Concatenated Fountain Codes

Zheng Wang and Jie Luo
Electrical & Computer Engineering Department
Colorado State University, Fort Collins, CO 80523
Email: {zhwang, rockey}@engr.colostate.edu

***Abstract*— This paper investigates fountain communication over discrete-time memoryless channels. Fountain error exponent achievable by linear complexity concatenated fountain codes is derived.**

I. Introduction

In a fountain communication system as illustrated in Figure 1, the encoder maps a message into an infinite sequence of channel input symbols and sends these symbols over a communication channel. Channel output symbols are passed

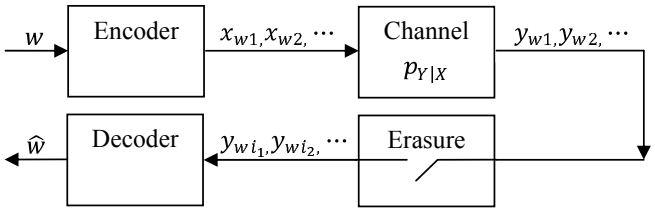

Fig. 1. Fountain communication over a memoryless channel.

through an erasure device which generates *arbitrary* erasures. The decoder outputs an estimated message once the number of received symbols exceeds a pre-determined threshold [1]. Fountain communication rate is defined as the number of transmitted information units normalized by the number of *received* symbols. As shown in [1], fountain communication is useful in many applications such as high rate data transmission over the internet, satellite broadcast, etc.

The first realization of fountain codes are LT codes introduced for binary erasure channels (BECs) by Luby in [2]. LT codes can recover k information bits from $k+O(\sqrt{k}\ln^2(k/\delta))$ encoded symbols with probability $1-\delta$ and a complexity of $O(k\ln(k/\delta))$, for any $\delta > 0$ [2]. Shokrollahi proposed Rapter codes in [3] which is the combination of an appropriate LT code and a pre-code. For BECs, Raptor codes can recover k information bits from $k(1+\epsilon)$ encoded symbols at high probability with complexity $O(k\log(1/\epsilon))$. In [4], Shamai, Telatar and Verdú studied fountain communication over a general stationary memoryless channel. It was shown that the maximum achievable fountain rate for reliable communication, defined as the fountain capacity, is equal to the Shannon capacity of the memoryless channel.

Coding complexity is a crucial concern in practical communication applications. For a conventional communication system, Forney proposed in [5] a one-level concatenated coding scheme that can achieve a positive error exponent, known as Forney's exponent, for any rate less than the Shannon capacity with a polynomial coding complexity. Forney's concatenated codes were generalized in [6] by Blokh and Zyablov to multi-level concatenated codes, whose maximum achievable error exponent is known as the Blokh-Zyablov exponent (or Blokh-Zyablov bound). It was shown in [7] that Forney's and Blokh-Zyablov error exponents can be arbitrarily approached by linear-time encodable/decodable codes.

In this paper, we extend one-level concatenated coding schemes to fountain communication systems over a general discrete-time memoryless channel. We define fountain error exponent in Section II and derive the error exponent achievable by one-level concatenated fountain codes, which concatenate a linear complexity nearly maximum distance separable (MDS) outer code (proposed in [8]) with random fountain inner codes (proposed in [4]). Encoding and decoding complexities of the concatenated fountain codes are linear in the number of transmitted symbols and the number of received symbols, respectively.

All logarithms in this paper are natural based.

II. System Model

Consider the fountain system illustrated in Figure 1. Assume the encoder uses a fountain coding scheme [4] with W codewords to map the source message $w \in \{1, 2, \cdots, W\}$ to an infinite channel input symbol sequence $\{x_{w1}, x_{w2}, \cdots\}$. Assume the channel is discrete-time memoryless, characterized by the conditional point mess function or probability density function $p_{Y|X}(y|x)$, where $x \in X$ and $y \in Y$ are the input and output symbols, X and Y are the input and output alphabets, respectively. Define schedule $\mathcal{N} = \{i_1, i_2, \cdots, i_{|\mathcal{N}|}\}$ as a subset of positive integers, where $|\mathcal{N}|$ is the cardinality of $\mathcal{N}$ [4]. Assume the erasure device generates an arbitrary schedule $\mathcal{N}$, whose elements are indices of the received symbols $\{y_{wi_1}, y_{wi_2}, \cdots y_{wi_{|\mathcal{N}|}}\}$. We say fountain rate of the system is $R = (\log W)/N$, if the decoder outputs an estimate $\hat{w}$ of the source message after observing N channel symbols, i.e., $|\mathcal{N}| = N$, based on $\{y_{wi_1}, y_{wi_2}, \cdots y_{wi_N}\}$ and $\mathcal{N}$. Decoding error happens when $\hat{w} \neq w$. Define error probability $P_e(N)$ as in [4],

$$P_e(N) = \sup_{\mathcal{N}, |\mathcal{N}| \geq N} Pr\{\hat{w} \neq w | \mathcal{N}\}. \quad (1)$$

