Finite-blocklength analysis of rate-compatible codes

Tsung-Yi Chen
Electrical Engineering & Computer Science Dept.
Northwestern University
Evanston, IL 60208
Email: tsungyi.chen@northwestern.edu

Dariush Divsalar
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109
Email: Dariush.Divsalar@jpl.nasa.gov

Richard D. Wesel
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, California 90095
Email: wesel@ee.ucla.edu

***Abstract*—This paper considers finite-blocklength achievability for rate-compatible codes. For a fixed number of messages, random coding analysis determines a sequence of achievable error probabilities for a sequence of blocklengths. However, traditional random coding achievability draws each code independently so that it does not show that a family of rate-compatible codes can achieve that sequence of error probabilities. Using random code extension, this paper shows achievable frame error rates for rate-compatible channel codes with finite blocklengths. This paper also shows that when a threshold decoder is used, the rate-compatible constraint does not affect the achievable error rates for a class of input-invariant channels. The results are applied to the binary-input symbol-wise independent fading channel with channel state information at the receiver, and binary-input AWGN channels using Chernoff bounds.**

I. Introduction

Incremental redundancy (IR) systems with receiver confirmation are widely used in modern communication systems. Receiver confirmation refers to a class of feedback systems where the confirmation (a decision to conclude a transmission session) is determined at the receiver. For detailed discussion on different types of confirmation see [1]. To achieve high expected throughput, modern incremental redundancy systems often use a good family of rate-compatible channel codes that provides better error protection as the number of received symbols increases.

Rate-compatible punctured convolutional (RCPC) codes and rate-compatible punctured turbo (RCPT) codes are among the most popular rate-compatible channel codes used in IR systems. Another ideal candidate to construct a family of rate-compatible codes is Low-Density Parity-Check (LDPC) codes. Aiming to achieve high throughput in Hybrid automatic repeat request (HARQ) systems in various classes of channels, numerous heuristics have been proposed to construct rate-compatible LDPC codes. The first work in the construction of rate-compatible LDPC codes appears to be [9]. See also [10]–[12] and the references therein. Recently we proposed a class of rate-compatible LDPC codes called protograph-based Raptor-like (PBRL) LDPC codes [13]. This class of rate-compatible LDPC codes is constructed by extending a base code, motivating the study of finite-blocklength achievability for rate-compatible codes based on code extension.

This material is based upon work supported by the National Science Foundation (NSF) under Grant Numbers 1162501 and 1161822. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF. This research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA JPL Task Plan 82-17473.

For a sequence of codes that asymptotically achieves capacity, finite-blocklength analysis can be considered from two perspectives: 1) the rate of convergence of the error probability as a function of blocklength and 2) the back-off of the channel coding rate from the asymptotic capacity as a function of the blocklengths and the desired error probabilities. From the first perspective, the channel coding rate is fixed and the analysis is focused on the error probability. This type of analysis appeared in the work of Feinstein, Shannon and Gallager [2]–[4], among others. The analysis from the second perspective first appeared in the work of Weiss, Dobrushin and Strassen [5]–[7]. Recently Polyanskiy, Poor and Verdú [8] revisited finite-blocklength analysis from the second perspective. A thorough review is provided and new non-asymptotic achievability bounds and converse are proposed in [8].

This paper studies the finite-blocklength performance for rate-compatible codes using random coding. The main contributions of this paper are:

1) Achievability for Rate-compatible Codes: This paper provides a channel coding achievability for rate-compatible codes. We show that for a class of channels where the information density is invariant of the input sequence, referred to as input-invariant channels, the achievable error probabilities for rate-compatible codes are identical to the error probabilities for regular codes using a threshold decoder. In other words, we show that the dependence testing (DT) bound in [8] is achievable over the input-invariant channels for a family of rate-compatible codes of various rates.

2) Computation of DT bounds using Chernoff bounds: For binary-input AWGN channel, numerical integrations is required to compute the DT bound exactly. We used Chernoff bounding technique to reduce the computational complexity. We provide an example for binary-input symbol-wise independent fading channel with channel state information (CSI) at the receiver, which includes binary-input additive white Gaussian noise (BI-AWGN) channel as a special. The bounding technique naturally provides lower bounds on the error exponent. The same bounding technique is applicable to other memoryless channels including binary symmetric channels (BSC).

