

Example 1 (Random Linear Encoder Ensemble): For fixed $N \in \mathbb{N}$ and rate R, let $\mathfrak{L}_N$ be the ensemble generated by the set of all generator $\lfloor RN \rfloor \times N$-binary matrices. This is equivalent to the ensemble formed by choosing each entry of a random generator matrix independent and identically distributed (i.i.d.) according to a Bernoulli with parameter $1/2$.

The average output weight enumerators for the linear encoder ensemble can be computed to be (see [26] and [27])

$$\overline{A}_d(\mathfrak{L}_N) = \begin{cases} 1 + \frac{2^{\lfloor RN \rfloor} - 1}{2^N} & d = 0 \\ \binom{N}{d} \frac{2^{\lfloor RN \rfloor} - 1}{2^N} & 1 \le d \le N. \end{cases} \tag{2}$$

Since the average number of weight-zero codewords is larger than one, there will always be some encoders in this ensemble which are not invertible.

Let now $\overline{\mathfrak{L}} = \{\mathfrak{L}_N\}_{N \in \mathbb{N}}$. It can be verified that the asymptotic spectral shape has the following expression:

$$\widehat{r}(\delta; \overline{\mathfrak{L}}) = H(\delta) - (1 - R) \ln 2 \tag{3}$$

where $H(\delta) = -\delta \ln \delta - (1 - \delta) \ln(1 - \delta)$ is the binary entropy function on the natural base.

We define the relative Gilbert–Varshamov distance as the unique number $\delta_{\mathrm{GV}}(R) \in [0, 1/2]$ such that

$$H(\delta_{\mathrm{GV}}(R)) = (1 - R) \ln 2. \tag{4}$$

Notice that $\widehat{r}(\delta_{\mathrm{GV}}(R); \overline{\mathfrak{L}}) = 0$ and that the spectral shape is negative for $\delta < \delta_{\mathrm{GV}}(R)$.

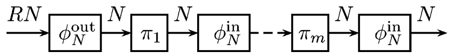

Fig. 1. Coding scheme: Multiple serial concatenated codes.

C. Estimation of Minimum Distance Distribution

One of the uses of the average weight enumerators and the corresponding spectral shapes is to obtain probabilistic information on the minimum distance of the encoders of the ensemble. Indeed the union bound leads to the following estimation (see also Lemma 1 in [18])

$$\mathbb{P}\left(d_{\min}(\mathfrak{E}) < d\right) \leq \sum_{h=1}^{d-1} \overline{A}_h(\mathfrak{E}), \tag{5}$$

where $d_{\min}(\mathfrak{E})$ denotes the minimum distance as a random variable on the ensemble.

We have the following simple result.

Proposition 1: Consider a sequence of encoder ensembles $\overline{\mathfrak{E}} = \{\mathfrak{E}_N\}_{N\in\mathbb{N}}$. If there exists δ_0 such that

$$\sup_{\sigma\leq\delta} \widehat{r}(\sigma;\overline{\mathfrak{E}}) < 0,\ \forall\delta < \delta_0$$

then, for any $\epsilon > 0$,

$$\mathbb{P}\left(d_{\min}(\mathfrak{E}_N) < (\delta_0 - \epsilon)N\right) \xrightarrow{N\to\infty} 0. \tag{6}$$

Proof: Straightforward application of inequality (5) considering that $\overline{A}_h(\mathfrak{E}_N) = \exp\{N r_N(h/N;\overline{\mathfrak{E}})\}$. □

Example 2 (Random Linear Encoder Ensemble): The use of Proposition 1 makes surprisingly easy the estimation of the minimum distance growth rate of a typical binary linear encoder, chosen uniformly from the set $\mathfrak{L}_N$.

Notice that the asymptotic spectral shape given in (3) is negative for $\delta < \delta_{\rm GV}(R)$, crosses zero at $\delta = \delta_{\rm GV}(R)$ then is positive for $\delta \in (\delta_{\rm GV}(R), 1-\delta_{\rm GV}(R))$. By Proposition 1, it follows that for any $\epsilon > 0$

$$\mathbb{P}\left(d_{\min}(\mathfrak{L}_N) < (\delta_{\rm GV}(R) - \epsilon)N\right) \xrightarrow{N\to\infty} 0. \tag{7}$$

III. The Serial Concatenation of Rate-1 Codes Through Uniform Random Interleavers

A. Ensemble Description

In this section, we consider a general class of concatenated coding systems of the type depicted in Fig. 1. The coding scheme is obtained by concatenating an arbitrary binary linear outer block encoder of rate $R < 1$ with a cascade of m identical rate-1 binary inner encoders through uniform random permutations.

Given $N \in R^{-1}\mathbb{N}\cap\mathbb{N}$, the outer encoder is a map such that $\phi_N^{\rm out}: \mathbb{Z}_2^{RN} \to \mathbb{Z}_2^N$ and the inner encoder $\phi_N^{\rm in}: \mathbb{Z}_2^N \to \mathbb{Z}_2^N$ is a rate-1 encoder.

Denote by S_N the group of permutations on N elements. Each $\pi \in S_N$ can naturally be interpreted as a linear isomorphism $\pi: \mathbb{Z}_2^N \to \mathbb{Z}_2^N$. For a fixed $m \in \mathbb{N}$ and a given $\boldsymbol{\pi} = (\pi_1,\ldots,\pi_m) \in S_N^m$, we can define the concatenated block encoder by the map composition

$$\phi_N^{\rm in}\circ\pi_m\circ\cdots\circ\phi_N^{\rm in}\circ\pi_1\circ\phi_N^{\rm out}. \tag{8}$$

Let $\mathfrak{S}_N^m$ be the set of all serial encoders (8) generated by varying the vector of permutations $\boldsymbol{\pi}$ over all possible elements in S_N^m. This type of ensemble lends itself to the average analysis introduced in [28] for turbo codes.

B. Weight Enumerators and Spectral Shapes for $\mathfrak{S}^m$

The average weight enumerators for $\mathfrak{S}_N^m$ are expressed in terms of the input-output weight distribution of its component codes (see [14], [29]). We have the following results.

Proposition 2:

$$\overline{\mathbf{A}}(\mathfrak{S}_N^m) = \mathbf{A}(\phi_N^{\rm out})\mathbf{P}(\phi_N^{\rm in})^m \tag{9}$$

where $\mathbf{P}(\phi_N^{\rm in})$ is a finite dimensional matrix given by

$$P_{w,d}(\phi_N^{\rm in}) \doteq \frac{A_{w,d}(\phi_N^{\rm in})}{\binom{N}{w}}.$$

$P_{w,d}(\phi_N^{\rm in})$ describes the probability that a randomly chosen input sequence of weight w is mapped by the inner encoder to an output sequence of weight d and for this reason is known as input–output weight transition probability. Notice that $\mathbf{P}(\phi_N^{\rm in})$ is thus a stochastic matrix.

We define now the Nth input-output weight distribution as

$$f_N(u,\delta;\phi^{\rm in}) \doteq \frac{1}{N}\ln P_{\lfloor uN\rfloor,\lfloor\delta N\rfloor}(\phi_N^{\rm in})$$

and the asymptotic input–output weight distribution as

$$f(u,\delta;\phi^{\rm in}) \doteq \limsup_{N\to\infty} f_N(u,\delta;\phi^{\rm in}).$$

From the fact that $P_{w,d}(\phi_N^{\rm in})$ represents a probability, it follows that the functions $f_N(u,\delta;\phi^{\rm in})$ and $f(u,\delta;\phi^{\rm in})$ are both not positive.

Whenever $\phi^{\rm in}$ is clear from the context, the input-output weight distributions will simply be denoted by $f_N(u,\delta)$ and $f(u,\delta)$, respectively.

Consider the sequence of ensembles $\overline{\mathfrak{S}}^m = \{\mathfrak{S}_N^m\}_{N\in R^{-1}\mathbb{N}\cap\mathbb{N}}$. It can be verified that the asymptotic spectral shapes satisfy the iterative relation:

$$\begin{aligned}\widehat{r}^{(m)}(\delta;\overline{\mathfrak{S}}) &= \max_{0\leq u\leq 1}\{\widehat{r}^{(m-1)}(u;\overline{\mathfrak{S}}) + f(u,\delta;\phi^{\rm in})\}\\ \widehat{r}^{(0)}(\delta;\overline{\mathfrak{S}}) &\doteq \limsup_{N\to\infty}\frac{1}{N}\ln A_{\lfloor\delta N\rfloor}(\phi_N^{\rm out}).\end{aligned} \tag{10}$$

C. Repeat Multiple–Accumulate Codes

Repeat Multiple–Accumulate codes, denoted by RA^m and introduced by Jin and McEliece in [1], are the simplest non trivial example of a serial concatenation of rate-1 codes through uniform random interleavers, where the outer encoder is the repeat code and inner encoders are truncated recursive convolutional accumulate codes.

Given $q \in \mathbb{N}$ and $N \in q\mathbb{N}$, the outer encoder $\mathrm{Rep}_N^q : \mathbb{Z}_2^{N/q} \to \mathbb{Z}_2^N$ repeats the information block q-times

$$\mathrm{Rep}_N^q([v_1,\dots,v_{N/q}]) = \underbrace{[v_1,\dots,v_{N/q},\dots,v_1,\dots,v_{N/q}]}_{q \text{ times}}$$

and the accumulator $\mathrm{Acc}_N : \mathbb{Z}_2^N \to \mathbb{Z}_2^N$ is the block encoder defined by

$$\mathrm{Acc}_N([u_1,\dots,u_N]) = [u_1, u_1+u_2,\dots,u_1+\cdots+u_N].$$

In this case, $\mathbf{A}(\mathrm{Rep}_N^q)$ and $\mathbf{P}(\mathrm{Acc}_N)$ can be explicitly computed (see [5]) and we obtain

$$A_{w,d}(\mathrm{Rep}_N^q) = \begin{cases} \binom{N/q}{w}, & qw = d \\ 0, & \text{otherwise} \end{cases}$$

$$P_{w,d}(\mathrm{Acc}_N) = \begin{cases} 1, & w = h = 0 \\ \frac{\binom{N-d}{\lfloor w/2 \rfloor}\binom{d-1}{\lceil w/2 \rceil - 1}}{\binom{N}{w}}, & w \geq 1 \text{ and } d \geq 1 \\ 0, & \text{otherwise.} \end{cases}$$

Consider the coefficients $P_{w,d}(\mathrm{Acc}_N)$ for $w \geq 1$ and $d \geq 1$ and notice that $P_{w,d}(\mathrm{Acc}_N)$ is nonzero if and only if

$$\lceil w/2 \rceil \leq d \quad \text{and} \quad \lfloor w/2 \rfloor \leq n - d$$

as one of the binomial coefficients in the numerator is zero if either condition is not satisfied.

Given $\delta \in [0,1]$, define the interval $\Omega_\delta = [0, 2\delta \wedge 2(1-\delta)]$. It can be verified [5] that the asymptotic spectral shapes satisfy the iterative relation in (10), with

$$f(u,\delta) \doteq f(u,\delta;\mathrm{Acc}) = \\ = \begin{cases} -H(u) + (1-\delta)H\left(\frac{u}{2(1-\delta)}\right) + \delta H\left(\frac{u}{2\delta}\right), \\ \qquad\qquad u \in \Omega_\delta \text{ and } \delta \in [0,1] \\ -\infty, \qquad\qquad \text{otherwise} \end{cases} \tag{11}$$

and

$$\widehat{r}^{(0)}(\delta) = H(\delta)/q. \tag{12}$$

IV. Previous Results

Kahale and Urbanke show [15] that for $m = 1$ the typical minimum distance of such coding schemes grows sublinearly in N with probability approaching one. Precisely, they prove the following result.

Theorem 1: Let ϕ^{in} be a recursive convolutional encoder. Then for every $\epsilon > 0$ we have

$$\lim_{N\to\infty} \mathbb{P}\left(d_{\min}(\mathfrak{S}_N^1) < N^{\frac{d_f^o - 2}{d_f^o} - \epsilon}\right) = 0$$

where d_f^o is the free distance of outer convolutional encoder.

Notice that the free distance d_f^o plays a crucial role in the estimation of the minimum distance: the more the parameter d_f^o is large, the more the minimum distance growth rate is close to be linear with high probability.

Multiple serial concatenations are studied in some depth in [18]. Indeed the authors, with arguments from the spectral theory of stochastic matrices applied to $\mathbf{P}(\phi_N^{\mathrm{in}})$, prove the following theorem.

Theorem 2 (Theorem 3 in [18]): For every $N \in R^{-1}\mathbb{N}$, it holds

$$\overline{A}_d(\mathfrak{S}_N^\infty) \doteq \lim_{m\to\infty} \overline{A}_d(\mathfrak{S}_N^m) = \begin{cases} 1, & \text{if } d = 0 \\ \binom{N}{d}\frac{2^{RN}-1}{2^N-1}, & \text{if } d \geq 1. \end{cases} \tag{13}$$

Notice that (2) and (13) are not identical. The main difference between them comes from the fact that all of the encoders in $\mathfrak{S}_N^m$ are invertible for all m, whereas the $\mathfrak{L}_N$ contains a small fraction of noninvertible encoders. However, both ensembles behave quite similarly: it can be verified that for any $\epsilon > 0$ there exists N_0 such that $\forall N \geq N_0$

$$|\overline{A}_d(\mathfrak{L}_N) - \overline{A}_d(\mathfrak{S}_N^\infty)| < \epsilon.$$

Theorem 2 yields the following result.

Corollary 1: There exists $\{m_N\}_{N\in R^{-1}\mathbb{N}\cap\mathbb{N}}$ such that for any $\epsilon > 0$

$$\mathbb{P}\left(d_{\min}(\mathfrak{S}_N^{m_N}) < (\delta_{\mathrm{GV}} - \epsilon)N\right) \xrightarrow{N\to\infty} 0.$$

Proof: Fix η such that $0 < \eta < 1/2^{1-R}$. It follows from Theorem 2 that, for every N and d, there exists $m_N(d)$ such that

$$\left|\overline{A}_d(\mathfrak{S}_N^m) - \overline{A}_d(\mathfrak{S}_N^\infty)\right| \leq \eta^N \\ \forall N \in R^{-1}\mathbb{N}\cap\mathbb{N},\ \forall d \geq 1,\ \forall m \geq m_N(d).$$

Let now $m_N \doteq \max\{m_N(d) : 1 \leq d \leq N\}$. Then

$$\left|\overline{A}_d(\mathfrak{S}_N^{m_N}) - \overline{A}_d(\mathfrak{S}_N^\infty)\right| \leq \eta^N \quad \forall N \in R^{-1}\mathbb{N}\cap\mathbb{N},\ \forall d \geq 1.$$

Equivalently

$$\frac{\ln\left(\overline{A}_d(\mathfrak{S}_N^\infty) - \eta^N\right)}{N} \leq \frac{\ln \overline{A}_d(\mathfrak{S}_N^{m_N})}{N} \leq \frac{\ln\left(\overline{A}_d(\mathfrak{S}_N^\infty) + \eta^N\right)}{N} \\ \forall N \in R^{-1}\mathbb{N}\cap\mathbb{N},\ \forall d \geq 1.$$

Denoting $r_N^{(m_N)}(\delta) = \frac{1}{N}\ln \overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^{m_N})$ we have that

$$r_N^{(m_N)}(\delta) \geq \frac{\ln \overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^\infty)}{N} + \frac{1}{N}\ln\left(1 - \frac{\eta^N}{\overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^\infty)}\right) \tag{14}$$

$$r_N^{(m_N)}(\delta) \leq \frac{\ln \overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^\infty)}{N} + \frac{1}{N}\ln\left(1 + \frac{\eta^N}{\overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^\infty)}\right). \tag{15}$$

From (37b) in Appendix A, we get the following estimations:

$$\overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^\infty) \geq \frac{\exp\{N[H(\delta) - (1-R)\ln 2]\}}{N+1} \tag{16}$$

$$\overline{A}_{\lfloor \delta N \rfloor}(\mathfrak{S}_N^\infty) \leq \exp\{N[H(\delta) - (1-R)\ln 2]\}. \tag{17}$$

From (14), (15), (16), and (17) we finally obtain

$$\begin{aligned} r_N^{(m_N)}(\delta) \geq & H(\delta) - (1-R)\ln 2 - \frac{1}{N}\ln(N+1) \\ & + \frac{1}{N}\ln\left[1 - \left(\frac{\eta}{\mathrm{e}^{H(\delta)-(1-R)\ln 2}}\right)^N (N+1)\right] \\ r_N^{(m_N)}(\delta) \leq & H(\delta) - (1-R)\ln 2 \\ & + \frac{1}{N}\ln\left[1 + \left(\frac{\eta}{\mathrm{e}^{H(\delta)-(1-R)\ln 2}}\right)^N (N+1)\right]. \end{aligned}$$

By the way η has been chosen, we have that $\eta/\mathrm{e}^{H(\delta)-(1-R)\ln 2} < 1$ for any $\delta \in [0,1]$. Therefore, we can conclude that

$$\lim_{N\to\infty} r_N^{(m_N)}(\delta) = H(\delta) - (1-R)\ln 2.$$

The assertion is now a straightforward application of (5). □

This result is very encouraging, as it puts into evidence that there exists a sequence of codes whose minimum distance converges to the GV bound. Notice that Theorem 2 holds for all the convolutional inner encoders (recursive or not recursive).