We say fountain rate R is achievable if there exists a fountain coding scheme with $\lim_{N\to\infty} P_e(N) = 0$ at rate R [4]. The

exponential rate at which error probability vanishes is defined as the fountain error exponent, $E_F(R)$,

$$E_F(R) = \lim_{N\to\infty} -\frac{1}{N}\log P_e(N). \tag{2}$$

Define fountain capacity $\mathcal{C}_F$ as the supremum of all achievable fountain rates. It was shown in [4] that $\mathcal{C}_F$ equals the Shannon capacity of the memoryless channel.

III. Random Fountain Codes

Random fountain coding scheme was firstly introduced in [4] to prove the capacity result. In a random fountain coding scheme, encoder and decoder share a fountain code library $\mathcal{L} = \{C_\theta : \theta \in \Theta\}$, which is a collection of fountain code books C_θ with θ being the code book index. All code books in the library have the same number of codewords and each codeword has infinite number of channel symbols. Let $C_\theta(m)_j$ be the j^{th} codeword symbol of message m in C_θ. To encode the message, the encoder first generates θ according to a distribution γ, such that the random variables $X_{m,j} : \theta \to C_\theta(m)_j$ are i.i.d. with a pre-determined input distribution p_X [4]. Then it uses codebook C_θ to map the message into a codeword. We assume the actual realization of θ is known to the decoder but is unknown to the erasure device[1]. Maximum likelihood decoding is assumed.

Theorem 1: Consider fountain communication over a discrete-time memoryless channel $p_{Y|X}$. Let $\mathcal{C}_F$ be the fountain capacity. For any fountain rate $R < C_F$, random fountain codes achieve the following random-coding fountain error exponent, $E_{Fr}(R)$.

$$E_{Fr}(R) = \max_{p_X} E_{FL}(R, p_X), \tag{3}$$

where $E_{FL}(R, p_X)$ is defined as follows

$$\begin{aligned} E_{FL}(R, p_X) &= \max_{0\le\rho\le 1}\{-\rho R + E_0(\rho, p_X)\}, \\ E_0(\rho, p_X) &= -\log\sum_y \left(\sum_x p_X(x) p_{Y|X}(y|x)^{\frac{1}{1+\rho}}\right)^{(1+\rho)}. \end{aligned} \tag{4}$$

If the channel is continuous, then summations in (4) should be replaced by integrals. ■

Theorem 1 was claimed implicitly in, and can be shown by, the proof of [4, Theorem 2].

$E_{Fr}(R)$ given in (3) equals the random-coding exponent for a conventional communication system over the same channel. For binary symmetric channels (BSCs), since random linear codes simultaneously achieve the random-coding exponent at high rates and the expurgated exponent at low rates [10], it can be easily shown that the same fountain error exponent is achievable by random linear fountain codes. However, because it is not clear whether there exists an expurgation operation, such as the one proposed in [9], that is robust to the observation of any subset of the channel outputs, whether expurgated exponent is achievable for fountain communication over a general discrete-time memoryless channel is therefore unknown.

[1]As demonstrated in [4], the capacity and error exponent results can be significantly different if the erasure device has partial information about θ and is trying to jam the communication.

IV. Concatenated Fountain Codes

Consider a one-level concatenated fountain coding scheme illustrated in Figure 2. Assume source message w can take

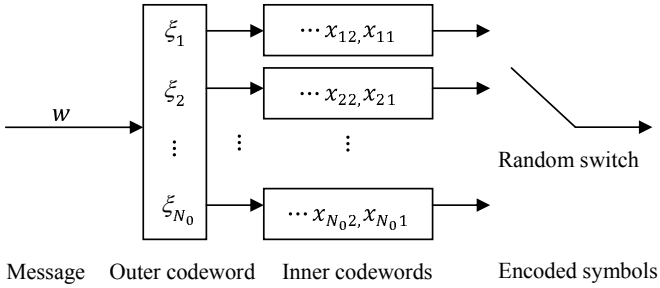

Fig. 2. One-level concatenated fountain codes.