The rest of the paper is organized as follows: Sec. II presents the main results of finite-blocklength analysis for rate-compatible codes. Sec. III presents the Chernoff bounding techniques through an example of binary-input symbol-wise independent fading channel with CSI at the receiver, which also applies to BI-AWGN channel as a special case. Numerical results of the error exponent for BI-AWGN channel are presented. Finally Sec. IV concludes the paper.

II. Rate-Compatible Channel Coding

A. Notation

This subsection briefly introduces our notation. Consider a pair of input and output product spaces $\mathsf{X} = \mathcal{X}^n$ and $\mathsf{Y} = \mathcal{Y}^n$. We denote a random variable (r.v.) by a capitalized letter and its associated instances by the corresponding italicized lowercase letter. For example X denotes a r.v. taking values x in X. We use the following shorthand when the dimension n of a product space is relevant: $x^n = (x_1, x_2, \ldots, x_n)$, where x^n denotes an n-dimensional vector, x_j the jth element of x^n, and x_i^j the ith to jth elements of x^n.

A codebook $\mathcal{C}$ is a collection of M elements in X: $\mathcal{C} = \{c_1, \ldots, c_M\}, c_j \in \mathsf{X}$. A decoder is a (possibly randomized) mapping from the output space Y to a decoder output set $\{0, 1, \ldots, M\}$ where 0 indicates that the decoder cannot decode and declares an error. Because of the possible randomization, the mapping is denoted by the conditional distribution $P_{D|Y}$ where D is the r.v. of the decoder output A codebook and a decoder that has the property $P_{D|X}(j|c_j) \geq 1-\epsilon, \forall j \in \{1, \ldots, M\}$ is called an (M, ϵ) code. In other words, an (M, ϵ) code has M codewords and a maximal error probability less than ϵ. To emphasize the dimension n, a code for the product-space channel $P_{Y^n|X^n} : \mathcal{X}^n \mapsto \mathcal{Y}^n$ is sometimes denoted as an (n, M, ϵ) code.

For a product-space channel $P_{Y|X} : \mathsf{X} \mapsto \mathsf{Y}$ and an input distribution P_X, let the distribution of $(X, \bar{X}, Y)$ be:

$$P_{X\bar{X}Y}(x, \bar{x}, y) = P_X(x)P_X(\bar{x})P_{Y|X}(y|x)\,, \tag{1}$$

i.e., the distribution of $\bar{X}$ is identical to X but independent of Y. Define the information density as

$$i(x; y) = \log \frac{dP_{Y|X}(y|x)}{dP_Y(y)}\,, \tag{2}$$

where $P_Y = \sum_{x\in\mathsf{X}} P_{Y|X}(y|x)P_X(x)$ is the marginal probability induced by P_X and $P_{Y|X}$. Whenever there are multiple r.v.'s involved in the expression we often use $\mathrm{P}[\cdot]$ to denote the probability measure with respect to (w.r.t.) the corresponding probability distributions.

B. Related Results

This subsection reviews the relevant theorem shown in [8].

Theorem 1 (Thm. 21, [8]): For any input distribution P_X and a measurable map $\gamma : \mathsf{X} \mapsto [0, \infty]$, there exists a code with M codewords such that the maximal error probability ϵ satisfies

$$\begin{aligned}\epsilon \leq\, & P_{XY}[i(X; Y) \leq \gamma(X)] \\ & + (M-1) \sup_x P_Y[i(x; Y) > \gamma(x)]\,.\end{aligned} \tag{3}$$

For the case of average error probability see [8]. This paper only studies the case for maximal error probability and will refer Thm. 1 as DT bound.

C. Achievability for Rate-Compatible Codes

This subsections derives the random-coding achievability for rate-compatible codes. We first define a family of rate-compatible codes as follows:

Definition 1: Let $n_1 < n_2 < \cdots < n_m$ be integers. A collection of codes $\{\mathcal{C}_j\}_{j=1}^m$ is said to be a family of rate-compatible codes if each $\mathcal{C}_j$ is an (n_j, M, ϵ_j) code that is the result of puncturing a common mother code, and all the symbols in the higher-rate code $\{\mathcal{C}_j\}$ are also in the lower rate code $\{\mathcal{C}_{j+1}\}$.