However, this argument does not give any information about the minimum distance distribution for the case of a finite number of inner encoders and it does not guarantee that the typical distance of all repeat-accumulate codes converges to the GV limit.

For this reason both analysis in [16] and [5] concern with the computation of the minimum distance distribution for the specific ensemble of RA^m. In particular, it is proved in [16] that if $m=2$ and the repetition factor is $q \geq 3$, then the typical minimum distance of RA^m grows linearly in the interleaver length N. An estimation of the linear growth rate is given in [5]. The result is summarized in the following theorem.

Theorem 3: For $q \geq 3$ and for any $\epsilon > 0$

$$\mathbb{P}\left(d_{\min}(RA_N^2) \leq (\overline{\delta} - \epsilon)N\right) \xrightarrow{N\to\infty} 0$$

where $\overline{\delta} = (4\mathrm{e}^{8/q})^{-1}$.

If we serially concatenate any encoder, whose minimum distance is growing like $\overline{\delta}N$, with an accumulate encoder through a uniform random interleaver, the minimum distance of the new encoder must grow faster than $\overline{\delta}N/2$ as $P_{h,d}(\mathrm{Acc}_N)$ is zero for every $d \leq \lceil h/2 \rceil$. Then Theorem 3 implies that if the minimum distance behaves linearly in N for $m=2$ then so must hold for every $m \geq 2$: for any $\epsilon > 0$

$$\mathbb{P}\left(d_{\min}(RA_N^m) \leq (\overline{\delta}/2^{m-2} - \epsilon)N\right) \xrightarrow{N\to\infty} 0.$$

Although this argument allows us to conclude that the typical minimum distance is growing linearly in N, it does not allow to prove that the minimum distance grows when m increases. However, simulation results in [18] show that the linear growth rate is increasing monotonically with m.

V. Summary of Our Results

The following theorems summarize our main results.

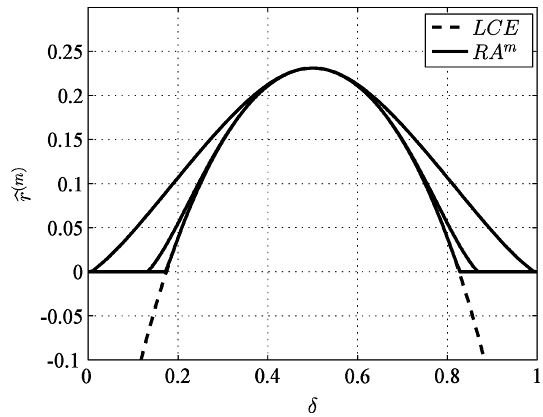

Fig. 2. Asymptotic spectral shapes for ensembles of RA^m with $m = 1,2,3$ (from top to bottom) and comparison to that of random linear coding ensemble (*LCE*).

Theorem 4: There exists a sequence of points $\{\epsilon_m\}_{m\in\mathbb{N}}$ (with $\epsilon_1 = 0$) strictly increasing in m such that

$$\begin{aligned} \widehat{r}^{(m)}(\delta) &= 0 \quad \forall\delta \in [0,\epsilon_m] \cup [1-\epsilon_m, 1] \\ \widehat{r}^{(m)}(\delta) &> 0 \quad \forall\delta \in (\epsilon_m, 1-\epsilon_m). \end{aligned}$$

Theorem 5: The sequence of asymptotic spectral shapes $\{\widehat{r}^{(m)}\}_{m\in\mathbb{N}} : [0,1] \to \mathbb{R}^+$ is monotonically nonincreasing in m and converges uniformly when $m \to \infty$ to

$$\widehat{r}^{(\infty)}(\delta) = \begin{cases} H(\delta) - (1-R)\ln 2, & \text{if } \delta \in (\delta_{\mathrm{GV}}, 1-\delta_{\mathrm{GV}}) \\ 0, & \text{otherwise} \end{cases}.$$

The spectral shapes are visualized in Fig. 2 with $q=3$ for $m = 1,2,3$ and compared to that of random linear coding ensemble.

We have to note that Theorem 4 corrects some wrong statements in [20], partially revised in [23]. Indeed, the authors in [20] overlook the fact that the maximizing value of

$$G^{(m)}(u,\delta) = \widehat{r}^{(m-1)}(u) + f(u,\delta)$$

with respect to the variable u can occur on the boundary. In fact, they only verify that the function at the local maximum is negative. Therefore they claim that the spectral shape is negative before some threshold ϵ_m and conclude that such point ϵ_m is the normalized minimum distance with high probability.

Actually, we will prove that $G^{(m)}(u,\delta)$ are not differentiable in general for the case $m > 2$ and that the floor of the spectral shapes $\widehat{r}^{(m)}(\delta)$ is zero. Therefore, we can not apply Proposition 1, in order to estimate the minimum distance distribution.

Nevertheless we shall prove the following theorem.

Theorem 6: We have that $\forall\epsilon > 0$

$$\lim_{N\to\infty} \mathbb{P}(d_{\min}(RA_N^m) \leq (\epsilon_m - \epsilon)N) = 0$$

for $m \geq 3$ and $q \geq 2$ or $m=2$ and $q \geq 3$.

We will see that this probability decreases to zero polinomially in the interleaver length N, in accordance with the previous results in [4] and [5].

TABLE I
NUMERICAL VALUES OF LINEAR GROWTH RATES ϵ_m FOR $m = 2, 3, 4$ AND COMPARISON TO THE NORMALIZED GILBERT-VARSHAMOV DISTANCE

R	ϵ_2	ϵ_3	ϵ_4	δ_{GV}
1/2	·	0.1033	0.1099	0.1100
1/3	0.1322	0.1731	0.1739	0.1740
1/4	0.1910	0.2143	0.2145	0.2145
1/5	0.2285	0.2429	0.2430	0.2430
1/6	0.2549	0.2643	0.2644	0.2644
1/7	0.2746	0.2812	0.2812	0.2812
1/8	0.2901	0.2949	0.2949	0.2949
1/9	0.3027	0.3063	0.3063	0.3063
1/10	0.3132	0.3160	0.3160	0.3160

These theorems guarantee together that the typical minimum distance of such coding schemes grows linearly in N with probability close to one. Moreover the minimum distance growth rate increases monotonically with m and converges to the limit implied by GV-bound when m tends to infinity. In Table I the normalized minimum distances ϵ_m are listed for $m = 2, 3, 4$ and compared to the GV-distance. These numerical results have been found using the approach in [31] to compute growth rates of the weight distributions of convolutional encoders. Notice that convergence looks quite fast: it is sufficient to put a small number of accumulate codes to get very close to the limit, i.e., to approach the normalized Gilbert–Varshamov distance.

Summarizing, Theorems 4 and 5 generalize those in [18], improve the earlier estimations of the growth rates in [16] and [5] for $m = 2$, and give a deeper insight into the problem of the aymptotic spectra of RA^m.

The monotonic increase with m of the minimum distance growth rate and the achievability of the Gilbert–Varshamov limit when m goes to infinity were conjectured but never analytically proved.

In Sections VI–VIII we shall prove, respectively, Theorem 4, 5, and 6 through intermediate steps.

VI. SPECTRAL SHAPE ANALYSIS

This section is devoted to the study of the asymptotic spectral shapes for a fixed number of accumulators m. The proofs of Propositions 3, 4, and 5 will be given in Appendix B.

Proposition 3: The following facts are true
1) $\widehat{r}^{(m)}(\delta) = \widehat{r}^{(m)}(1-\delta)$;
2) $\widehat{r}^{(m)}(\delta) : [0,1] \to \mathbb{R}$ is continuous;
3) $\widehat{r}^{(m)}(\delta) \geq 0,\ \forall \delta \in [0,1]$;
4) $\widehat{r}^{(m)}(\delta)$ is increasing in $\delta \in [0,1/2]$ and $\widehat{r}^{(m)}\left(\frac{1}{2}\right) = R\ln 2$.

Proposition 4: The sequence of functions $\{\widehat{r}^{(m)}(\delta)\}_{m\geq 1}$ is decreasing in m.

In the case $m = 1$ we can strengthen the properties 2) and 3) of Proposition 3.

Proposition 5: The following facts are true
1) $\widehat{r}^{(1)}(\delta) > 0,\ \forall \delta \in (0,1)$;
2) $\widehat{r}^{(1)}$ is differentiable.
3)
$$\frac{\mathrm{d}}{\mathrm{d}\delta}\widehat{r}^{(1)}(\delta)\big|_{\delta=0} \begin{cases} = 0, & \text{for } q \geq 3 \\ \leq 2, & \text{for } q = 2. \end{cases}$$

Define the sequence of points $\{\epsilon_m\}_{m\geq 1}$ such that
$$\epsilon_m = \max\{\epsilon \in [0,1/2) : \widehat{r}^{(m)}(\delta) = 0\ \forall\ \delta \leq \epsilon\}. \tag{18}$$
From the property 1) of Proposition 5 it follows that $\epsilon_1 = 0$. From Proposition 4 it is trivial to see that $\{\epsilon_m\}_{m\geq 1}$ is increasing in m. It can actually be shown that monotonicity is strict (see Appendix C for details).

Theorem 4 follows trivially from Proposition 3 and the monotonicity of $\{\epsilon_m\}_{m\geq 1}$.

The next results provide, respectively, a lower bound and an upper bound on the thresholds ϵ_m.

Proposition 6: If $\widehat{r}^{(m-1)}(\delta) \leq c\delta$ with $c \in \mathbb{R}$ then $\widehat{r}^{(m)}(\delta) = 0,\ \forall \delta \leq \frac{1}{2}(1 - \sqrt{1-\mathrm{e}^{-2c}})$.

Proof:
$$\begin{aligned}\widehat{r}^{(m)}(\delta) &= \max_{u\in\Omega_\delta}\{\widehat{r}^{(m-1)}(u) + f(u,\delta)\} \\ &\leq \max_{u\in\Omega_\delta}\{cu + f(u,\delta)\}.\end{aligned}$$
As $\frac{\partial^2}{\partial u^2} f(u,\delta) \leq 0,\ \forall\delta,\ u \in \Omega_\delta$ then for any fixed δ the maximizing value $\tilde{u}$ is unique:
$$\tilde{u}(\delta) = 1 - \frac{1-2\delta}{\sqrt{1-\mathrm{e}^{-2c}}} \in \Omega_\delta \iff \delta \leq \frac{1}{2}(1 - \sqrt{1-\mathrm{e}^{-2c}}).$$
It can be easily verified that
$$c\tilde{u}(\delta) + f\left(\tilde{u}(\delta),\delta\right) \leq 0 \qquad \forall \delta \leq \frac{1}{2}(1 - \sqrt{1-\mathrm{e}^{-2c}}).$$
The statement is proved, by using property 3) of Proposition 3. □

Corollary 2: $\widehat{r}^{(m)}(\delta) = 0\ \forall \delta \leq \frac{1}{2}(1-\sqrt{1-\mathrm{e}^{-4}}), m \geq 2.$

Proof: Consider the case with $m = 2$ and $q = 2$. From the inequality (43) and by the fact that $\{R_q(\delta)\}_{q\in\mathbb{N}}$ form a non-increasing sequence of functions in q, we have that
$$\widehat{r}^{(1)}(\delta) \leq R_2(\delta) = \frac{1}{2}\ln\left[1 + 4\delta(1-\delta)\right] \leq 2\delta \qquad \forall\delta.$$
From Proposition 6 we get that
$$\widehat{r}^{(2)}(\delta) = 0 \qquad \forall\delta \leq \frac{1}{2}(1 - \sqrt{1-\mathrm{e}^{-4}}).$$
The statement also holds for $m > 2$ from Proposition 4. □

Proposition 7:
$$\widehat{r}^{(m)}(\delta) \geq H(\delta) - (1-R)\ln 2 \qquad \forall m.$$

Proof: We prove it by induction on m. Consider the case with $m = 1$. From (10), (11), and (12) we have
$$\begin{aligned}\widehat{r}^{(1)}(\delta) &\geq q^{-1}H(2\delta(1-\delta)) + f(2\delta(1-\delta),\delta) \\ &= q^{-1}H(2\delta(1-\delta)) + H(\delta) - H(2\delta(1-\delta)) \\ &\geq -(1-q^{-1})\ln 2 + H(\delta).\end{aligned}$$
Suppose now that the inequality holds for m. We have
$$\widehat{r}^{(m+1)}(\delta) \geq \widehat{r}^{(m)}(2\delta(1-\delta)) + f(2\delta(1-\delta),\delta).$$

Using the inductive assumption on $\widehat{r}^{(m)}$ and again the fact that

$$f(2\delta(1-\delta),\delta) = H(\delta) - H(2\delta(1-\delta))$$

we prove that the inequality also holds for $m+1$. The proof is thus complete. □

Corollary 3: Let $\{\epsilon_m\}_{m\in\mathbb{N}}$ be the sequence defined in (18). We have that $\epsilon_m \leq \delta_{\mathrm{GV}}$, $\forall m$.

VII. Asymptotic Analysis

In the previous section we have studied the properties of the asymptotic spectral shape for a fixed number of accumulators m. We now analyze the behavior for $m \to +\infty$. The lower bound derived in Proposition 7 together with Proposition 4 guarantee that the sequence $\{r^{(m)}\}_{m\in\mathbb{N}}$ has a finite limit when m tends to infinity.

The recursive expression in (10) allows us to track the evolution of the spectral shape as it passes through each accumulate encoder. In this case the spectral shape in the new iteration can be expressed through a dynamical system. Through some techniques of non smooth analysis and the study of fixed points of the dynamical system, we will be able to study the convergence of the sequence of spectral shapes and to complete the proof of Theorem 5.

A. Dynamical System Formulation

We start by considering the operator

$$\begin{aligned}\Psi &: C([0,1]) \longrightarrow C([0,1])\\ \Psi[g](\delta) &= \max_{u\in\Omega_\delta}\{g(u)+f(u,\delta)\}, \quad \forall\delta\in[0,1]. \qquad (19)\end{aligned}$$

Given $\widehat{r}^{(0)}$ as initial condition, the sequence of asymptotic spectral shapes can be obtained recursively by

$$\widehat{r}^{(t+1)} = \Psi\left[\widehat{r}^{(t)}\right]. \qquad (20)$$

In order to describe the evolution (20), we study now some properties of Ψ.