$\exp(NR)$ possible values with an equal probability, where R is the targeted fountain information rate, and decoder decodes the source message after receiving N channel symbols. The encoder first encodes the message using an outer code into an outer codeword, $\{\xi_1, \xi_2, \cdots, \xi_{N_o}\}$, with N_o outer symbols. We assume the outer code is a linear-time encodable/decodable nearly MDS error-correction code of rate $r_o \in [0, 1]$. That is, the outer code can recover the source message from a codeword with d symbol erasures and t symbol errors, so long as $2t + d \le (1 - r_o - \zeta_0)N_o$, where $\zeta_0 > 0$ is a positive constant that can be made arbitrarily small. An example of such linear complexity error-correction code was presented by Guruswami and Indyk in [8]. Each outer symbol ξ_k can take $\exp\left(\frac{N}{N_o}\frac{R}{r_o}\right)$ possible values. Define $N_i = \frac{N}{N_o}$, $R_i = \frac{R}{r_o}$. The encoder then uses a set of random fountain codes (as introduced in [4] and in Section III) each with $\exp(N_i R_i)$ codewords to map each outer symbol ξ_k into an inner codeword, which is an infinite sequence of channel input symbols $\{x_{k1}, x_{k2}, \cdots\}$. Let $C_\theta^{(k)}(\xi_k)_j$ be the j^{th} codeword symbol of the k^{th} inner code in codebook $C_\theta^{(k)}$, where θ is the codebook index as introduced in Section III. We assume θ is generated according to a distribution such that the random variables $X_{k,\xi_k,j} : \theta \to C_\theta^{(k)}(\xi_k)_j$ are i.i.d. with a pre-determined input distribution p_X. To simplify the notations, we have assumed N_i, N_o, NR, and $N_i R_i$ should all be integers. We also assume $N_o \gg N_i \gg 1$.

After encoding, the inner codewords are regarded as N_o channel symbol queues, as illustrated in Figure 2. In the l^{th} time unit, the encoder uses a random switch to pick one inner code with index $k_l(\theta)$ uniformly, and sends the first channel input symbol in the corresponding queue through the channel. The transmitted symbol is then removed from the queue. We assume random variables $k_l : \theta \to \{1, 2, \ldots, N_o\}$ are i.i.d. uniform. We assume the decoder knows the outer codebook

and the code libraries of the inner codes. We also assume the encoder and the decoder share the realization of θ such that the decoder knows the exact codebook used in each inner code and the exact order in which channel input symbols are transmitted.

Decoding starts after $N = N_o N_i$ channel output symbols are received. The decoder first distributes the received symbols to the corresponding inner codewords. Assume $z_k N_i$ channel output symbols are received from the kth inner codeword, where $z_k > 0$ and $z_k N_i$ is an integer. We term z_k the normalized effective codeword length of the kth inner code. Based on z_k, and the received channel output symbols, $\{y_{ki_1}, y_{ki_2}, \ldots, y_{ki_{z_k N_i}}\}$, the decoder computes the maximum likelihood estimate of the outer symbol $\hat{\xi}_k$ together with an optimized reliability weight $\alpha_k \in [0, 1]$. We assume, given z_k and $\{y_{ki_1}, y_{ki_2}, \ldots, y_{ki_{z_k N_i}}\}$, reliability weight α_k is computed using Forney's algorithm presented in [5, Section 4.2]. After that, the decoder carries out a general minimum distance (GMD) decoding of the outer code and outputs an estimate $\hat{w}$ of the source message. GMD decoding of the outer code here is the same as that in a conventional communication system, the detail of which can be found in [7].

Compared to a conventional communication system where all inner codes have the same length, in a concatenated fountain coding scheme, the number of received symbols from different inner codes may be different. Consequently, error exponent achievable by one-level concatenated fountain codes is less than Forney's exponent, as shown in the following theorem.

Theorem 2: Consider fountain communication over a discrete-time memoryless channel $p_{Y|X}$ with fountain capacity $\mathcal{C}_F$. For any fountain rate $R < \mathcal{C}_F$, the following fountain error exponent can be arbitrarily approached by one-level concatenated fountain codes.

$$E_{Fc}(R) = \max_{p_X, \frac{R}{C_F} \le r_o \le 1, 0 \le \rho \le 1} (1 - r_o) \left(-\rho \frac{R}{r_o} + E_0(\rho, p_X) \left[1 - \frac{1 + r_o}{2} E_0(\rho, p_X) \right] \right). \quad (5)$$

where $E_0(\rho, p_X)$ is defined in (4).