A family of rate-compatible codes can be constructed by finding a collection of compatible puncturing patterns satisfying Def. 1 from an (N, M) mother code, $N \geq n_j, j = 1, \ldots, m$. Note that the puncturing becomes straightforward if we reorder the symbols of the mother code so that the symbols of $\mathcal{C}_1$ are first, followed by the symbols of $\mathcal{C}_2$ and so on. In view of this observation, we assume without loss of generality (w.l.o.g.) that the coordinates are reordered such that the puncturing of the mother code is in a sequential order. In other words $x_1^{n_j} \in \mathcal{C}_j, j = 1, 2, \ldots m$.

Let $\mathcal{N} = \{n_1, n_2, \ldots, n_m\}$ be a set of integers such that $n_i < n_j$ if $i < j$. An $(\mathcal{N}, M, \varepsilon)$ code is a family of rate-compatible codes with a collection of possibly randomized decoders $\mathcal{D} = \{D_j\}_{j=1}^m$ and a collection of error probabilities $\varepsilon = \{\epsilon_j\}_{j=1}^m$ such that:

$$P_{D_j|X^{n_j}}\left(i|x_1^{n_j}\right) \geq 1 - \epsilon_j\,, \tag{4}$$

where we let $x_1^{n_j}$ be the ith codeword in $\mathcal{C}_j$. We will focus on the case with a collection of two codes: $\mathcal{N} = \{n_1, n_2\}, \mathcal{D} = \{D_1, D_2\}$ and $\varepsilon = \{\epsilon_1, \epsilon_2\}$. Using induction generalizes the case to $m > 2$.

The following theorem gives the achievable improvement in error probability by extending a code with an increment of d symbols, i.e. letting $n_1 = n, n_2 = n + d$. For simplicity in notation we denote $X \in \mathsf{X}$ and $X' \in \mathsf{X}'$ where $\mathsf{X}' = \mathcal{X}^{n+d}$. Similarly x and x' follow the same convention, i.e., $x = {x'}_1^n$.

Theorem 2: For any input distribution $P_{X^{n_2}} = P_{X^{n_1}}P_{X^d}$ and measurable maps $\gamma : \mathsf{X} \mapsto [0, \infty], \gamma' : \mathsf{X}' \mapsto [0, \infty]$, there exists an $(\{n_1, n_2\}, M, \{\epsilon_1, \epsilon_2\})$ rate-compatible code such that

$$\epsilon_1 \leq \mathrm{P}[i(X; Y) \leq \gamma(X)] + (M-1) \sup_{x\in\mathsf{X}} P_Y[i(x; Y) > \gamma(X)] \tag{5}$$

$$\begin{aligned}\epsilon_2 \leq\, & \mathrm{P}[i(X; Y) \leq \gamma(X)' - \delta]\mathrm{P}[E] + \mathrm{P}[E^c] \\ & + (M-1) \sup_{x'\in\mathsf{X}'} P_{Y'}[i(x'; Y') > \gamma'(x')]\end{aligned} \tag{6}$$

where $E = \left\{i\left({X'}_{n_1+1}^{n_2}; {Y'}_{n_1+1}^{n_2}\right) > \delta\right\}$ for some $\delta > 0$.

Proof: The bound (5) follows from Thm 1 hence it suffices to show (6). For a given code $\{c_1, \ldots, c_M\}$ that satisfies (3), we draw the incremental codeword segments $r_1, \ldots, r_M$

randomly and sequentially according to the marginal probability P_{X^d}. The generated new codewords are denoted as $\{c'_1, \ldots, c'_M\}$, $c'_j \in \mathsf{X}'$. Note that since $P_{\mathsf{X}'} = P_{X^n} P_{X^d}$ and the channel is static and memoryless, we have

$$i\left(x'^n_1; y'^n_1\right) + i\left(x'^{n+d}_{n+1}; y'^{n+d}_{n+1}\right) = i(x'; y'). \tag{7}$$

Let F_j be defined as

$$F_j = \{y' \in \mathsf{Y}' : i(c'_j; y') > \gamma'(c'_j)\}, \tag{8}$$

and let $\mathsf{F}_j = \bigcup_{i=1}^{j} F_i$. The decoder outputs the smallest j such that $y' \in F_j$. The decoding region for codeword c'_j is therefore $F_j \setminus \mathsf{F}_{j-1}$. Letting $r_1 = x^d$ be the incremental codeword segment for the first codeword, the error probability for $x' = (c_1, x^d)$ is given as

$$\epsilon_2^{(1)}(x') = \mathsf{P}[i(x'; Y') \le \gamma'(x') | x' = (c_1, x^d)]. \tag{9}$$