We start with some simple properties.

Lemma 1: Let $g,h\in C([0,1])$, then

1) $\|\Psi[g]-\Psi[h]\|_\infty \leq \|g-h\|_\infty$.
2) If $g(\delta)\leq h(\delta)\ \forall\delta\in[0,1]$, then $\Psi[g](\delta)\leq\Psi[h](\delta)\ \forall\delta\in[0,1]$.
3) $\Psi[g+C] = C+\Psi[g]$, for any $C\in\mathbb{R}$.

Proof: 1): The result is an immediate consequence of the following fact

$$\begin{aligned}\Psi[g](\delta) &\leq \max_{u\in\Omega_\delta}[g(u)-h(u)] + \max_{u\in\Omega_\delta}[h(u)+f(u,\delta)] =\\ &= \max_{u\in\Omega_\delta}[g(u)-h(u)] + \Psi[h](\delta).\end{aligned}$$

2) and 3) are obvious. □

We say that $g\in C([0,1])$ is a *fixed point* for Ψ if $g=\Psi[g]$. It follows from (3) of Lemma 1 that, if g is a fixed point for Ψ, then the same holds for $g+C$. Another interesting way to modify fixed points is illustrated in the following result.

Proposition 8: If g is a fixed point for Ψ, then $g_+(x)=0 \vee g(x)$ is a fixed point for Ψ.

Proof: Consider the subset of maximizing points

$$\Gamma^+(\delta) = \operatorname*{argmax}_{u\in\Omega_\delta}[g_+(u)+f(u,\delta)].$$

For each $\delta\in[0,1]$ choose $u^+(\delta)\in\Gamma^+(\delta)$. We have

$$\begin{aligned}\Psi[g_+](\delta) &= g_+(u^+(\delta)) + f(u^+(\delta),\delta) =\\ &= f(u^+(\delta),\delta)\vee[g(u^+(\delta))+f(u^+(\delta),\delta)] \leq\\ &\leq f(u^+(\delta),\delta)\vee\left\{\max_{u\in\Omega_\delta}[g(u)+f(u,\delta)]\right\} =\\ &= f(u^+(\delta),\delta)\vee g(\delta). \qquad (21)\end{aligned}$$

Suppose now that $\delta\in[0,1]$ is such that $g(\delta)\leq 0$. Then, from (21) we have

$$0\leq\Psi[g_+](\delta)\leq f(u^+(\delta),\delta)\vee g(\delta)\leq 0.$$

We conclude that $\Psi[g_+](\delta)=0=g_+(\delta)$.

If instead δ is such that $g(\delta)>0$, we have

$$g(u)\leq g_+(u) \Longrightarrow g=\Psi[g]\leq\Psi[g_+].$$

As f is nonpositive, it follows that

$$g(\delta)\leq\Psi[g_+](\delta)\leq f(u^+(\delta),\delta)\vee g(\delta) = g(\delta)$$

and we conclude that $\Psi[g_+](\delta)=g(\delta)=g_+(\delta)$. This completes the proof. □

Proposition 9: The following functions are fixed points for Ψ, for any arbitrary constant C:

1) $g(\delta)=C$;
2) $g(\delta)=H(\delta)+C$.

Proof:

1) The result follows trivially by noticing that $g=0$ is a fixed point for Ψ as $f(u,\delta)$ is nonpositive. It then follows from property 3) of Lemma 1.
2) Consider

$$\begin{aligned}L(u,\delta) &= H(u)+f(u,\delta)\\ &= \delta H\left(\frac{u}{2\delta}\right) + (1-\delta)H\left(\frac{u}{2(1-\delta)}\right). \qquad (22)\end{aligned}$$

Since, for any fixed δ, $L(u,\delta)$ is concave in u, it is maximized at the only stationary point

$$u_{\max}(\delta) = 2\delta(1-\delta). \qquad (23)$$

It is straightforward to verify that $L(u_{\max}(\delta),\delta)=H(\delta)$. □

An important consequence of Propositions 8 and 9 is that both $H(\delta)-(1-R)\ln 2$ and $[H(\delta)-(1-R)\ln 2]_+$ are fixed points for Ψ.

The following is the key technical result of this section: proof will be given in Appendix D.

Lemma 2: There exists a constant $K\in\mathbb{R}$ such that

$$|\widehat{r}^{(m)}(\delta_2)-\widehat{r}^{(m)}(\delta_1)| \leq K|\delta_2-\delta_1| \qquad \forall\delta_1,\delta_2,\forall m.$$

Theorem 7: The sequence $\{\widehat{r}^{(m)}\}_{m\geq 1}$ converges uniformly to the limit $\widehat{r}^{(\infty)}$.

Proof: Since the sequence of functions $\{\widehat{r}^{(m)}\}_{m\geq 1}$ is decreasing in m and is lower bounded, it converges to the limit function $\widehat{r}^{(\infty)}$. Let

$$a_m = \max_{\delta\in[0,1]}[\widehat{r}^{(m)}(\delta) - \widehat{r}^{(\infty)}(\delta)].$$

Then, the sequence $\{a_m\}_{m\in\mathbb{N}}$ is monotonically decreasing in m and has a limit when $m \to \infty$.

From Proposition 3 and Lemma 2 the family $\{\widehat{r}^{(m)}\}_{m\geq 1}$ consists of uniformly bounded Lipschitz functions. Therefore Ascoli Arzelá's theorem (see [30]) guarantees that there exists a subsequence $\{m_j\}_{j\in\mathbb{N}}$ such that $a_{m_j} \to 0$ when $j \to \infty$. For the uniqueness of this limit we conclude that $a_m \to 0$. □

Corollary 4: $\widehat{r}^{(\infty)}(\delta)$ is a fixed point for Ψ.

Proof: It follows from Theorem 7, (20) and Lemma 1. □

B. Analysis of Limit Function $\widehat{r}^{(\infty)}(\delta)$

As we know the family $\{\widehat{r}^{(m)}\}_{m\geq 1}$ consists of a sequence of continuous and non negative functions converging uniformly to the limit function $\widehat{r}^{(\infty)}$. The next proposition characterizes some properties of it.

Proposition 10: The following facts are true

1) $\widehat{r}^{(\infty)}(\delta) = \widehat{r}^{(\infty)}(1-\delta)$.
2) $\widehat{r}^{(\infty)}(\delta) : [0,1] \to \mathbb{R}$ is continuous.
3) There exists $\epsilon_\infty > 0$ such that $\widehat{r}^{(\infty)}(\delta) = 0, \forall\delta \leq \epsilon_\infty$;
4) $\widehat{r}^{(\infty)}(\delta)$ is increasing in $\delta \in [0,1/2]$ and $\widehat{r}^{(\infty)}(1/2) = \frac{\ln 2}{q}$.

Proof: These are trivial consequences of Proposition 3 and, for the only case of continuity, also of Theorem 7. □

Notice that we already know a fixed point of Ψ satisfying all properties stated in Proposition 10: it is the function $[H(\delta)-(1-R)\ln 2]_+$. For the moment, from Proposition 7 and Theorem 7, we only know that, for any $\delta \in [0,1]$, $\widehat{r}^{(\infty)}(\delta) \geq [H(\delta) - (1-R)\ln 2]_+$. In the rest of this section, we will prove that they are in fact equal.

Consider g such that $\Psi[g] = g$ and such that it satisfies all properties listed in Proposition 10. Let

$$\Gamma_g(\delta) = \operatorname*{argmax}_{u\in\Omega_\delta}[g(u) + f(u,\delta)].$$

Then, for any $u \in \Gamma_g(\delta)$ it clearly holds

$$g(\delta) = \Psi[g] = g(u) + f(u,\delta). \tag{24}$$

We start with a technical result.

Lemma 3: The following facts are true.

1) For any $\delta \in [0,1/2)$, $\Gamma_g(\delta) \subseteq [0,1/2]$.
2) If $0 \leq \delta_1 < \delta_2 \leq 1/2$ and $u_i \in \Gamma_g(\delta_i)$ $(i = 1,2)$, then $u_1 \leq u_2$.
3) For any $\delta \in (\epsilon_\infty, 1/2)$ and $u \in \Gamma_g(\delta)$, we have $u \geq \delta$. Moreover, $\delta \in \Gamma_g(\delta)$ if and only if $\delta \in \{0,1/2\}$.
4) If $\delta_n \xrightarrow{n\to\infty} \delta_\infty$ and, $u_n \in \Gamma_g(\delta_n)$ is such that $u_n \xrightarrow{n\to\infty} u_\infty$, then $u_\infty \in \Gamma_g(\delta_\infty)$.

Proof:

1) From (47) we have

$$\frac{\partial}{\partial u}f(u,\delta) = \frac{1}{2}\ln\frac{4\delta(1-\delta)-2u+u^2}{(1-u)^2} < 0, \quad \forall\delta \neq 1/2, \forall u \in \Omega_\delta. \tag{25}$$

Hence for any fixed $\delta \neq 1/2$ $f(u,\delta)$ is monotonic decreasing for $u \in \Omega_\delta$. From this fact and property 1) of Proposition 10 follows immediately that

$$\begin{aligned} g(u) + f(u,\delta) &= g(1-u) + f(u,\delta) \\ &\geq g(1-u) + f(1-u,\delta) \quad \forall u\in[0,1/2]\cap\Omega_\delta. \end{aligned}$$

Then we conclude that

$$\Gamma_g(\delta) = \operatorname*{argmax}_{u\in\Omega_\delta}[g(u) + f(u,\delta)] \subseteq [0,1/2].$$

2) Since

$$\begin{aligned} \frac{\partial^2}{\partial\delta\partial u}f(u,\delta) &= \frac{\partial}{\partial\delta}\left\{\frac{1}{2}\ln\frac{4\delta(1-\delta)-2u+u^2}{(1-u)^2}\right\} \\ &= \frac{2(1-2\delta)}{4\delta(1-\delta)-2u+u^2} > 0 \quad \forall\delta<1/2,\ u \in \Omega_\delta \end{aligned}$$

it follows that if $0 \leq \delta_1 < \delta_2 \leq 1/2$ then

$$\frac{\partial}{\partial u}[f(u,\delta_2) - f(u,\delta_1)] = \int_{\delta_1}^{\delta_2} f_{\delta u}(u,\delta)\,\mathrm{d}\delta > 0. \tag{26}$$

We now prove the result by contradiction. Suppose that $\exists\, u_1 \in \Gamma_g(\delta_1)$ and $u_2 \in \Gamma_g(\delta_2)$ such that $u_2 < u_1$. From (24) we have

$$\begin{aligned} g(\delta_2) &= g(u_2) + f(u_2,\delta_2) \\ &= [g(u_2) + f(u_2,\delta_2)] + [f(u_2,\delta_1) - f(u_2,\delta_1)] \\ &= [g(u_2) + f(u_2,\delta_1)] + [f(u_2,\delta_2) - f(u_2,\delta_1)]. \end{aligned}$$

If $0 \leq \delta_1 < \delta_2 \leq 1/2$ then from (26) we get

$$f(u_2,\delta_2) - f(u_2,\delta_1) < f(u_1,\delta_2) - f(u_1,\delta_1).$$

from which it follows that

$$g(\delta_2) < [g(u_2) + f(u_2,\delta_1)] + [f(u_1,\delta_2) - f(u_1,\delta_1)].$$

Since $u_1 \in \Gamma_g(\delta_1)$ then

$$[g(u_2) + f(u_2,\delta_1)] \leq [g(u_1) + f(u_1,\delta_1)]$$

from which it follows that

$$\begin{aligned} g(\delta_2) &< g(u_1) + f(u_1,\delta_1) + f(u_1,\delta_2) - f(u_1,\delta_1) \\ &= g(u_1) + f(u_1,\delta_2). \end{aligned}$$

We conclude that $u_2 \notin \Gamma_g(\delta_2)$, which contradicts our assumption.

3) Notice first that if $\delta \in (\epsilon_\infty, 1/2)$, for sure $0 \notin \Gamma_g(\delta)$. Since $f(u,\delta) < 0$ for any $\delta \neq 1/2$ and $u \neq 0$, it follows from (24) that, necessarily, $g(\delta) < g(u)$. It now follows by property 4) of Proposition 10, that, $\delta < u$. Finally notice that (24) holds with $u = \delta$ if and only if $f(\delta,\delta) = 0$, and this happens if and only if $\delta \in \{0, 1/2\}$.
4) Let $\tilde{u} \in [0,1]$, then

$$g(\tilde{u}) + f(\tilde{u}, \delta_n) \leq g(u_n) + f(u_n, \delta_n).$$

By letting $n \to \infty$ and from the continuity of g and f we get

$$g(\tilde{u}) + f(\tilde{u}, \delta_\infty) \leq g(u_\infty) + f(u_\infty, \delta_\infty).$$

This yields the result. □

Theorem 8: Let g_1 and g_2 be fixed points of Ψ satisfying the properties listed in Proposition 10, then $g_1(\delta) = g_2(\delta), \forall \delta$.

Proof: Let

$$\overline{\epsilon}_i = \max\{\epsilon < 1/2 : g_i(\delta) = 0, \ \forall \delta \leq \epsilon\}.$$

For any $\delta \in [0,1]$, choose arbitrarily $\tilde{u}_i(\delta) \in \Gamma_{g_i}(\delta)$ with the only constraint that $\tilde{u}_1(1/2) = \tilde{u}_2(1/2) = 1/2$. For any δ, we can estimate as follows

$$\begin{aligned} g_2(\delta) &= g_2(\tilde{u}_2(\delta)) + f(\tilde{u}_2(\delta), \delta) + g_1(\tilde{u}_2(\delta)) - g_1(\tilde{u}_2(\delta)) \\ &\leq g_2(\tilde{u}_2(\delta)) + f(\tilde{u}_1(\delta), \delta) + g_1(\tilde{u}_1(\delta)) - g_1(\tilde{u}_2(\delta)) \\ &= g_2(\tilde{u}_2(\delta)) + g_1(\delta) - g_1(\tilde{u}_2(\delta)) \\ \Longrightarrow & g_2(\delta) - g_1(\delta) \leq g_2(\tilde{u}_2(\delta)) - g_1(\tilde{u}_2(\delta)). \end{aligned}$$

Repeating the argument k times we get

$$g_2(\delta) - g_1(\delta) \leq g_2(\tilde{u}_2^{(k)}(\delta)) - g_1(\tilde{u}_2^{(k)}(\delta)) \tag{27}$$

where $\tilde{u}_2^{(k+1)}(\delta) \in \Gamma_{g_2}\left(\tilde{u}_2^{(k)}(\delta)\right)$ with $k = 1, 2, \ldots$

In the same way, we get that

$$\begin{aligned} g_2(\delta) &\geq g_2(\tilde{u}_1(\delta)) + f(\tilde{u}_1(\delta), \delta) + g_1(\tilde{u}_1(\delta)) - g_1(\tilde{u}_1(\delta)) \geq \\ &\geq g_2(\tilde{u}_1(\delta)) + g_1(\delta) - g_1(\tilde{u}_1(\delta)). \end{aligned}$$

Iterating the argument k times, we have

$$g_2(\delta) - g_1(\delta) \geq g_2(\tilde{u}_1^{(k)}(\delta)) - g_1(\tilde{u}_1^{(k)}(\delta)) \tag{28}$$

where $\tilde{u}_1^{(k+1)}(\delta) \in \Gamma_{g_1}\left(\tilde{u}_1^{(k)}(\delta)\right)$ with $k = 1, 2, \ldots$

Fix now $\delta \in [\overline{\epsilon}_1 \vee \overline{\epsilon}_2, 1/2]$ and consider the recursive systems, for $i = 1, 2$:

$$\delta_i^{(k+1)} = \tilde{u}_i\left(\delta_i^{(k)}\right) \quad \delta_i^{(0)} = \delta. \tag{29}$$

By the way $\tilde{u}_i$ have been constructed and from items 1), 2), and 3) of Lemma 3 we know that both sequences $\{\delta_i^{(k)}\}_{k \in \mathbb{N}}$, $i = 1, 2$ are upper bounded by $1/2$ and increasing in k. Using 4) of Lemma 3, it follows that they both converge to $1/2$. From inequalities in (27) and (28), we have

$$\begin{aligned} g_2(\delta) - g_1(\delta) &\geq \lim_{k \to \infty} \left[g_2\left(\delta_1^{(k)}\right) - g_1\left(\delta_1^{(k)}\right) \right] \\ g_2(\delta) - g_1(\delta) &\leq \lim_{k \to \infty} \left[g_2\left(\delta_2^{(k)}\right) - g_1\left(\delta_2^{(k)}\right) \right]. \end{aligned}$$

As the functions g_2 and g_1 are both continuous

$$\begin{aligned} g_2(\delta) - g_1(\delta) &\geq g_2\left(\lim_{k\to\infty} \delta_1^{(k)}\right) - g_1\left(\lim_{k\to\infty} \delta_1^{(k)}\right) \\ g_2(\delta) - g_1(\delta) &\leq g_2\left(\lim_{k\to\infty} \delta_2^{(k)}\right) - g_1\left(\lim_{k\to\infty} \delta_2^{(k)}\right) \end{aligned}$$

then

$$0 = g_2\left(\frac{1}{2}\right) - g_1\left(\frac{1}{2}\right) \leq g_2(\delta) - g_1(\delta) \leq g_2\left(\frac{1}{2}\right) - g_1\left(\frac{1}{2}\right) = 0$$

and we conclude that $g_2(\delta) = g_1(\delta)$ for every $\delta \in [\overline{\epsilon}_1 \vee \overline{\epsilon}_2, 1/2]$.