Encoding and decoding complexities of the one-level concatenated codes are linear in the number of transmitted symbols and the number of received symbols, respectively. ■

The proof of Theorem 2 is given in Appendix A.

Corollary 1: $E_{Fc}(R)$ is upper-bounded by Forney's error exponent $E_c(R)$ given in [5]. $E_{Fc}(R)$ is lower bounded by $\tilde{E}_{Fc}(R)$, which is defined as,

$$\tilde{E}_{Fc}(R) = \max_{p_X, \frac{R}{C_F} \le r_o \le 1, 0 \le \rho \le 1} (1 - r_o) \left(-\rho \frac{R}{r_o} + E_0(\rho, p_X) \left[1 - E_0(\rho, p_X) \right] \right). \quad (6)$$

As R approaches C_F, the upper and lower bounds are asymptotically equal in the sense of $\lim_{R \to C_F} \frac{\tilde{E}_{Fc}(R)}{E_c(R)} = 1$. ■

The proof of Corollary 1 is skipped.

In Figure 3, we illustrate $E_{Fc}(R)$, $E_c(R)$, and $\tilde{E}_{Fc}(R)$ for a BSC with crossover probability $q = 0.1$. We can

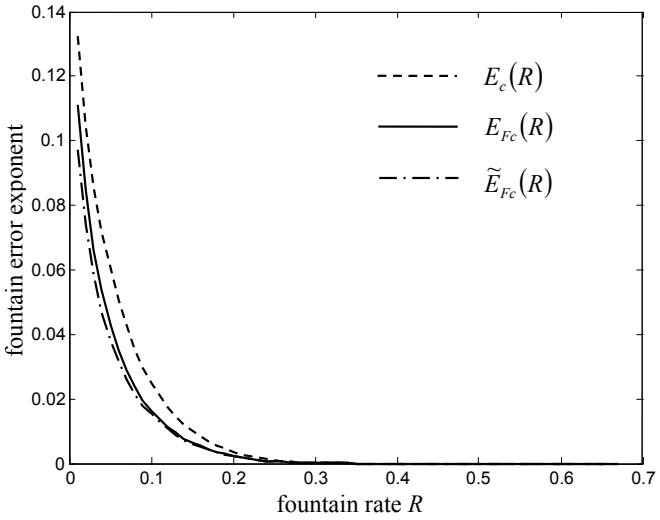

Fig. 3. Comparion of fountain error exponent $E_{Fc}(R)$, its upper bound $E_c(R)$, and its lower bound $\tilde{E}_{Fc}(R)$.

see that $E_{Fc}(R)$ is closely approximated by $\tilde{E}_{Fc}(R)$. This approximation is useful in fountain exponent derivation when the one-level concatenated codes are extended to multi-level concatenated fountain codes.

APPENDIX

A. Proof of Theorem 2

Proof: Assume the decoder starts decoding after receiving $N = N_o N_i$ symbols, where N_o is the length of the outer codeword, N_i is the *expected* number of received symbols from each inner code. In the following fountain error exponent analysis, asymptotic results are obtained by first taking N_o to infinity and then taking N_i to infinity.

Let $\boldsymbol{z}$ be an N_o-dimensional vector whose kth element z_k is the normalized effective codeword length of the kth inner code, from which the conditional empirical distribution function $F_{Z|\theta}$ can be induced, as a function of variable $z \ge 0$, given the random variable θ specified in Section IV. Let the conditional density function of $F_{Z|\theta}$ be $f_{Z|\theta}$. Because the total number of received channel symbols equals $N = N_o N_i$, we must have

$$\int_0^\infty z f_{Z|\theta}(z) dz = 1. \quad (7)$$

Note that $f_{Z|\theta}$ may not be different for each value of θ. Regard $f_{Z|\theta}$ as a random variable and denote its distribution by G_F, as a function of $f_{Z|\theta}$. Assume, given θ, the conditional error probability of the concatenated code can be written as $P_{e|\theta}(f_{Z|\theta}) = \exp(-N_i N_o E_f(f_{Z|\theta}, R))$, where the conditional error exponent $E_f(f_{Z|\theta}, R)$ is a function of $f_{Z|\theta}$. The overall error probability can therefore be written as

$$P_e = \int_\theta \exp(-N_i N_o E_f(f_{Z|\theta}, R)) dG_F(f_{Z|\theta}). \quad (8)$$