Averaging over all possible choice of x^d for (9) according to the distribution P_{X^d} we have

$$\mathbb{E}[\epsilon_2(x')] = \mathsf{P}[i(c_1; Y) + i(X^d; Y^d) \le \gamma'(x')] \tag{10}$$

$$\begin{aligned} &= \mathsf{P}[i(c_1; Y) \le \gamma'(X') - i(X^d; Y^d), E] \\ &\quad + \mathsf{P}[i(c_1; Y) \le \gamma'(X') - i(X^d; Y^d), E^c] \end{aligned} \tag{11}$$

$$\le \mathsf{P}[i(X; Y) \le \gamma(X)' - \delta]\mathsf{P}[E] + \mathsf{P}[E^c], \tag{12}$$

where $E = \{i(X^d; Y^d) > \delta\}$. Hence there must exist a choice of incremental code segment r_1 such that the resulting new codeword $c'_1 = (c_1, r_1)$ satisfies (10) and (12). Now suppose that we have found the incremental code segments for the first $j-1$ codewords. If we choose x^d as the jth codeword segment, the error probability for $x' = (c_j, x^d)$ is given as

$$\epsilon_2^{(j)}(x') = \mathsf{P}[\{i(x'; Y') \le \gamma'(x')\} \cup \mathsf{F}_{j-1}]. \tag{13}$$

Using the same averaging argument w.r.t. P_{X^d}, there must exist a choice of x^d such that for $x' = (c_j, x^d)$

$$\epsilon_2^{(j)}(x') \le \mathsf{P}[i(c_j; Y) + i(X^d; Y^d) \le \gamma'(X')] + \mathsf{P}[\mathsf{F}_{j-1}]. \tag{14}$$

Upper bounding the first term as in (10)-(12) yields:

$$\begin{aligned} \epsilon_2^{(j)} &\le \mathsf{P}[i(X; Y) \le \gamma(X)' - \delta]\mathsf{P}[E] + \mathsf{P}[E^c] \\ &\quad + \mathsf{P}[\mathsf{F}_{j-1}] \end{aligned} \tag{15}$$

$$\begin{aligned} &\le \mathsf{P}[i(X; Y) \le \gamma(X)' - \delta]\mathsf{P}[E] + \mathsf{P}[E^c] \\ &\quad + (j-1) \sup_{x' \in \mathsf{X}'} P_{Y'}[i(x'; Y') > \gamma'(x')]. \end{aligned} \tag{16}$$

Hence there exists an incremental code segment x^d for c_j such that $c'_j = (c_j, x^d)$ has an error probability upper bounded as:

$$\begin{aligned} \epsilon_2^{(j)} &\le \mathsf{P}[i(X; Y) \le \gamma(X)' - \delta]\mathsf{P}[E] + \mathsf{P}[E^c] \\ &\quad + (j-1) \sup_{x' \in \mathsf{X}'} P_{Y'}[i(x'; Y') > \gamma'(x')]. \end{aligned} \tag{17}$$

By induction we arrive at (6). ∎

For some special cases, e.g. binary-input symbol-wise independent fading channel, BI-AWGN channel, and BSC, the achievability that can be shown using extension with random codeword segments becomes more powerful. Specifically, we have the following:

Theorem 3: If a memoryless channel has the input invariance properties that 1) $P_{Y|X=x}[i(x, Y) \le \alpha]$ is independent of x for any α and 2) $P_Y[i(x, Y) \le \alpha]$ is independent of x for any α, then there exists an $(\{n_1, n_2\}, M, \{\epsilon_1, \epsilon_2\})$ rate-compatible code such that

$$\epsilon_1 \le \mathbb{E}\left[\exp\{-[i(X; Y) - \log(M-1)]^+\}\right], \tag{18}$$

and

$$\epsilon_2 \le \mathbb{E}\left[\exp\{-[i(X'; Y') - \log(M-1)]^+\}\right]. \tag{19}$$