Since the functions g_1 and g_2 are both symmetric with respect $\delta = 1/2$, continuous (see items 1) and 2) of Proposition 10)

$$g_1(\delta) = 0 \quad \forall \delta \leq \overline{\epsilon}_1, \qquad g_2(\delta) = 0 \quad \forall \delta \leq \overline{\epsilon}_2$$

and $g_1(\delta) = g_2(\delta), \forall \delta \in [\overline{\epsilon}_1 \vee \overline{\epsilon}_2, 1/2]$ then $\overline{\epsilon}_1 = \overline{\epsilon}_2$. □

Corollary 5:

$$r^{(\infty)}(\delta) = [H(\delta) - (1-R)\ln 2]_+.$$

VIII. Estimation of Minimum Distance Distribution

As we have already noticed, the floor of the spectral shapes is zero (see Fig. 2) and we cannot apply Proposition 1, in order to estimate the minimum distance distribution.

This means that $\forall \epsilon > 0$ the minimum distance is upper bounded by $d_{\min}(RA_N^m) \leq (\epsilon_m - \epsilon)N$ with a probability that does not decay exponentially in N.

We will prove that this probability decreases to zero polinomially in the interleaver length N. Inspired by asymptotic techniques devised in [19], we split the computation of the probability into two parts. The first part considers the contribution of the codewords with small weight in the last accumulate encoder $h < h_N$ and the second part refers to those codewords with weight $h \geq h_N$. The sequence $\{h_N\}_{N \in \mathbb{N}}$ can be chosen in such a way that the first term dominates the behavior of the overall probability.

Lemma 4: Let $\{h_N\}_{N \in \mathbb{N}}$ be a sequence of integers such that for all $\eta > 0$

$$\lim_{N \to \infty} \frac{h_N}{N^\eta} = 0. \tag{30}$$

Then

$$\sum_{h=1}^{h_N - 1} \overline{A}_h\left(RA_N^m\right) = O\left(N^{\beta_m + \eta}\right) \tag{31}$$

where $\beta_m = 1 - \sum_{i=1}^m \lceil q/2^i \rceil$.

The proof will be given in Appendix E.

Lemma 5:

$$r_N^{(m)}(\delta) \leq 2m \frac{\ln(N+1)}{N} + \widehat{r}^{(m)}(\delta).$$

Proof: We give the proof by induction on m. As an initial step, we take $m = 0$: by using the inequalities (37b) in Appendix A we get

$$A_d(\mathrm{Rep}_N^q) \leq \binom{N/q}{\lfloor d/q \rfloor} \leq \mathrm{e}^{N \frac{H(d/N)}{q}}$$

and the statement of this Lemma trivially holds [see (12)].

For the inductive step, assume that the statement of this lemma is true for $m-1$: from (37b) in Appendix A we have

$$\begin{aligned}
&\overline{A}_d(RA_N^m)\\
&=\sum_{h=1}^{N}\overline{A}_h(RA_N^{m-1})P_{h,d}(\mathrm{Acc}_N)\\
&\le N\max_{1\le h\le N}\left\{\overline{A}_h(RA_N^{m-1})P_{h,d}(\mathrm{Acc}_N)\right\}\\
&\le N\max_{\frac{h}{N}\in\left[\frac{1}{N},2\frac{d}{N}\wedge 2\left(1-\frac{d}{N}\right)\right]}\left\{\mathrm{e}^{N\left[\widehat{r}^{(m-1)}(h/N)+2(m-1)\frac{\ln(N+1)}{N}\right]}\right.\\
&\qquad\left.\times(N+1)\frac{\mathrm{e}^{N\left[(1-d/N)H\left(\frac{h/N}{2(1-d/N)}\right)+\frac{d}{N}H\left(\frac{h/N}{2d/N}\right)\right]}}{\mathrm{e}^{NH(h/N)}}\right\}\\
&\le(N+1)^2\exp\left\{N\left[\max_{\frac{h}{N}\in\left[\frac{1}{N},2\frac{d}{N}\wedge 2\left(1-\frac{d}{N}\right)\right]}\left[\widehat{r}^{(m-1)}(h/N)\right.\right.\right.\\
&\qquad\left.\left.\left.+f(h/N,d/N)\right]+2(m-1)\frac{\ln(N+1)}{N}\right]\right\}\\
&\le\exp\left\{N\left[\widehat{r}^{(m)}(d/N)+2m\frac{\ln(N+1)}{N}\right]\right\}
\end{aligned}$$

where the last equality follows from (10). Then statement is proved also for m. □

Theorem 9: Let $\{\epsilon_m\}_{m\ge 2}$ be the sequence defined in (18). We have that $\forall\epsilon>0$

$$\lim_{N\to\infty}\mathbb{P}(d_{\min}(RA_N^m)\le(\epsilon_m-\epsilon)N)=0.$$

Proof: Fix $\epsilon>0$ and let $d_N=(\epsilon_m-\epsilon)N$. Pick a sequence of integers $\{h_N\}_{N\in\mathbb{N}}$ satisfying (30) and such that

$$\lim_{N\to\infty}\frac{\ln N}{h_N}=0. \tag{32}$$

From (5) we have

$$\begin{aligned}
&\mathbb{P}(d_{\min}(RA_N^m)\le d_N)\\
&\le\sum_{d=1}^{d_N}\sum_{h=1}^{2d}\overline{A}_h\left(RA_N^{m-1}\right)P_{h,d}(\mathrm{Acc}_N)\\
&\le\sum_{h=1}^{2d_N}\sum_{d=1}^{d_N}\overline{A}_h\left(RA_N^{m-1}\right)P_{h,d}(\mathrm{Acc}_N)\\
&=\sum_{h=1}^{h_N-1}\overline{A}_h\left(RA_N^{m-1}\right)\sum_{d=1}^{d_N}P_{h,d}(\mathrm{Acc}_N)+\\
&\quad+\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\overline{A}_h\left(RA_N^{m-1}\right)P_{h,d}(\mathrm{Acc}_N)\\
&\le\sum_{h=1}^{h_N-1}\overline{A}_h(RA_N^{m-1})+\\
&\quad+\underbrace{\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\overline{A}_h(RA_N^{m-1})P_{h,d}(\mathrm{Acc}_N)}_{S^m(N)}.
\end{aligned}\tag{33}$$

Let $G^{(m)}(x,y)=\widehat{r}^{(m-1)}(x)+f(x,y)$.

From Lemma 5 and (37b) in Appendix A we can estimate as follows:

$$\begin{aligned}
&S^m(N)\\
&=\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\overline{A}_h(RA_N^{m-1})P_{h,d}(\mathrm{Acc}_N)\\
&\le\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\mathrm{e}^{Nr_N^{(m-1)}(h/N)}(N+1)\\
&\quad\times\frac{\mathrm{e}^{N\left[\left(1-\frac{d}{N}\right)H\left(\frac{h/N}{2(1-d/N)}\right)+\frac{d}{N}H\left(\frac{h/N}{2d/N}\right)\right]}}{\mathrm{e}^{NH(h/N)}}\\
&=(N+1)\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\mathrm{e}^{N\left[r_N^{(m-1)}(h/N)+f(h/N,d/N)\right]}\\
&\le(N+1)\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\mathrm{e}^{N\left[\widehat{r}^{(m-1)}(h/N)+2(m-1)\frac{\ln(N+1)}{N}+f(h/N,d/N)\right]}\\
&=(N+1)^{2m-1}\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\mathrm{e}^{N\left[\widehat{r}^{(m-1)}(h/N)+f(h/N,d/N)\right]}\\
&=(N+1)^{2m-1}\sum_{h=h_N}^{2d_N}\sum_{d=1}^{d_N}\mathrm{e}^{NG^{(m)}(h/N,d/N)}.
\end{aligned}$$

Using the fact that for fixed u the function $f(u,\delta)$ is an increasing function in $u/2<\delta<1/2$, then we get

$$\begin{aligned}
S^m(N)&\le d_N(N+1)^{2m-1}\sum_{h=h_N}^{2d_N}\mathrm{e}^{N\max_{y\in[1/N,d_N/N]}G^{(m)}(h/N,y)}\\
&=d_N(N+1)^{2m-1}\sum_{h=h_N}^{2d_N}\mathrm{e}^{NG^{(m)}(h/N,\epsilon_m-\epsilon)}\\
&\le(\epsilon_m-\epsilon)(N+1)^{2m}\sum_{h=h_N}^{2d_N}\mathrm{e}^{NG^{(m)}(h/N,\epsilon_m-\epsilon)}.
\end{aligned}$$

Moreover, for $h_N\le h\le 2d_N$

$$NG^{(m)}(h/N,\epsilon_m-\epsilon)\le h\tau^{(m)}$$

where

$$\tau^{(m)}\doteq\max_{h_N/N<u\le 2(\epsilon_m-\epsilon)}\frac{\widehat{r}^{(m-1)}(u)+f(u,\epsilon_m-\epsilon)}{u}.$$

Being $\widehat{r}^{(m-1)}(u)=0\ \forall u\le\epsilon_{m-1}$, we can compute the optimizing value, by splitting the computation of $\tau^{(m)}$ as follows:

$$\tau^{(m)}=\max\left(L_1^{(m)},L_2^{(m)}\right)$$

where

$$\begin{aligned}
L_1^{(m)}&\doteq\max_{u\in[h_N/N,\epsilon_{m-1}]}\frac{f(u,\epsilon_m-\epsilon)}{u}\\
L_2^{(m)}&\doteq\max_{u\in(\epsilon_{m-1},2(\epsilon_m-\epsilon)]}\frac{\widehat{r}^{(m-1)}(u)+f(u,\epsilon_m-\epsilon)}{u}.
\end{aligned}$$

Fixed $\delta>0$, from (40) and (25) we have that

$$\frac{\partial}{\partial u}\left(\frac{f(u,\delta)}{u}\right)=\frac{1}{u^2}\left[u\frac{\partial}{\partial u}\Big(f(u,\delta)\Big)-f(u,\delta)\right]$$

$$
\begin{aligned}
&= \frac{1}{u^2}\left[\frac{u}{2}\ln\left(\frac{4\delta(1-\delta)-2u+u^2}{(1-u)^2}\right)\right.\\
&\quad - u\ln\left(2\sqrt{\delta(1-\delta)}\right) - (1-u)\ln(1-u)\\
&\quad + \frac{2\delta-u}{2}\ln\left(\frac{2\delta-u}{2\delta}\right)\\
&\quad \left.+\frac{2-2\delta-u}{2}\ln\left(\frac{2-2\delta-u}{2-2\delta}\right)\right]\\
&= \frac{1}{u^2}\left[\frac{u}{2}\ln\frac{(2\delta-u)(2-2\delta-u)}{(1-u)^2}\right.\\
&\quad -\frac{u}{2}\ln(4\delta(1-\delta))-\ln(1-u)+u\ln(1-u)\\
&\quad + \frac{2\delta-u}{2}\ln(2\delta-u) - \frac{2\delta-u}{2}\ln(2\delta)\\
&\quad + \frac{2-2\delta-u}{2}\ln(2-2\delta-u)\\
&\quad \left.-\frac{2-2\delta-u}{2}\ln(2-2\delta)\right]\\
&= \frac{1}{u^2}\left[-\ln(1-u)+\delta\ln\frac{(2\delta-u)}{2\delta}\right.\\
&\quad \left.+(1-\delta)\ln\frac{(2(1-\delta)-u)}{2(1-\delta)}\right].
\end{aligned}
$$

The function $-\ln(u)$ is convex, hence, Jensen's inequality gives

$$
\begin{aligned}
-\ln(1-u) &= -\ln\left(\frac{2-2\delta-u}{2}+\frac{2\delta-u}{2}\right)\\
&= -\ln\left((1-\delta)\frac{2-2\delta-u}{2(1-\delta)}+\delta\frac{2\delta-u}{2\delta}\right)\\
&\leq -(1-\delta)\ln\left(\frac{2-2\delta-u}{2(1-\delta)}\right)-\delta\ln\left(\frac{2\delta-u}{2\delta}\right)
\end{aligned}
$$

from which it follows

$$
\begin{aligned}
\frac{\partial}{\partial u}\left(\frac{f(u,\delta)}{u}\right) &= \frac{1}{u^2}\left[-\ln(1-u)+\delta\ln\frac{(2\delta-u)}{2\delta}\right.\\
&\quad \left.+(1-\delta)\ln\frac{(2(1-\delta)-u)}{2(1-\delta)}\right] \leq 0 \quad \forall u\in\Omega_\delta.
\end{aligned}
$$

Therefore, $f(u,\epsilon_m-\epsilon)/u$ is monotonic decreasing and we conclude that

$$
\begin{aligned}
L_1^{(m)} &= \frac{f(h_N/N,\epsilon_m-\epsilon)}{h_N/N} \leq \lim_{N\to\infty}\frac{f(h_N/N,\epsilon_m-\epsilon)}{h_N/N} =\\
&= \frac{\partial}{\partial u}f(u,\epsilon_m-\epsilon)\Big|_{u=0} < 0.
\end{aligned} \tag{34}
$$

Using the fact that $\widehat{r}^{(m-1)}$ and f are both continuous and $f(u,\delta)$ is strictly increasing in $\delta<1/2$, by the way ϵ_m has been defined (18), we obtain that

$$
\widehat{r}^{(m-1)}(u)+f(u,\epsilon_m-\epsilon)<0 \quad \text{for } u\in[\epsilon_{m-1},2(\epsilon_m-\epsilon)].
$$

Hence

$$
L_2^{(m)} = \max_{u\in(\epsilon_{m-1},2(\epsilon_m-\epsilon)]}\frac{\widehat{r}^{(m-1)}(u)+f(u,\epsilon_m-\epsilon)}{u} < 0. \tag{35}
$$

From (34) and (35) we get $\tau^{(m)} = L_1^{(m)} \vee L_2^{(m)} < 0$ and

$$
\begin{aligned}
S^m(N) &\leq (\epsilon_m-\epsilon)(N+1)^{2m}\sum_{h=h_N}^{2d_N} \mathrm{e}^{\tau^{(m)}h} \leq\\
&\leq (\epsilon_m-\epsilon)(N+1)^{2m}\frac{\mathrm{e}^{h_N\tau^{(m)}}}{1-e^{\tau^{(m)}}}.
\end{aligned}
$$

It follows from (32) that

$$
\begin{aligned}
S^m(N) &\leq C\exp\left[2m\ln(N+1)+\tau^{(m)}h_N\right]\\
&= C\exp\left[h_N\left(2m\frac{\ln(N+1)}{h_N}+\tau^{(m)}\right)\right] \xrightarrow{N\to\infty} 0
\end{aligned}
$$

with $C=\frac{\epsilon_m-\epsilon}{1-e^{\tau}}$.