Proof: For the stated input invariance conditions and a fixed γ independent of the input, Thm 2 gives

$$\epsilon_1 \le \mathsf{P}[i(X; Y) \le \gamma] + (M-1) \sup_{x \in \mathsf{X}} P_Y[i(x; Y) > \gamma] \tag{20}$$

$$= P_{Y|X=x}[i(x; Y) \le \gamma] + (M-1) P_Y[i(x; Y) > \gamma]. \tag{21}$$

Starting from (14) in the proof of Thm 2 for the extended code $\mathcal{C}_2$ and for $j = M$, we have

$$\begin{aligned} \epsilon_2 &\le \mathbb{E}_{X^d}\left[P_{Y'|X'=(c_j, x^d)}[i(c_j; Y) + i(x^d; Y^d) \le \gamma']\right] \\ &\quad + (M-1) \sup_{x' \in \mathsf{X}'} [P_{Y'}[i(x'; Y') > \gamma']] \end{aligned} \tag{22}$$

$$\begin{aligned} &= P_{Y'|X'=x'}[i(x'; Y') \le \gamma'] \\ &\quad + (M-1) P_{Y'}[i(x'; Y') > \gamma']. \end{aligned} \tag{23}$$

We can optimize the thresholds by viewing (21) and (23) as the weighted average error of Bayesian hypothesis testing problems between $P_{Y|X=x}$ vs. P_Y and $P_{Y'|X'=x'}$ vs. $P_{Y'}$, respectively. Hence similar to the analysis in [8] the optimal thresholds for both γ and γ' are $\log(M-1)$. ∎

Thus, applying the input invariance conditions to Thm. 2 results in a generalization of the DT bound in (3) to a family of rate-compatible codes with different blocklengths as in (21) and (23). Some examples of input-invariance channels are BSC, BI-AWGN and binary-input symbol-wise independent fading channel.

As shown in [8], DT bound can be computed exactly for BSC and BEC. For other channels such as BI-AWGN channel, however, the computation is intractable. We provide Chernoff bounding technique through an example of binary-input symbol-wise independent fading channel in the following section. BI-AWGN channel is treated as a special case and the error exponent of BI-AWGN channel is also provided.

III. Chernoff Bounding Techniques for Computation of DT Bounds

This section provides a concrete derivation for binary-input symbol-wise independent fading channel verifying that the conditions stated in Thm 3 hold. This channel is discrete input continuous output memoryless channel. The channel produces two outputs one y^n and the other ϕ^n. The output ϕ^n provides CSI to the receiver. Thus the results in Thm 3 applies. The channel can be modeled as

$$y_j = \phi_j x_j + z_j \; ; j = 1, \ldots, n \tag{24}$$

where ϕ_js are independent identically distributed fading samples available at decoder with density function $P_\Phi(\phi)$, and $\mathbb{E}\{\Phi^2\} = 1$. The samples z_js are independent zero-mean, unit variance Gaussian noise samples. Let P be the averaged received power. Thus the averaged received signal-to-noise ratio is $SNR = 2E_s/N_o = P$ where Es/No is signal-to-noise ratio for a coded symbol. In practice this model is valid if a coherent receiver is used. The main goal is to show that the information density $i(x^n; Y^n, \Phi^n)$ is independent of x^n when $(Y^n, \Phi^n) \sim P_{Y^n|\Phi^n=\phi^n} P_{\Phi^n}$ and when $(Y^n, \Phi^n) \sim P_{Y^n|X^n=x^n,\Phi^n=\phi^n} P_{\Phi^n}$. Let the input distribution be i.i.d. where P_X is uniform on $\{-\sqrt{P}, \sqrt{P}\}$. The output distribution $P_{Y^n|\Phi^n} P_{\Phi^n}$ is induced by the input and the channel with unit noise power. Given a set of input and output sequence (x^n, y^n, ϕ^n), the information density is given as

$$\begin{aligned} &i(x^n; y^n, \phi^n) \\ &= \sum_{j=1}^{n} \log \frac{\exp\{-(y_j - \phi_j x_j)^2/2\}}{\frac{1}{2}\left(\exp\left\{\frac{-(y_j-\phi_j\sqrt{P})^2}{2}\right\} + \exp\left\{\frac{-(y_j+\phi_j\sqrt{P})^2}{2}\right\}\right)} \\ &= n\log 2 - \frac{1}{2}\sum_{j=1}^{n} z_j^2 \\ &- \sum_{j=1}^{n} \log\left[\exp\{\frac{-(y_j-\phi_j\sqrt{P})^2}{2} + \exp\{\frac{-(y_j+\phi_j\sqrt{P})^2}{2}\}\right] \end{aligned}$$

where $z_j = y_j - \phi_j x_j$.