From (33) and Lemma 4, we conclude that $\forall\eta>0$

$$
\mathbb{P}(d_{\min}(RA_N^m)\leq(\epsilon_m-\epsilon)N) = O\left(N^{\beta_{m-1}+\eta}\right), \tag{36}
$$

from which, if we choose η sufficiently small, it follows that

$$
\lim_{N\to\infty}\mathbb{P}(d_{\min}(RA_N^m)\leq(\epsilon_m-\epsilon)N) = 0
$$

for $m\geq 3$ and $q\geq 2$ or $m=2$ and $q\geq 3$. □

IX. Beyond Repeat Multiple–Accumulate Codes

The analysis in the previous sections has been focused on a restricted class of turbo codes. However, we expect many of the properties to hold for more general ensembles of codes. In this section we introduce in the encoding scheme some simple variations, which we hope will produce better codes, in the sense of the minimum distance distribution. We discuss here some numerical examples: first of all, we replace the outer encoder (the repetition encoder) with a convolutional encoder; then we consider the class of codes of Repeat Multiple-Convolute codes, obtained by choosing a general convolutional encoder of rate-1 in place of the accumulator.

A. Convolutional Multiple–Accumulate Codes

Convolutional Multiple–Accumulate codes (CA^m) are an extension of the RA^m codes. They consist of a serial concatenation of a truncated convolutional code ϕ_N^{out} with m interleaved rate-1 accumulate codes.

Replacing the repetition code with any other encoder we get that the sequence of asymptotic spectral shapes are defined recursively by the dynamical system Ψ in (19) with $f(u,\delta)$ given by (11) with generic initial condition

$$
\widehat{r}^{(0)}(\delta) = \limsup_{N\to\infty}\frac{1}{N}\ln A_{\lfloor\delta N\rfloor}(\phi_N^{\text{in}}).
$$

Based on the computation of asymptotic growth rates of the weight distributions of the convolutional encoders in [31] [Examples 6 and 7], the asymptotic spectral shapes of the CA^m with the outer convolutional encoder $G_1(D)=[1,1+D]$ and $G_2(D)=[1+D,1+D+D^2]$ are plotted for $m=1,2,3$,

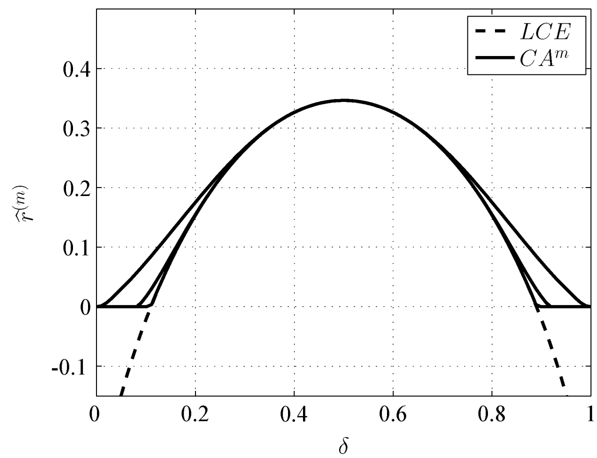

Fig. 3. Asymptotic spectral shapes corresponding to the ensembles of CA^m with $m = 1, 2, 3$ (from top to bottom) and outer convolutional encoder $G_1(D) = [1, 1 + D]$ (thick lines) and comparison to that of linear coding ensemble (dashed line).

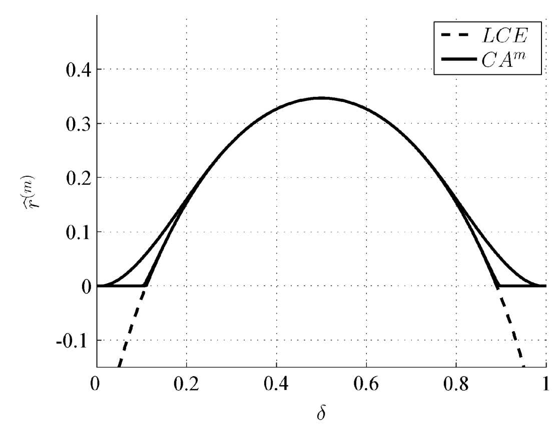

Fig. 4. Asymptotic spectral shapes corresponding to the ensembles of CA^m with $m = 1, 2, 3$ (from top to bottom) and outer convolutional encoder $G_2(D) = [1 + D^2, 1 + D + D^2]$ (thick lines) and comparison to that of linear coding ensemble (dashed line).

TABLE II
NUMERICAL VALUES OF LINEAR GROWTH RATES ϵ_m FOR CA^m OF RATE 1/2 FOR $m = 2, 3, 4$

Ensemble	ϵ_2	ϵ_3	ϵ_4	δ_{GV}
RA^m	·	0.1033	0.1099	0.1100
$C_{G_1}A^m$	0.0831	0.1092	0.1100	0.1100
$C_{G_2}A^m$	0.1044	0.1100	0.1100	0.1100

and compared with that of the random ensemble of binary linear block codes (dashed line) in Figs. 3 and 4.

In Table II linear growth rate in the code length N of the minimum distance are evaluated numerically and compared.

These numerical results are very encouraging and suggest that Theorem 4, 5, and 6 are likely also to hold for ensembles of CA^m codes.

B. Repeat Multiple–Convolute Codes

Repeat Multiple–Convolute codes are obtained from RA^m codes by replacing the inner accumulate encoders with m convolutional encoders.

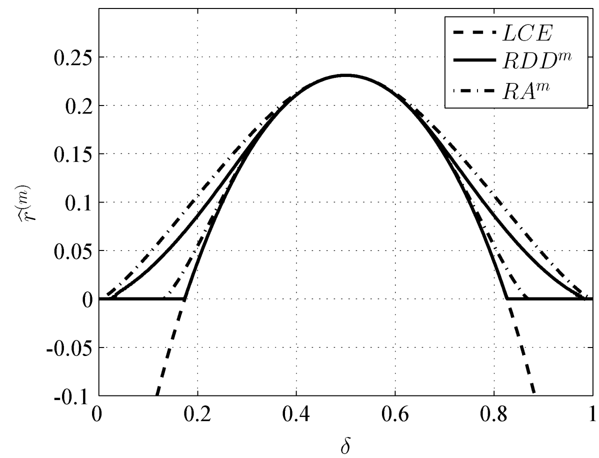

Fig. 5. Asymptotic spectral shapes of RDD^m with $m = 1, 2$ (thick lines from top to bottom) and comparison to those of RA^m (dashed-dot lines) and linear coding ensemble (dashed line).

The asymptotic spectral shapes for Repeat Multiple–Delay codes (RDD^m), introduced in [2] and obtained by interconnecting the repeat code with m recursive convolutional encoders with transfer function $G_i(D) = (1 + D + D^2)^{-1}$, are plotted in Fig. 5 and compared with that of binary linear random code ensemble (dashed curve) with $m = 1, 2, 3$.

The dynamical system defined in (19) depends exclusively on the accumulator, and it would be different with another convolutional encoder. The numerical results and the results in [18] lead us to conjecture that the dynamical systems analysis is likely to hold true for more general convolutional encoders. Instead in order to get linear growth of minimum distances it is clear that the inner encoder must necessarily be recursive.

X. CONCLUDING REMARKS

In this paper, we have studied some properties of the spectral shapes for uniformly interleaved repeat multiple–accumulate codes and their relation to the minimum distance growth rate. In particular, we have shown that for $m \geq 2$, the asymptotic spectral shapes exhibit some different features as compared to the case where $m = 1$. The main difference is that for $m \geq 2$ there exists a positive point ϵ_m such that the function is zero below it and positive beyond it. Moreover, by tracking the evolution of the asymptotic spectral shapes, we have shown that these functions converge uniformly, when m tends to infinity, to that of random linear code ensemble for $\delta \in [\delta_{\mathrm{GV}}, 1 - \delta_{\mathrm{GV}}]$.

Although the floor of the spectral shapes is zero, by combining the asymptotic spectral shapes with the use of the union bound we have proved that the ensemble of Repeat-Accumulate-m codes is asymptotically good, in the sense that the typical minimum distance grows linearly in N with probability one. Moreover we have provided a numerical method to estimate with arbitrarily small accuracy the linear growth rate: except for the case of $q = 2$ and $m = 2$, we have proved that the growth rate is at least ϵ_m. This implies that the normalized minimum distance increases monotonically with m and meets the limit implied by the Gilbert–Varshamov bound on the minimum distance when m tends to infinity. Notice that the minimum distance ratios computed using this method are quite close to the empirical growth rates listed in [18].

Although our results are obtained for a restricted class of turbo codes, we believe that our techniques can be applied to a much wider class of interleaved concatenated codes. Theoretically, our mathematical tools provide a new general framework for estimating the minimum distance distribution of multiple serially concatenated codes.

This paper leaves some open problems to study:

- How fast the sequence of the spectral shapes converges to the limit function?
- What is the effect of the inner encoders of the encoding scheme on this convergence?

The results presented in this paper are very encouraging and suggest that even a few accumulators are sufficient to approach the asymptotic behavior.

Moreover the dynamical system we have defined depends exclusively on the accumulator and it would be different if we replace it with another convolutional encoder. The numerical results and the fact that Theorem 6 and Corollary 7 hold for any choice of the convolutional encoder (as long it was not the identity) lead us to conjecture that the dynamical systems analysis is likely to hold true for all convolutional encoders both recursive and not recursive.

Instead for the final part on the estimation of distances, the role of recursivity must necessarily come up since, if it is not recursive can not certainly exhibit linear growth of minimum distances. Indeed, it is easy to verify that for any nonrecursive rate-1 convolutional encoder with an impulse response of weight d the output weight will be at most d times the input weight. If the desired output weight is γN and the input weight is 1, then the minimum numbers of concatenations needed is $\log_d \gamma N$. We conclude that, for fixed m and asymptotically large N, the convolutional encoder never maps an input word of weight 1 to an output word of weight γN and we expect that the ensemble has low weight codewords.

The main difficulty in the extension of the proofs is the computation of the asymptotic spactral shapes. Indeed a basic request is to determine the average spectrum of the given ensemble. Although there are analytic approaches to determine the weight distribution exactly for relatively small lengths (see [7]), none of them allow a direct computation of the asymptotic growth rate. Sason *et al.* in [31] present a method for the determination of the asymptotic input-output weight distribution of convolutional encoders, but only in some cases this method is able to derive analytic expressions. In general cases, it requires to solve numerically a system of polynomial equations.

Appendix A
Some Useful Inequalities

Proposition 11: For any integers $1 \leq k \leq n$ the following inequalities hold true:

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{n\mathrm{e}}{k}\right)^k \tag{37a}$$

$$\frac{\mathrm{e}^{nH(k/n)}}{n+1} \leq \binom{n}{k} \leq \mathrm{e}^{nH(k/n)} \tag{37b}$$

$$\frac{\binom{n}{i}\binom{m}{j}}{\binom{n+m}{i+j}} \leq \left(\frac{n}{n+m}\right)^i \left(\frac{m}{n+m}\right)^j \left(\frac{i+j}{i}\right)^i \left(\frac{i+j}{j}\right)^j. \tag{37c}$$

The proof is given in [5, Appendices 3A and 3C].

Appendix B
Proofs of Proposition 3, 4, and 5

Proof of Proposition 3:

1) From (12) we have that $\widehat{r}^{(0)}(\delta) = \widehat{r}^{(0)}(1-\delta)$. Consider now the case $m \geq 1$. From (10) we get that

$$\begin{aligned}\widehat{r}^{(m)}(1-\delta) &= \max_{u\in[0,1]}\{\widehat{r}^{(m-1)}(u) + f(u,1-\delta)\}\\ &= \max_{u\in[0,1]}\{\widehat{r}^{(m-1)}(u) + f(u,\delta)\} = \widehat{r}^{(m)}(\delta)\end{aligned}$$

where the second equality follows from the fact that $f(u,\delta) = f(u,1-\delta), \forall u \in [0,1]$ [see (11)].

2) The asymptotic spectral shape $\widehat{r}^{(0)}$ in (12) is continuous as the entropy function is continuous. Considering (10) and the continuity of f defined in (11), the general case can be proved by induction on m.

3) From (10), (11), and (12) we have

$$\begin{aligned}\widehat{r}^{(1)}(\delta) &= \max_{u\in[0,1]}\left\{\frac{H(u)}{q} + f(u,\delta)\right\}\\ &\geq f(0,\delta) + H(0)/q = 0.\end{aligned}$$

Then proceed again by induction on m.

4) From (12) we have that $\widehat{r}^{(0)}(\delta)$ is strictly increasing in $\delta \in [0,1/2]$. Since

$$\begin{aligned}\frac{\partial}{\partial\delta}f(u,\delta) &= H\left(\frac{u}{2\delta}\right) - H\left(\frac{u}{2(1-\delta)}\right)\\ &\quad - \frac{u}{2\delta}\ln\frac{1-\frac{u}{2\delta}}{\frac{u}{2\delta}} + \frac{u}{2(1-\delta)}\ln\left(\frac{1-\frac{u}{2(1-\delta)}}{\frac{u}{2(1-\delta)}}\right)\\ &= \ln\left(1-\frac{u}{2(1-\delta)}\right) - \ln\left(1-\frac{u}{2\delta}\right) \geq 0,\\ &\qquad \forall\delta \in [0,1/2],\ u \in \Omega_\delta\end{aligned} \tag{38}$$

if $0 \leq \delta_1 \leq \delta_2 \leq 1/2$ then we have that

$$\begin{aligned}\widehat{r}^{(m)}(\delta_1) &= \max_{u\in[0,1]}\left\{\widehat{r}^{(m-1)}(u) + f(u,\delta_1)\right\}\\ &= \max_{u\in[0,2\delta_1]}\left\{\widehat{r}^{(m-1)}(u) + f(u,\delta_1)\right\}\\ &\leq \max_{u\in[0,2\delta_1]}\left\{\widehat{r}^{(m-1)}(u) + f(u,\delta_2)\right\}\\ &\leq \max_{u\in[0,2\delta_2]}\left\{\widehat{r}^{(m-1)}(u) + f(u,\delta_2)\right\} = \widehat{r}^{(m)}(\delta_2).\end{aligned}$$

Moreover, from (12) we have that $\widehat{r}^{(0)}(1/2) = R\ln 2$. The general case can be proved by induction on m, using the fact that $f(u,1/2) = 0$.

Proof of Proposition 4: We prove the assertion by induction on m.

Let us consider first the case $m = 1$. We prove that

$$\widehat{r}^{(1)}(\delta) \leq \widehat{r}^{(0)}(\delta) \quad \forall\delta \in [0,1] \tag{39}$$

and the equality holds if and only if $\delta = 0, 1/2, 1$.

From the expression in (11) we have

$$\begin{aligned}
&f(u,\delta)\\
&= -H(u)+(1-\delta)H\left(\frac{u}{2(1-\delta)}\right)+\delta H\left(\frac{u}{2\delta}\right)\\
&= u\ln u+(1-u)\ln(1-u)\\
&\quad+(1-\delta)\left[-\frac{u}{2(1-\delta)}\ln\left(\frac{u}{2(1-\delta)}\right)\right.\\
&\quad\left.-\left(1-\frac{u}{2(1-\delta)}\right)\ln\left(1-\frac{u}{2(1-\delta)}\right)\right]\\
&\quad+\delta\left[-\frac{u}{2\delta}\ln\left(\frac{u}{2\delta}\right)-\left(1-\frac{u}{2\delta}\right)\ln\left(1-\frac{u}{2\delta}\right)\right]\\
&= u\ln u+(1-u)\ln(1-u)-\frac{u}{2}\ln\left(\frac{u}{2(1-\delta)}\right)\\
&\quad-\frac{2(1-\delta)-u}{2}\ln\left(\frac{2(1-\delta)-u}{2(1-\delta)}\right)\\
&\quad-\frac{u}{2}\ln\left(\frac{u}{2\delta}\right)-\frac{2\delta-u}{2}\ln\left(\frac{2\delta-u}{2\delta}\right)\\
&= u\ln u+(1-u)\ln(1-u)-\frac{u}{2}\ln u+\frac{u}{2}\ln\left(2(1-\delta)\right)\\
&\quad-\frac{2-2\delta-u}{2}\ln\left(\frac{2-2\delta-u}{2(1-\delta)}\right)-\frac{u}{2}\ln u+\frac{u}{2}\ln(2\delta)\\
&\quad-\frac{2\delta-u}{2}\ln\left(\frac{2\delta-u}{2\delta}\right)\\
&= u\ln\left(2\sqrt{\delta(1-\delta)}\right)+(1-u)\ln(1-u)\\
&\quad-\frac{2\delta-u}{2}\ln\left(\frac{2\delta-u}{2\delta}\right)-\frac{2-2\delta-u}{2}\ln\left(\frac{2-2\delta-u}{2-2\delta}\right).
\end{aligned}\tag{40}$$

Jensen's inequality and the fact that $g(u)=u\ln u$ is strictly convex imply that

$$\begin{aligned}
(1-u)\ln(1-u) &= g(1-u)\\
&= g\left(\frac{2\delta-u}{2\delta}\delta+\frac{2-2\delta-u}{2(1-\delta)}(1-\delta)\right)\\
&\le \delta g\left(\frac{2\delta-u}{2\delta}\right)+(1-\delta)g\left(\frac{2-2\delta-u}{2(1-\delta)}\right)\\
&= \frac{2\delta-u}{2}\ln\left(\frac{2\delta-u}{2\delta}\right)\\
&\quad+\frac{2-2\delta-u}{2}\ln\left(\frac{2-2\delta-u}{2(1-\delta)}\right)
\end{aligned}$$

from which it follows that

$$f(u,\delta)\le u\ln\left(2\sqrt{\delta(1-\delta)}\right)\tag{41}$$

and equality holds if and only if $u=0$ or $\delta=1/2$.

By (41) we get that

$$\begin{aligned}
\widehat{r}^{(1)}(\delta) &= \max_{0\le u\le 1}\{\widehat{r}^{(0)}(u)+f(u,\delta)\}\\
&= \max_{0\le u\le 1}\left\{\frac{H(u)}{q}+f(u,\delta)\right\}\\
&\le \max_{0\le u\le 1}\left\{\frac{H(u)}{q}+u\ln\left(2\sqrt{\delta(1-\delta)}\right)\right\}\\
&= \max_{0\le u\le 1}\left\{\frac{H(u)}{q}+\frac{u}{q}\ln\left(2\sqrt{\delta(1-\delta)}\right)^q\right\}.
\end{aligned}\tag{42}$$

Differentiating this expression with respect to the variable u

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}u}\left\{\frac{H(u)}{q}+\frac{u}{q}\ln\left(2\sqrt{\delta(1-\delta)}\right)^q\right\}\\
&\qquad=\frac{1}{q}\ln\left(\frac{1-u}{u}\right)+\frac{1}{q}\ln\left(2\sqrt{\delta(1-\delta)}\right)^q=0
\end{aligned}$$

we get

$$\frac{1-u}{u}=\frac{1}{\left(2\sqrt{\delta(1-\delta)}\right)^q}$$

from which the optimizing value in the computation is

$$u=\frac{1}{1+\frac{1}{\left(2\sqrt{\delta(1-\delta)}\right)^q}}=\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}.$$

Substituting it in the right-hand side of (42) gives

$$\begin{aligned}
\widehat{r}^{(1)}(\delta) &\le \frac{1}{q}H\left(\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\right)\\
&\quad+\frac{1}{q}\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\ln\left(2\sqrt{\delta(1-\delta)}\right)^q\\
&= -\frac{1}{q}\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\ln\left(\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\right)\\
&\quad-\frac{1}{q}\frac{1}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\ln\left(\frac{1}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\right)\\
&\quad+\frac{1}{q}\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\ln\left(2\sqrt{\delta(1-\delta)}\right)^q\\
&= \frac{1}{q}\frac{\left(2\sqrt{\delta(1-\delta)}\right)^q}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\ln\left(1+\left(2\sqrt{\delta(1-\delta)}\right)^q\right)\\
&\quad+\frac{1}{q}\frac{\ln\left(1+\left(2\sqrt{\delta(1-\delta)}\right)^q\right)}{1+\left(2\sqrt{\delta(1-\delta)}\right)^q}\\
&= \frac{1}{q}\ln\left[1+\left(2\sqrt{\delta(1-\delta)}\right)^q\right]\doteq R_q(\delta).
\end{aligned}\tag{43}$$

In order to prove (39), it is now sufficient to show that $R_q(\delta)\le\widehat{r}^{(0)}(\delta)$.

Define the auxiliary function

$$F_q(\delta)=q[\widehat{r}^{(0)}(\delta)-R_q(\delta)]=H(\delta)-\ln\left[1+\left(2\sqrt{\delta(1-\delta)}\right)^q\right].\tag{44}$$

Since the sequence of functions $\{F_q(\delta)\}_{q\ge 2}$ is increasing in q, then we have that $F_q(\delta)\ge F_2(\delta)$ for $0\le\delta\le 1$. So it is sufficient to verify that $F_2(\delta)\ge 0$, $\forall\delta\in[0,1]$.

From (44) we have that

$$\begin{aligned}
F_2(\delta) &= H(\delta)-\ln\left[1+\left(2\sqrt{\delta(1-\delta)}\right)^2\right]\\
&= H(\delta)-\ln\left[1+4\delta(1-\delta)\right].
\end{aligned}$$

Since $F_2(\delta)=F_2(1-\delta)$, it is sufficient to verify that $F_2(\delta)\geq 0,\ \forall\delta\in[0,1/2]$. We have that

$$F_2(0)=0 \quad\text{and}\quad F_2(1/2)=0. \tag{45}$$

Differentiating F_2 we get

$$\frac{\mathrm{d}}{\mathrm{d}\delta}F_2(\delta)=\ln\left(\frac{1-\delta}{\delta}\right)-\frac{4(1-2\delta)}{1+4\delta(1-\delta)}$$

from which it follows that

$$\lim_{\delta\to 0}\frac{\mathrm{d}}{\mathrm{d}\delta}F_2(\delta)=\infty \qquad \frac{\mathrm{d}}{\mathrm{d}\delta}F_2(\delta)\Big|_{\delta=1/2}=0. \tag{46}$$

Moreover

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}F_2(\delta)&=-\frac{1}{1-\delta}-\frac{1}{\delta}-\frac{[1+4\delta(1-\delta)](-8)-(4-8\delta)^2}{[1+4\delta(1-\delta)]^2}\\&=-\frac{1}{\delta(1-\delta)}+\frac{(4-8\delta)^2+8(1+4\delta(1-\delta))}{[1+4\delta(1-\delta)]^2}\\&=\frac{-[1+4\delta(1-\delta)]^2+8\delta(1-\delta)(3-4\delta(1-\delta))}{\delta(1-\delta)[1+4\delta(1-\delta)]^2}\\&=\frac{-3[4\delta(1-\delta)]^2+4[4\delta(1-\delta)]-1}{\delta(1-\delta)[1+4\delta(1-\delta)]^2}\end{aligned}$$

from which we get that

$$\begin{aligned}&\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}F_2(\delta)>0 && \forall\delta\in\left(1/2-1/\sqrt{6},1/2\right)\\&\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}F_2(\delta)=0 && \delta=1/2-1/\sqrt{6},\ \delta=1/2\\&\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}F_2(\delta)<0 && \forall\delta\in\left[0,1/2-1/\sqrt{6}\right).\end{aligned}$$

Since F_2 is concave in the interval $\left[0,1/2-1/\sqrt{6}\right]$ and convex in the interval $\left[1/2-1/\sqrt{6},1/2\right]$, and $F_2(1/2-1/\sqrt{6})>0$ we conclude that $F_2(\delta)>0$ for $0<\delta<1/2$. This and the symmetry of F_2 completes the proof for the case with $m=1$.

For the inductive step, we assume that the statement is true for m: from the recursive expression (10) and inductive hypothesis we have

$$\begin{aligned}\widehat{r}^{(m+1)}(\delta)&=\max_{0\leq u\leq 1}\left\{\widehat{r}^{(m)}(u)+f(u,\delta)\right\}\\&\leq\max_{0\leq u\leq 1}\left\{\widehat{r}^{(m-1)}(u)+f(u,\delta)\right\}\\&=\widehat{r}^{(m)}(\delta)\qquad\forall\delta\in[0,1]\end{aligned}$$

and the statement is proved also for $m+1$.

Proof of Proposition 5: Consider

$$G^{(1)}(u,\delta)=\frac{H(u)}{q}+f(u,\delta),\qquad u\in[0,1],\ \delta\in[0,1].$$

1) Differentiating the function $G^{(1)}(u,\delta)$ with respect to the variable u, we get that

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{H(u)}{q}\right)=\frac{1}{q}\ln\frac{1-u}{u}\xrightarrow{u\to 0^+}+\infty.$$

and

$$\begin{aligned}\frac{\partial}{\partial u}f(u,\delta)&=\frac{\partial}{\partial u}\left(-H(u)+(1-\delta)H\left(\frac{u}{2(1-\delta)}\right)\right.\\&\qquad\left.+\delta H\left(\frac{u}{2\delta}\right)\right)\\&=-\ln\left(\frac{1-u}{u}\right)+\frac{1}{2}\ln\left(\frac{2(1-\delta)-u}{u}\right)\\&\quad+\frac{1}{2}\ln\left(\frac{2\delta-u}{u}\right)\\&=\ln u-\ln(1-u)+\frac{1}{2}\ln(2(1-\delta)-u)\\&\quad-\frac{1}{2}\ln u+\frac{1}{2}\ln(2\delta-u)-\frac{1}{2}\ln u\\&=-\ln(1-u)+\frac{1}{2}\ln(2(1-\delta)-u)\\&\quad+\frac{1}{2}\ln(2\delta-u)\\&=\frac{1}{2}\ln\frac{(2(1-\delta)-u)(2\delta-u)}{(1-u)^2}\\&=\frac{1}{2}\ln\frac{4\delta(1-\delta)-2u+u^2}{(1-u)^2}\end{aligned}\tag{47}$$

from which we have

$$\begin{aligned}\frac{\partial}{\partial u}f(u,\delta)\Big|_{u=0}&=\frac{1}{2}\ln(2\delta)+\frac{1}{2}\ln[2(1-\delta)]\\&=\frac{1}{2}\ln\left(4\delta(1-\delta)\right)\leq 0\qquad\forall\delta\in(0,1).\end{aligned}$$

As $G^{(1)}(0,\delta)=0$ and there exists ϵ_δ such that

$$\frac{\partial}{\partial u}G^{(1)}(u,\delta)>0,\ \forall u\in(0,\epsilon_\delta)$$

then $G^{(1)}(u,\delta)>0,\ \forall\delta$ and for u sufficiently small. From the recursive expression in (10) we conclude that

$$\begin{aligned}\widehat{r}^{(1)}(\delta)&=\max_{0\leq u\leq 1}\{\widehat{r}^{(0)}(u)+f(u,\delta)\}\\&=\max_{0\leq u\leq 1}\{G^{(1)}(u,\delta)\}>0,\quad\forall\delta\in(0,1).\end{aligned}$$

2) By concavity of $H(u)$ and by the fact that

$$\begin{aligned}\frac{\partial^2}{\partial u^2}f(u,\delta)&=\frac{1}{1-u}-\frac{1}{2(2\delta-u)}-\frac{1}{2\,[2(1-\delta)-u]}\\&=\frac{1}{1-u}-\frac{1-u}{(2\delta-u)\,[2(1-\delta)-u]}\\&=\frac{1}{1-u}\left[1-\frac{(1-u)^2}{(2\delta-u)\,[2(1-\delta)-u]}\right]\\&=\frac{1}{1-u}\left[\frac{(2\delta-u)\,[2(1-\delta)-u]-(1-u)^2}{(2\delta-u)\,[2(1-\delta)-u]}\right]\\&=\frac{1}{1-u}\left[\frac{4\delta(1-\delta)-2u+u^2-(1-u)^2}{(2\delta-u)\,[2(1-\delta)-u]}\right]\\&=-\frac{1-4\delta(1-\delta)}{(1-u)(2\delta-u)\,[2(1-\delta)-u]}\leq 0,\\&\quad\forall\delta,\ u\in\Omega_\delta\end{aligned}$$

and

$$\frac{\partial^2}{\partial u^2}f(u,\delta)=0\iff\delta=1/2$$

we conclude that, for fixed δ, $G^{(1)}(u,\delta)$ is strictly concave in $u \in \Omega_\delta$. As

$$\frac{\partial}{\partial u} G^{(1)}(u,\delta)\Big|_{u=0} = +\infty \quad \frac{\partial}{\partial u} G^{(1)}(u,\delta)\Big|_{u=2\delta} = -\infty$$

we deduce that the maximizing value $u^{(1)}$ of the function $G^{(1)}(u,\delta)$ is unique and $u^{(1)} \in (0, 2\delta \wedge 2-2\delta)$.
Define the function

$$u^{(1)}(\delta) = \operatorname*{argmax}_{u\in\Omega_\delta} G^{(1)}(u,\delta).$$

If we differentiate $G^{(1)}(u,\delta)$ with respect to u, we get that $u^{(1)}(\delta)$ must satisfy the following condition:

$$\frac{\partial}{\partial u} G^{(1)}(u,\delta) = -\left(1-\frac{1}{q}\right)\ln(1-u) - \frac{1}{q}\ln u + \\ +\frac{1}{2}\ln(2-2\delta-u) + \frac{1}{2}\ln(2\delta-u) = 0. \quad (48)$$

Rearranging and defining

$$F(u,\delta) = (u^2 - 2u + 4\delta(1-\delta))^{q/2} - (1-u)^{q-1}u$$

we have that $F(u^{(1)}(\delta),\delta) = 0$.
It can be verified that $F(0,0) = 0$, $\frac{\partial}{\partial u}F(u,\delta) < 0$ and F is C^1, then the theorem of implicit function guarantees that $u_q^{(1)}(\delta)$ is C^1, $\forall q \in \mathbb{N}$.