Suppose without loss of generality that the first $\nu \in [0, 1]$ portion of an input x^n is $+\sqrt{P}$. The distribution $i(x^n; Y^n, \phi^n)$ when $Y^n|\Phi^n = \phi^n \sim P_{Y^n|X^n=x^n,\Phi^n=\phi^n}$ is the same as

$$\begin{aligned} S_n &= n\log 2 - \frac{1}{2}\sum_{j=1}^{n} Z_j^2 \\ &- \sum_{j=1}^{\nu n} \log\left[\exp\left\{\frac{-Z_j^2}{2}\right\} + \exp\left\{\frac{-(Z_j+2\phi_j\sqrt{P})^2}{2}\right\}\right] \\ &- \sum_{j=\nu n+1}^{n} \log\left[\exp\left\{\frac{-Z_j^2}{2}\right\} + \exp\left\{\frac{-(Z_j-2\phi_j\sqrt{P})^2}{2}\right\}\right] \end{aligned}$$

where $Z_j \sim \mathcal{N}(0, 1)$. Note that Z_j and $-Z_j$ have the same distribution. Hence after some manipulation the distribution of the information density is the same as

$$S_n = n\log 2 - \sum_{j=1}^{n} \log\left(1 + \exp\left\{2\phi_j\sqrt{P}Z_j - 2\phi_j^2 P\right\}\right).$$

If $Y^n|\Phi^n = \phi^n \sim P_{Y^n|\Phi^n=\phi^n}$ then we have

$$\begin{aligned} \bar{S}_n &= n\log 2 - \sum_{j=1}^{\nu n} \frac{(Y_j-\phi_j\sqrt{P})^2}{2} - \sum_{j=\nu n+1}^{n} \frac{(Y_j+\phi_j\sqrt{P})^2}{2} \\ &- \sum_{j=1}^{n} \log\left[\exp\left\{\frac{-(Y_j-\phi_j\sqrt{P})^2}{2}\right\} + \exp\left\{\frac{-(Y_j+\phi_j\sqrt{P})^2}{2}\right\}\right] \end{aligned}$$

Note that since $P_{Y|\Phi=\phi}$ is symmetric, Y_j and $-Y_j$ given $\Phi_j = \phi_j$ have the same distribution. Therefore we have

$$\bar{S_n} = n\log 2 - \sum_{j=1}^{n} \log\left[1 + \exp\left\{2\phi_j\sqrt{P}Y_j\right\}\right] \tag{25}$$

Observe that both S_n and $\bar{S}_n$ are independent of the input x and hence Thm 3 holds for binary input independent fading channel. In other words, there exists an $(\{n_1, n_2\}, M, \{\epsilon_1, \epsilon_2\})$ code such that

$$\epsilon_1 \le \mathrm{P}[S_{n_1} \le \gamma_1] + (M-1)\mathrm{P}[\bar{S}_{n_1} > \gamma_1] \tag{26}$$

$$\epsilon_2 \le \mathrm{P}[S_{n_2} \le \gamma_2] + (M-1)\mathrm{P}[\bar{S}_{n_2} > \gamma_2] \tag{27}$$

In oder to evaluation the error probability for binary input independent fading channel, we need to evaluate the following for a fixed γ':

$$\epsilon \le \mathrm{P}[S_n \le \gamma'] + (M-1)\mathrm{P}[\bar{S}_n > \gamma'] \,. \tag{28}$$

Let $Z_j \sim \mathcal{N}(0, 1)$ and let $A_j, B_j, j = 1, \ldots, n$ be i.i.d. r.v.'s defined as :

$$A_j = \log 2 - \log\left[1 + \exp\{2\phi_j\sqrt{P}Z_j - 2\phi_j^2 P\}\right] \,, \tag{29}$$

$$B_j = \log 2 - \log\left[1 + \exp\{2\phi_j\sqrt{P}Y_j\}\right] \,, \tag{30}$$

where $Y_j, j = 1, \ldots, n$ are i.i.d. r.v.'s distributed as the marginal distribution induced by the uniform input distribution. The optimal γ' that minimizes (28) is $\gamma' = \log(M-1)$ if the probability can be evaluated without further bounding the expressions in (28).

For the binary-input fading channel, however, the probability cannot be evaluated without a large number of multiple numerical integrations. Therefore we focus on finding the optimal threshold γ' in terms of the error exponent. For independent fading channel we make the threshold γ' to be dependent on fading samples. In particular we assume the threshold $\gamma' = \gamma + \zeta\sum_{j=1}^{n}\phi_j^2$. We will optimize both γ and ζ after applying Chernoff bounds to (28).