3) From the bound in (43) we have

$$0 \le \lim_{\delta\to 0^+} \frac{\widehat{r}^{(1)}(\delta)}{\delta} \le \lim_{\delta\to 0^+} \frac{R_q(\delta)}{\delta}$$

where

$$\frac{R_q(\delta)}{\delta} = \frac{\frac{1}{q}\ln\left(1 + \left(2\sqrt{\delta(1-\delta)}\right)^q\right)}{\delta} \\ \sim \frac{1}{q}2^q\delta^{q/2-1} \xrightarrow{\delta\to 0^+} \begin{cases} 0, & \text{for } q \ge 3 \\ 2, & \text{for } q = 2. \end{cases}$$

APPENDIX C
STRICT MONOTONICITY OF $\{\epsilon_m\}_{m\in\mathbb{N}}$

Lemma 6: Let $\delta < 1/2$ and $\Gamma^{(m)}(\delta)$ the set of points such that

$$\Gamma^{(m)}(\delta) = \operatorname*{argmax}_{0\le u\le 2\delta}\{\widehat{r}^{(m-1)}(u) + f(u,\delta)\}. \quad (49)$$

If $\tilde{u} \in \Gamma^{(m)}(\delta)$ then $\tilde{u} \le 1/2$.

Proof: The statement is trivially proved if $\delta \le 1/4$. Consider the case with $1/4 < \delta < 1/2$ and suppose at the contrary that $\tilde{u} \in (1/2, 2\delta]$. As $\widehat{r}^{(m-1)}(u) = \widehat{r}^{(m-1)}(1-u)$ then there exists a point $y \in [1/2-(2\delta-1/2), 1/2]$ such that

$$\widehat{r}^{(m-1)}(\tilde{u}) = \widehat{r}^{(m-1)}(y)$$

and by the fact that $f(u,\delta)$ is decreasing in u we get that

$$\widehat{r}^{(m-1)}(\tilde{u}) + f(\tilde{u},\delta) < \widehat{r}^{(m-1)}(y) + f(y,\delta)$$

and therefore $\tilde{u} \notin \Gamma^{(m)}(\delta)$, which contradicts our assumption. □

Proposition 12: Let $\{\epsilon_m\}_{m\in\mathbb{N}}$ be the sequence of points such that

$$\epsilon_m = \max\{\epsilon \in [0,1/2) : \widehat{r}^{(m)}(\delta) = 0, \forall\delta \le \epsilon\}.$$

The sequence is strictly increasing in $m \in \mathbb{N}$.

Proof: From Proposition 4 follows that $\epsilon_{m+1} \ge \epsilon_m$. We now prove by induction on m that a strictly inequality holds.

As first step, choose $m = 2$: from Corollary 2 it is proved that $\epsilon_2 > 0 = \epsilon_1$.

For the inductive step, assume the statement is true for m, namely $\epsilon_m > \epsilon_{m-1}$.

Let $\Gamma^{(m+1)}(\delta)$ be the set of points defined in (49) and we prove preliminarily that $\Gamma^{(m+1)}(\epsilon_m) = \{0\}$. From Lemma 6, we have that if $u^{(m+1)} \in \Gamma^{(m+1)}(\epsilon_m)$ then $u^{(m+1)} < \frac{1}{2} \wedge 2\epsilon_m$. Suppose at the contrary that $u^{(m+1)} \in (\epsilon_{m-1}, 1/2 \wedge 2\epsilon_m]$ then

$$0 = \widehat{r}^{(m+1)}(\epsilon_m) = \widehat{r}^{(m)}(u^{(m+1)}) + f(u^{(m+1)},\epsilon_m) = 0.$$

From Proposition 4 and from the inductive hypothesis we get

$$0 = \widehat{r}^{(m+1)}(\epsilon_m) < \widehat{r}^{(m-1)}(u^{(m+1)}) + f(u^{(m+1)},\epsilon_m) \\ \le \max_{0\le u\le 2\epsilon_m} \widehat{r}^{(m-1)}(u) + f(u,\epsilon_m) = \widehat{r}^{(m)}(\epsilon_m)$$

then

$$\epsilon_m \ne \max\{\epsilon \in [0,1/2) : \widehat{r}^{(m)}(\delta) = 0, \forall\delta \le \epsilon\}.$$

which contradicts the definition of ϵ_m.

As $\widehat{r}^{(m)}(u) = 0$, for every $u \in [0,\epsilon_m]$ and the function $f(u,\epsilon_m)$ is decreasing in $u \in [0,\epsilon_{m-1}]$ then $\Gamma^{(m+1)}(\epsilon_m) = \{0\}$. We conclude that there exists $\eta > 0$ for which

$$\widehat{r}^{(m)}(u) + f(u,\epsilon_m) \le -\eta \quad \forall u \in (\epsilon_{m-1}, 2\epsilon_m].$$

Using the fact that $\widehat{r}^{(m)}$ and f are both continuous and $f(u,\delta)$ is strictly increasing in $\delta < 1/2$, by the way ϵ_m has been defined (18), we get that there exists $\epsilon' > 0$ such that

$$\widehat{r}^{(m)}(u) + f(u,\epsilon_m + \epsilon') \le 0 \quad \forall u \in (\epsilon_{m-1}, 2(\epsilon_m + \epsilon')]$$

and the statement is proved also for $m+1$. □

APPENDIX D
PROOF OF LEMMA 2

In order to prove Lemma 2 we need to establish some intermediate results. Lemma 7 allows us to get some information about the monotony of nonsmooth functions. The result is surely not original but we give the assertion, as we don't have any reference.

Lemma 7: Let $y : \mathbb{R} \to \mathbb{R}$ be a bounded Lipschitz function such that

$$\limsup_{\eta\to 0} \frac{y(x+\eta) - y(x)}{\eta} \le 0 \quad (50)$$

for all $x \in \mathbb{R}$. Then $y(x)$ is a monotonically decreasing function.

Proof: Notice first that Rademacher's theorem (see [32]) guarantees that y is differentiable at almost every point in $\mathbb{R}$. Let $y' : \mathbb{R} \to \mathbb{R}$ be any bounded measurable function coinciding with the derivative of y when this exists. Clearly $y' \leq 0$ almost surely and it is also easy to see that y' coincides with the distributional derivative of y.

Let now $\{\psi_n\}_{n\in\mathbb{N}}$ be a sequence of C^∞ functions such that

$$\mathrm{supp}(\psi_n) = \left[-\frac{1}{n}, \frac{1}{n}\right] \qquad \int_{-\infty}^{\infty} \psi_n(x)\mathrm{d}x = 1$$

where $\mathrm{supp}(\psi_n)$ is the set of points where the function ψ_n is not zero. Clearly

$$\int_{-\infty}^{\infty} \psi_n(x)y(x)\mathrm{d}x \xrightarrow{n\to\infty} y(0).$$

Fix $a < b$ and consider now the sequence of functions $\{J_n(x)\}_{n\in\mathbb{N}}$ defined by

$$J_n(x) = \int_{-\infty}^{x} [\psi_n(s-b) - \psi_n(s-a)]\,\mathrm{d}s \qquad \forall x.$$

We have that the functions $J_n(x)$ are C^∞, compactly supported and $J_n(x) \leq 0$ for every n. We now have

$$\begin{aligned} 0 \leq \int_{-\infty}^{\infty} J_n(x)y'(x)\mathrm{d}x &= -\int_{-\infty}^{\infty} J_n'(x)y(x)\mathrm{d}x = \\ &= -\int_{-\infty}^{\infty} \psi_n(x-b)y(x)\mathrm{d}x + \int_{-\infty}^{\infty} \psi_n(x-a)y(x)\mathrm{d}x \\ &\xrightarrow{n\to\infty} -y(b) + y(a). \end{aligned}$$

Hence, $y(a) \geq y(b)$. This proves the result. □

For every δ, define the following set:

$$\Gamma^{(m)}(\delta) = \underset{u\in\Omega_\delta}{\mathrm{argmax}}\,\{\widehat{r}^{(m-1)}(u) + f(u,\delta)\} \tag{51}$$

and choose $u^{(m)}(\delta) \in \Gamma^{(m)}(\delta)$. Moreover, we know that $u^{(1)}(\delta)$ is unique.

Lemma 8: For any arbitrary $\epsilon \in]0, 1/2]$, we have:

1) if $u^{(m)}(\delta) \leq 2\delta(1-\delta)$ and $\widehat{r}^{(m)}(\delta)$ is Lipschitz in $\delta \in [\epsilon, 1/2]$, then $\widehat{r}^{(m)}(\delta) - H(\delta)$ is decreasing in $\delta \in [\epsilon, 1/2]$;
2) if $\widehat{r}^{(m)}(\delta) - H(\delta)$ decreases in $\delta \in [\epsilon, 1/2]$, then $u^{(m+1)}(\delta) \leq 2\delta(1-\delta)$ and $\widehat{r}^{(m+1)}(\delta)$ is Lipschitz in $\delta \in [\epsilon, 1/2]$;

Proof:

1) From the hypothesis we know that $u^{(m)}(\delta) \leq 2\delta(1-\delta)$ and we can write that for any arbitrary $\eta > 0$

$$\widehat{r}^{(m)}(\delta+\eta) = \max_{0\leq u\leq 2(\delta+\eta)(1-\delta-\eta)} [\widehat{r}^{(m-1)}(u) + f(u,\delta+\eta)].$$

Using the fact that $\frac{\partial^2}{\partial\delta^2}f(u,\delta) \leq 0$, and $\frac{\partial^2}{\partial\delta\partial u}f(u,\delta) \geq 0\ \forall\delta \leq 1/2, \forall u \in \Omega_\delta$, we can estimate, for $u \leq 2(\delta+\eta)(1-\delta-\eta)$

$$\begin{aligned} f(u,\delta+\eta) &\leq f(u,\delta) + f_\delta(u,\delta)\eta \\ &\leq f(u,\delta) + f_\delta(2(\delta+\eta)(1-\delta-\eta),\delta)\eta. \end{aligned}$$

Hence

$$\begin{aligned} \widehat{r}^{(m)}(\delta+\eta) \leq &\max_{0\leq u\leq 2(\delta+\eta)(1-\delta-\eta)} [\widehat{r}^{(m-1)}(u) + f(u,\delta) + \\ &+ f_\delta(2(\delta+\eta)(1-\delta-\eta),\delta)\eta] \\ \leq &\,\widehat{r}^{(m)}(\delta) + f_\delta(2(\delta+\eta)(1-\delta-\eta),\delta)\eta \end{aligned}$$

where the last inequality follows by the fact that $u^{(m)}(\delta) \leq 2\delta(1-\delta)$. So we have

$$\frac{\widehat{r}^{(m)}(\delta+\eta) - \widehat{r}^{(m)}(\delta)}{\eta} \leq f_\delta(2(\delta+\eta)(1-\delta-\eta),\delta)$$

and

$$\limsup_{\eta\to 0} \frac{\widehat{r}^{(m)}(\delta+\eta) - \widehat{r}^{(m)}(\delta)}{\eta} \leq f_\delta(2\delta(1-\delta),\delta) = H'(\delta).$$

From Lemma 7 we conclude that $\widehat{r}^{(m)}(\delta) - H(\delta)$ decreases in $\delta \in [\epsilon, 1/2]$.

2) We prove it by contradiction. If we assume that, for some $\delta \in [\epsilon, 1/2]$, it holds $u^{(m+1)}(\delta) > 2\delta(1-\delta)$, then

$$\begin{aligned} \widehat{r}^{(m+1)}(\delta) &= \widehat{r}^{(m)}(u^{(m+1)}(\delta)) + f(u^{(m+1)}(\delta),\delta) = \\ &= \widehat{r}^{(m)}(u^{(m+1)}(\delta)) - H(u^{(m+1)}(\delta)) + \\ &\quad + H(u^{(m+1)}(\delta)) + f(u^{(m+1)}(\delta),\delta). \end{aligned}$$

From the hypothesis and from (23) it can be upper bounded as follows:

$$\begin{aligned} \widehat{r}^{(m+1)}(\delta) &< \widehat{r}^{(m)}(2\delta(1-\delta)) - H(2\delta(1-\delta)) + \\ &\quad + H(2\delta(1-\delta)) + f(2\delta(1-\delta),\delta) = \\ &= \widehat{r}^{(m)}(2\delta(1-\delta)) + f(2\delta(1-\delta),\delta). \end{aligned}$$

This is absurd by the definition of $\widehat{r}^{(m+1)}$.

We now prove the second part of 2). Let $\delta_1 < \delta_2 \in [\epsilon, 1/2]$. We have

$$u^{(m+1)}(\delta_2) \in [0, 2\delta_2(1-\delta_2)].$$

Since

$$\frac{\partial}{\partial\delta}f(u,\delta) = \ln\left(1 - \frac{u}{2(1-\delta)}\right) - \ln\left(1 - \frac{u}{2\delta}\right)$$

is continuous in $\delta \in [\epsilon, 1/2]$ and $u \in [0, 2\delta(1-\delta)]$, Weierstrass's theorem guarantees that $|\frac{\partial f}{\partial\delta}|$ attains its maximum $K \in \mathbb{R}$ over a closed bounded domain. By applying Lagrange's theorem in the variable δ we have that $\exists\,\xi \in (\delta_1, \delta_2)$ such that

$$\begin{aligned} |f(u,\delta_2) - f(u,\delta_1)| &= \left|\frac{\partial f}{\partial\delta}(u,\xi)(\delta_2-\delta_1)\right| \\ &= \left|\frac{\partial f}{\partial\delta}(u,\xi)\right| |\delta_2-\delta_1| \leq K|\delta_2-\delta_1| \end{aligned}$$

and we conclude that $f(u,\delta)$ is Lipschitz in $\delta \in [\epsilon, 1/2]$ uniformly in $u \in [0, 2\delta(1-\delta)]$.

It follows that

$$\begin{aligned} \widehat{r}^{(m+1)}(\delta_2) &= \max_{0\leq u\leq 2\delta_2(1-\delta_2)} \{\widehat{r}^{(m)}(u) + f(u,\delta_2)\} \\ &\leq \max_{0\leq u\leq 2\delta_1(1-\delta_1)} \{\widehat{r}^{(m)}(u) + f(u,\delta_1)\} + K|\delta_2-\delta_1| \\ &= \widehat{r}^{(m+1)}(\delta_1) + K|\delta_2-\delta_1|. \end{aligned}$$

Similarly, we can estimate,

$$\widehat{r}^{(m+1)}(\delta_2) \geq \widehat{r}^{(m+1)}(\delta_1) - K|\delta_2-\delta_1|.$$

We conclude that

$$|\hat{r}^{(m+1)}(\delta_2) - \hat{r}^{(m+1)}(\delta_1)| \leq K|\delta_2 - \delta_1| \quad \forall \delta_1, \delta_2 \in [\epsilon, 1/2].$$

Notice that the constant K only depends on f and ϵ, and not on m. □

Proof of Lemma 2: We first consider the case $m = 1$. Let

$$u_q^{(1)}(\delta) = \operatorname*{argmax}_{u \in \Omega_\delta} [H(u)/q + f(u, \delta)].$$

If we consider the case $q = 2$, we find the analytical expression

$$u_2^{(1)}(\delta) = \frac{3 - \sqrt{9 - 32\delta(1-\delta)}}{4} \qquad \forall \delta \in [0, 1/2]$$

by which $u_2^{(1)}(\delta) \leq 2\delta(1-\delta)$ and $u_2^{(1)}(\delta) = 2\delta(1-\delta) \iff \delta = 0$ or $\delta = 1/2$.

We prove now that $\{u_q^{(1)}(\delta)\}_{q \in \mathbb{N}}$ is a decreasing sequence of functions in q. Supposing *ab absurdo* that $u_q^{(1)}(\delta) < u_{q+1}^{(1)}(\delta)$,

$$\begin{aligned}
\hat{r}_q^{(1)}(\delta) &= \frac{H(u_q^{(1)}(\delta))}{q} + f(u_q^{(1)}(\delta), \delta) + \\
&\quad + \frac{H(u_q^{(1)}(\delta))}{q+1} - \frac{H(u_q^{(1)}(\delta))}{q+1} \leq \\
&\leq \frac{H(u_q^{(1)}(\delta))}{q} - \frac{H(u_q^{(1)}(\delta))}{q+1} + \\
&\quad + f(u_{q+1}^{(1)}(\delta), \delta) + \frac{H(u_{q+1}^{(1)}(\delta))}{q+1} < \\
&< \frac{H(u_{q+1}^{(1)}(\delta))}{q} - \frac{H(u_{q+1}^{(1)}(\delta))}{q+1} \\
&\quad + f(u_{q+1}^{(1)}(\delta), \delta) + \frac{H(u_{q+1}^{(1)}(\delta))}{q+1} = \\
&= f(u_{q+1}^{(1)}(\delta), \delta) + \frac{H(u_{q+1}^{(1)}(\delta))}{q}
\end{aligned}$$

we get that $u_q^{(1)}(\delta) \neq \operatorname*{argmax}_{u \in \Omega_\delta} [r^{(1)}(u) + f(u, \delta)]$.

So we have $u_q^{(1)}(\delta) \leq u_2^{(1)}(\delta) \leq 2\delta(1-\delta)$, $\forall q \in \mathbb{N}$.

Notice that $\hat{r}^{(1)}(\delta)$ is differentiable and $u^{(1)}(\delta) \leq 2\delta(1-\delta)$. Applying, inductively, Lemma 8 for some $\epsilon \in (0, \epsilon_2)$ we obtain that $\hat{r}^{(m)}(\delta)$ are all Lipschitz in $\delta \in [\epsilon, 1/2]$, $\forall m$. As $\hat{r}^{(m)}(\delta)$ is symmetric respect to axis $\delta = 1/2$ and $\hat{r}^{(m)}(\delta) = 0 \; \forall \delta \leq \epsilon$, $\hat{r}^{(m)}(\delta)$ is Lipschitz in every point in $[0, 1]$, $\forall m$.

Notice that the Lipschitz's constant K is the same for every spectral shape.

Appendix E
Proof of Lemma 4

We prove the assertion by induction on m. Consider the case $m = 1$ as initial step: we have that

$$\begin{aligned}
\sum_{h=1}^{h_N} \overline{A}_h(RA) &= \sum_{h=1}^{h_N} \sum_{w=q}^{2h_N} A_w(\mathrm{Rep}_N^q) P_{w,h}(\mathrm{Acc}_N) \\
&= \sum_{w=q}^{2h_N} A_w(\mathrm{Rep}_N^q) \sum_{h=1}^{h_N} P_{w,h}(\mathrm{Acc}_N) \\
&\leq \sum_{w=q}^{2h_N} \binom{N/q}{\lfloor w/q \rfloor} \sum_{h=1}^{h_N} P_{w,h}(\mathrm{Acc}_N)
\end{aligned}$$

and from (37c) in Appendix A we get

$$\begin{aligned}
&\sum_{h=1}^{h_N} \overline{A}_h(RA) \\
&\quad \leq \sum_{w=q}^{2h_N} \binom{N/q}{\lfloor w/q \rfloor} \sum_{h=1}^{h_N} \frac{\lceil w/2 \rceil}{h} \frac{\binom{N-h}{\lfloor w/2 \rfloor}\binom{h}{\lceil w/2 \rceil}}{\binom{N}{w}} \\
&\quad \leq \sum_{w=q}^{2h_N} \binom{N/q}{\lfloor w/q \rfloor} \sum_{h=1}^{h_N} \frac{\lceil w/2 \rceil}{h} 2^w \\
&\quad \times \left(\frac{h}{N}\right)^{\lceil w/2 \rceil} \left(1 - \frac{h}{N}\right)^{\lfloor w/2 \rfloor} \\
&\quad \leq \sum_{w=q}^{2h_N} \lceil w/2 \rceil 2^w N^{\lfloor w/q \rfloor - \lceil w/2 \rceil} \sum_{h=1}^{h_N} h^{\lceil w/2 \rceil - 1} \\
&\quad \leq \sum_{w=q}^{2h_N} \lceil w/2 \rceil 2^w N^{\lfloor w/q \rfloor - \lceil w/2 \rceil} \int_0^{h_N} x^{\lceil w/2 \rceil - 1} \mathrm{d}x \\
&\quad = \sum_{w=q}^{2h_N} \lceil w/2 \rceil 2^w N^{\lfloor w/q \rfloor - \lceil w/2 \rceil} \frac{h_N^{\lceil w/2 \rceil}}{\lceil w/2 \rceil} \\
&\quad \leq 2h_N \max_{q \leq w \leq 2h_N} \left\{ 2^w N^{\lfloor w/q \rfloor - \lceil w/2 \rceil} h_N^{\lceil w/2 \rceil} \right\} \\
&\quad = 2h_N \max_{q \leq w \leq 2h_N} \left\{ N^{w/\log_2 N + \lfloor w/q \rfloor} \left(\frac{h_N}{N}\right)^{\lceil w/2 \rceil} \right\}.
\end{aligned}$$

Let now $\eta > 0$ be an arbitrary small number, and $\eta' = \frac{\eta}{1 + 2\lceil q/2 \rceil}$. From (30), then we get

$$\lim_{N \to \infty} \frac{h_N}{N^{\eta'}} = 0.$$

Notice that

$$\begin{aligned}
2h_N \left(\frac{h_N}{N}\right)^{\lceil w/2 \rceil} &= 2N^{\eta'} \left(\frac{h_N}{N^{\eta'}}\right) \left(\frac{h_N}{N^{\eta'} N^{1-\eta'}}\right)^{\lceil w/2 \rceil} \\
&= 2N^{\eta'} \left(\frac{h_N}{N^{\eta'}}\right)^{\lceil w/2 \rceil + 1} N^{-(1-\eta')\lceil w/2 \rceil}
\end{aligned}$$

then we have

$$\begin{aligned}
&2h_N \max_{q \leq w \leq 2h_N} \left\{ N^{w/\log_2 N + \lfloor w/q \rfloor} \left(\frac{h_N}{N}\right)^{\lceil w/2 \rceil} \right\} = \\
&N^{\eta'} \max_{q \leq w \leq 2h_N} \left\{ 2N^{\frac{w}{\log_2 N} + \lfloor w/q \rfloor - (1-\eta')\lceil \frac{w}{2} \rceil} \left(\frac{h_N}{N^{\eta'}}\right)^{\lceil \frac{w}{2} \rceil + 1} \right\}.
\end{aligned}$$

Let $C > 0$ be a constant. Since $h_N/N^{\eta'} \xrightarrow{N \to \infty} 0$, we have that for $w \geq q$ and N large enough

$$2\left(\frac{h_N}{N^{\eta'}}\right)^{1 + \lceil w/2 \rceil} \leq 2\left(\frac{h_N}{N^{\eta'}}\right)^{1 + \lceil q/2 \rceil} \leq C.$$

Also for $q \leq w \leq 2h_N$ and N large enough

$$\frac{w}{\log_2 N} + \left\lfloor \frac{w}{q} \right\rfloor - (1-\eta') \left\lceil \frac{w}{2} \right\rceil \leq \left\lfloor \frac{w}{q} \right\rfloor - (1-2\eta') \left\lceil \frac{w}{2} \right\rceil \leq 1 - (1-2\eta') \left\lceil \frac{q}{2} \right\rceil.$$

Hence, for fixed $C > 0$ and for N large enough we get that

$$\begin{aligned}\sum_{h=1}^{h_N} \overline{A}_h(RA) &\leq 2h_N \max_{q \leq w \leq 2h_N} \left\{ N^{w/\log_2 N + \lfloor w/q \rfloor} \left(\frac{h_N}{N}\right)^{\lceil w/2 \rceil} \right\} \\ &\leq CN^{\eta' + 1 - (1-2\eta')\lceil q/2 \rceil} \\ &= CN^{1-\lceil q/2 \rceil + \eta'(1+2\lceil q/2 \rceil)} \\ &= CN^{1-\lceil q/2 \rceil + \eta}.\end{aligned}$$

Assume now that this statement is true for the case $m-1$, we have

$$\begin{aligned}\sum_{h=1}^{h_N} \overline{A}_h(RA^m) &= \sum_{h=1}^{h_N} \sum_{w=1}^{2h_N} \overline{A}_w(RA^{m-1}) P_{w,h}(\mathrm{Acc}_N) \\ &= \sum_{w=1}^{2h_N} \overline{A}_w(RA^{m-1}) \sum_{h=1}^{h_N} P_{w,h}(\mathrm{Acc}_N).\end{aligned}$$

Let now $\eta > 0$ be an arbitrary small number and $\eta' = \eta/(1+2\lceil q/2^m \rceil)$.

From inductive hypothesis we have that for fixed $C > 0$ and for large enough N

$$\sum_{h=1}^{2h_N} \overline{A}_h(RA^{m-1}) \leq CN^{1 - \sum_{i=1}^{m-1} \lceil q/2^i \rceil + \eta'}.$$

It follows that

$$\begin{aligned}\sum_{h=1}^{h_N} \overline{A}_h(RA^m) &\leq CN^{1 - \sum_{i=1}^{m-1} \lceil q/2^i \rceil + \eta'} \sum_{w=\lceil q/2^{m-1} \rceil}^{2h_N} 2^w \left(\frac{h_N}{N}\right)^{\lceil w/2 \rceil} \\ &\leq CN^{1 - \sum_{i=1}^{m-1} \lceil q/2^i \rceil + \eta'} \times \\ &\quad \times 2h_N \max_{\lceil q/2^{m-1} \rceil \leq w \leq 2h_N} 2^w N^{-\lceil w/2 \rceil} h_N^{\lceil w/2 \rceil} \\ &\leq CN^{1 - \sum_{i=1}^{m} \lceil q/2^i \rceil + \eta}.\end{aligned}$$

Then the statement is proved also for m.

ACKNOWLEDGMENT

The authors express our sincere gratitude to the Associate Editor for providing more accurate numerical results used in Tables I and II and for improving the readability of the proofs. They wish to thank Prof. R. Urbanke for useful discussions and Prof. H. Pfister for sharing material and ideas with us. They also thank the Reviewers for carefully reading the manuscript and for their valuable suggestions.

REFERENCES

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

[2] H. Jin, "Analysis and design of turbo-like codes," Ph.D. dissertation, Calif. Inst. Technol., Pasadena, CA, 2001.

[3] H. Jin and R. J. McEliece, "RA codes achieve AWGN channel capacity," in *Proc. 13th Int. Symp. Appl. Algebra, Algebraic Algorithms and Error Correcting Codes*, Honolulu, HI, Nov. 1999, pp. 10–18.

[4] L. Bazzi, "Minimum distance of error correcting codes versus complexity, simmetry and pseudorandomness," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 2003.

[5] H. D. Pfister, "On the capacity of the finite state channels and the analysis of convolutional accumulate-m codes," Ph.D. dissertation, Univ. Calif. San Diego, La Jolla, CA, 2003.

[6] J. Feldman and D. Karger, "Decoding turbo-like codes via linear programming," *J. Comp. Syst. Sci.*, vol. 68, no. 4, pp. 733–752, Jun. 2004.

[7] M. Peleg, I. Sason, S. Shamai, and A. Elia, "On interleaved, differentially encoded convolutional codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2572–2582, Nov. 1999.

[8] A. Brown, M. Luby, and A. Shokrollahi, "Repeat-accumulate codes that approach the Gilbert-Varshamov bound," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 2005, pp. 169–173.

[9] C. Hsu and A. Anastasopoulos, "Asymptotic weight distributions of irregular repeat-accumulate codes," in *Proc. 2005 IEEE Global Telecommun. Conf.*, Nov. 2005, vol. 3, pp. 1147–1151.

[10] H. Jin, A. Khandekar, and R. McEliece, "Irregular repeat-accumulate codes," in *Proc. 2nd Int. Symp. Turbo Codes Related Top.*, Brest, France, Sep. 2000.

[11] I. Sason and I. Goldenberg, "Coding for parallel channels: Gallager bounds and applications to turbo-like codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2394–2428, Jul. 2007.

[12] J. Li, K. Narayanan, and C. Georghiades, "Generalized product accumulate codes: Analysis and performance," in *Proc. IEEE Global Telecommun. Conf., 2001*, Nov. 2001, vol. 2, pp. 975–979.

[13] J. Li, K. R. Narayanan, and C. N. Georghiades, "Product accumulate codes: A class of codes with near-capacity performance and low decoding complexity," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 31–46, Jan. 2004.

[14] 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.

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

[16] L. Bazzi, M. Mahdian, and A. Spielman, "The minimum distance of turbo-like codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 6–15, Jan. 2009.

[17] D. Costello, J. Kliewer, C. Koller, and F. Vatta, "On the design of double serially concatenated codes," in *Proc. 7th Int. ITG Conf. Source Channel Coding*, Ulm, Germany, Jan. 2008.

[18] 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, Jun. 2003.

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

[20] J. Kliewer, K. Zigangirov, and D. Costello, "New results on the minimum distance of repeat multiple accumulate codes," in *Proc. 45th Ann. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, Sep. 2007.

[21] D. J. Costello, C. Koller, J. Kliewer, and K. S. Zigangirov, "On the distance growth properties of double serially concatenated convolutional codes," in *Proc. Inf. Theory Appl. Workshop*, Jan. 2008.

[22] C. Koller, J. Kliewer, K. Zigangirov, and D. Costello Jr., "Minimum distance bounds for multiple-serially concatenated code ensembles," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1888–1892.

[23] J. Kliewer, K. Zigangirov, C. Koller, and D. Costello, Coding Theorems for Repeat Multiple Accumulate Codes Oct. 2008 [Online]. Available: arXive.org

[24] F. Fagnani and C. Ravazzi, "Spectra and minimum distances of repeat multiple accumulate codes," in *Proc. Inf. Theory Appl. Workshop*, La Jolla, CA, Jan. 2008, pp. 77–86.

[25] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[26] A. Barg and G. Forney, "Random codes: Minimum distances and error exponents," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2568–2573, Sep. 2002.

[27] J. N. Pierce, "Limit distrubution of the minimum distance of random linear codes," *IEEE Trans. Inf.. Theory*, vol. IT-13, no. 4, pp. 595–599, Oct. 1967.

[28] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.

[29] 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 Commun.*, vol. 16, pp. 231–244, Feb. 1998.

[30] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.

[31] I. Sason, E. Telatar, and R. Urbanke, "On the asymptotic input output weight distributions and thresholds of convolutional and turbo-like encoders," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3052–3061, Dec. 2002.

[32] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*. Berlin, Germany: Springer, 1998.

Chiara Ravazzi received the B.Sc. and M.Sc. degrees in applied mathematics from Politecnico di Torino, Torino, Italy, in 2005 and 2007, respectively. She is currently working towards the Ph.D. degree at the Department of Mathematics (DIMAT), Politecnico di Torino.

Her current interests include asymptotic analysis of turbo-like codes and structured LDPC, graphical models, Belief Propagation algorithm, and its applications.

Fabio Fagnani received the Laurea degree in mathematics from the University of Pisa, Pisa, Italy, and Scuola Normale Superiore, Pisa, in 1986, and the Ph.D. degree in mathematics from the University of Groningen, Groningen, The Netherlands, in 1991.

During 1991–1998, he was an Assistant Professor with the Scuola Normale Superiore. Since 1998, he has been with the Politecnico di Torino, Torino, Italy, where he is currently Full Professor of Mathematical Analysis. His research activities are on the fundamental mathematical aspects of systems and control theory and of coding theory on which he is author of about 30 papers on international journals. Specific themes of current interest are: control under communication constraints and coordinated control and their connection with graph theory and symbolic dynamics; inverse problems and recursive deconvolution techniques; codes over groups and their use in high-performance schemes.