For an input sequence x^n, we can write the information density $i(x^n; Y^n, \phi^n)$ when $Y^n \sim P_{Y^n|X^n=x^n,\Phi^n=\phi^n}$ as $S_n = \sum_{j=1}^{n} A_j$, and when $Y^n \sim P_{Y^n|\Phi^n=\phi^n}$ as $\bar{S}_n = \sum_{j=1}^{n} B_j$. To find the best error exponent we reserve the choice of the threshold $\gamma' = \gamma + \zeta\sum_{j=1}^{n}\phi_j$ and upper bound the two terms by Chernoff bounds. Denoting $M_n = \frac{\log(M-1)}{n}$, for any $\rho > 0, \eta > 0$, $\gamma > 0$ and ζ we have

$$\epsilon \le \mathrm{P}\left[S_n < \gamma + \zeta\sum_{j=1}^{n}\phi_j^2\right] + (M-1)\mathrm{P}\left[\bar{S}_n \ge \gamma + \zeta\sum_{j=1}^{n}\phi_j^2\right] \tag{31}$$

$$\le \exp\{-nf_n(\gamma, \zeta, \rho)\} + \exp\{-ng_n(\gamma, \zeta, \eta, M)\} \,, \tag{32}$$

where $f_n(\gamma, \zeta, \rho)$ and $g_n(\gamma, \zeta, \eta, M)$ are given as

$$f_n(\gamma, \zeta, \rho) = -\log\mathbb{E}[\exp\{-\rho A_1 + \rho\zeta\phi_1\}] - \frac{\rho\gamma}{n} \,, \tag{33}$$

$$g_n(\gamma, \zeta, \eta, M) = -\log\mathbb{E}[\exp\{\eta B_1 - \eta\zeta\phi_1\}] - M_n + \frac{\eta\gamma}{n} \,. \tag{34}$$

Taking derivative with respect to γ and finding the root yields the following optimal γ^*:

$$\gamma^* = \frac{n}{\rho+\eta}\left(\log\frac{\mathbb{E}[\exp\{\eta B_1 - \eta\zeta\phi_1\}]}{\mathbb{E}[\exp\{-\rho A_1 + \rho\zeta\phi_1\}]} + M_n + \frac{\log\frac{\eta}{\rho}}{n}\right). \tag{35}$$

Plugging it into (33) we get f_n

$$f_n(\zeta,\rho,\eta,M) = -\frac{\eta}{\rho+\eta}\log\mathbb{E}[\exp\{-\rho A_1 + \rho\zeta\phi_1\}$$
$$-\frac{\rho}{\rho+\eta}\log\mathbb{E}[\exp\{\eta B_1 - \eta\zeta\phi_1\}] - \frac{\rho}{\rho+\eta}M_n - \frac{\rho}{\rho+\eta}\frac{\log\frac{\eta}{\rho}}{n} \tag{36}$$

Plugging it into (34) we get g_n

$$g_n(\zeta,\rho,\eta,M) = f_n(\zeta,\rho,\eta,M) + \frac{\log\frac{\eta}{\rho}}{n} \tag{37}$$

To simplify the expression and the optimization of the error exponent, we perform the following change of variables $\eta = \frac{\alpha-\beta}{2\beta}$, and $\rho = \frac{\alpha-\beta}{2(1-\beta)}$, where $\alpha \geq \beta$ and $\beta \in [0,1]$. Then we can calculate (31) as

$$\epsilon \leq \exp\{-nE(n,M,\alpha,\beta,\zeta)\}\ , \tag{38}$$

where the exponent is

$$E(n,M,\alpha,\beta,\zeta)$$
$$= -(1-\beta)\log\mathbb{E}[\exp\{-\frac{\alpha-\beta}{2(1-\beta)}(A_1-\zeta\phi_1)\}$$
$$-\beta\log\mathbb{E}[\exp\{\frac{\alpha-\beta}{2\beta}(B_1-\zeta\phi_1)\}] - \beta M_n - \frac{h(\beta)}{n}$$

where $h(x)$ is the binary entropy function $h(x) = -x\log x - (1-x)\log(1-x)$. Note that $h(\beta) \leq \log(2)$. Finally we get

$$\epsilon \leq \exp\{-nE(n,M)\}\ , \tag{39}$$

The exponent $E(n,M)$ is obtained by maximizing over ζ, α, and β where $0 \leq \beta \leq \alpha$:

$$E(n,M) = \sup_{0\leq\beta\leq\alpha,\zeta} E(n,M,\alpha,\beta,\zeta) \tag{40}$$

Note that the derivation here applies to other channels that have the input invariance property. Setting $\zeta = 0$ and $\phi_j = 1$ for all j yields the BI-AWGN channel and hence the result also applied to BI-AWGN channel. Fig. 1 shows the error exponents of the DT bound and Gallager's random coding bound for the 0.187dB BI-AWGN channel. The gaps between the two exponents are negligible when the rates are close to capacity and the gap increases as the rates decreases.

IV. Conclusion

This paper studied the finite-blocklength analysis for rate-compatible channel codes. In particular, this paper shows that the DT bound in [8] is achievable for rate-compatible codes for a class of input-invariant channels. Chernoff bounding technique is used to compute these bounds for channels where exact computation of DT bound is intractable. This paper also provides a numerical example for the error exponent of the BI-AWGN channel.

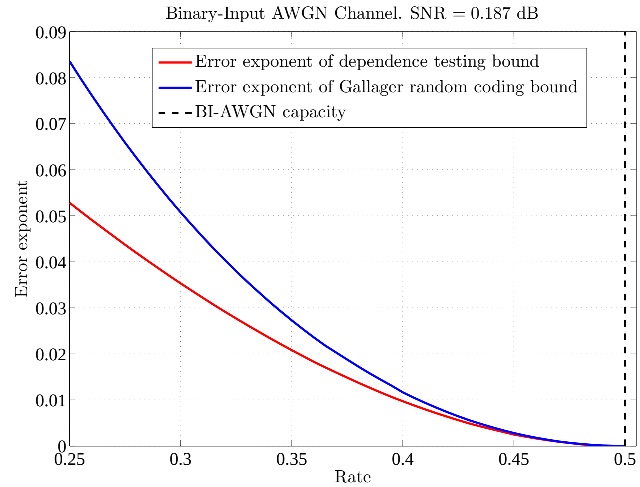

Fig. 1. Error exponents for a binary-input AWGN channel with SNR = 0.187 dB.

References

[1] T.-Y. Chen, A. Williamson, N. Seshadri, and R. Wesel, "Feedback communication systems with limitations on incremental redundancy," *IEEE Trans. Inf. Theory*, (submitted) 2013.

[2] A. Feinstein, "A new basic theorem of information theory," *Information Theory, Transactions of the IRE Professional Group on*, vol. 4, no. 4, pp. 2–22, 1954.

[3] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. PGIT-2, pp. 611–656, 1959.

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[5] L. Weiss, "On the strong converse of the coding theorem for symmetric channels without memory," *Quart. Appl. Math.*, 1960.

[6] R. L. Dobrushin, "Mathematical problems in the shannon theory of optimal coding of information," in *Proc. 4th Berkeley Symp. Mathematics, Statistics, and Probability*, vol. 1, 1961, pp. 211 – 252.

[7] V. Strassen, "Asymptotische abschätzungen in shannons informationstheorie," in *Trans. 3d Prague Conf. Inf. Theory*, Prague, 1962, pp. 689 – 723.

[8] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307 –2359, May. 2010.

[9] J. Li and K. Narayanan, "Rate-compatible low density parity check codes for capacity-approaching ARQ schemes in packet data communications," in *Proc. Int. Conf. on Comm., Internet, and Info. Tech. (CIIT)*, Nov. 2002.

[10] M. Yazdani and A. Banihashemi, "On construction of rate-compatible low-density parity-check codes," *Communications Letters, IEEE*, vol. 8, no. 3, pp. 159–161, 2004.

[11] M. El-Khamy, J. Hou, and N. Bhushan, "Design of rate-compatible structured LDPC codes for hybrid ARQ applications," *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 6, pp. 965–973, 2009.

[12] B. Vellambi and F. Fekri, "Finite-length rate-compatible LDPC codes: a novel puncturing scheme," *Communications, IEEE Transactions on*, vol. 57, no. 2, pp. 297–301, 2009.

[13] T.-Y. Chen, D. Divsalar, J. Wang, and R. D. Wesel, "Protograph-based raptor-like LDPC codes for rate-compatibility with short blocklengths," in *Proc. IEEE Global Communications Conference*, Houston, TX, Dec. 2011.