Consequently, error exponent of the concatenated code is given by

$$E_{Fc}(R) = \lim_{N_i\to\infty}\lim_{N_o\to\infty} -\frac{1}{N_iN_o}\log\int_\theta \exp(-N_iN_oE_f(f_{Z|\theta},R))dG_F(f_{Z|\theta}) = \min_{f_Z}\left\{E_f(f_Z,R) - \lim_{N_i,N_o\to\infty}\frac{1}{N_iN_o}\log dG_F(f_Z)\right\}, \quad (9)$$

where in the second equality we wrote $f_{Z|\theta}$ as f_Z to simplify the notation.

Next, we will obtain the expression of $-\lim_{N_i,N_o\to\infty}\frac{1}{N_iN_o}\log dG_F(f_Z)$. We will make several approximations during the derivation. These approximations are valid in the sense of not affecting the final result given in (14).

Let $\boldsymbol{z}(i)$ be an N_o-dimensional vector with only one non-zero element corresponding to the ith received symbol. If the ith received symbol belongs to the kth inner code, then we let the kth element of $\boldsymbol{z}(i)$ equal 1 and let all other elements equal 0. Since the random switch (illustrated in Figure 2) picks inner codes uniformly, we have

$$E[\boldsymbol{z}(i)] = \frac{1}{N_o}\mathbf{1}, \quad \text{cov}[\boldsymbol{z}(i)] = \frac{1}{N_o}\boldsymbol{I}_{N_o} - \frac{1}{{N_o}^2}\mathbf{1}\mathbf{1}^T, \quad (10)$$

where $\mathbf{1}$ is an N_o-dimensional vector with all elements being one. According to the definitions, we have $\boldsymbol{z} = \frac{1}{N_i}\sum_{i=1}^{N_iN_o}\boldsymbol{z}(i)$. Note that the $\boldsymbol{z}(i)$ vectors are i.i.d.. According to the central limit theorem, $\boldsymbol{z}$ is approximately Gaussian with mean and covariance matrix given by

$$E[\boldsymbol{z}] = \mathbf{1}, \quad \text{cov}[\boldsymbol{z}] = \frac{1}{N_i}\boldsymbol{I}_{N_o} - \frac{1}{N_oN_i}\mathbf{1}\mathbf{1}^T. \quad (11)$$

Since the total number of received symbols equal N_iN_o, we must have $\mathbf{1}^T\boldsymbol{z} = N_o$. The density function of $\boldsymbol{z}$, denoted by $g(\boldsymbol{z})$, can therefore be approximated by

$$g(\boldsymbol{z}) = \left(\sqrt{\frac{N_i}{2\pi}}\right)^{N_o}\exp\left(\frac{N_i\|\mathbf{1}-\boldsymbol{z}\|^2}{2}\right). \quad (12)$$

As explained before, given $\boldsymbol{z}$, we can obtain the conditional empirical inner codeword length density function f_Z. The density function of f_Z, denoted by g_F, can therefore be written as

$$g_F(f_Z) = K_0(N_i,N_o)g(\boldsymbol{z}), \quad \lim_{N_i,N_o\to\infty}\frac{\log K_0(N_i,N_o)}{N_iN_o} = 0. \quad (13)$$

This consequently yields

$$-\lim_{N_i,N_o\to\infty}\frac{1}{N_iN_o}\log dG_F(f_Z) = \int_0^\infty \frac{(1-z)^2}{2}f_Z(z)dz. \quad (14)$$

Substitute (14) and (7) into (9), we get

$$E_{Fc}(R) = \min_{f_Z,\int_0^\infty zf_Z(z)dz=1}\left\{E_f(f_Z,R) + \int_0^\infty\frac{(1-z)^2}{2}f_Z(z)dz\right\}. \quad (15)$$

Next, we will derive the expression of $E_f(f_Z,R)$, which is the error exponent conditioned on an inner codeword length density f_Z.

Let $\boldsymbol{z}$ be a particular N_o-dimensional inner codewords length vector, which follows the empirical density function f_Z. Since error probability conditioned on f_Z can be written as $P_e(f_Z) = \exp(-N_iN_oE_f(f_Z,R))$, error probability given $\boldsymbol{z}$ can be written as

$$P_e(\boldsymbol{z}) = \frac{\exp(-N_iN_oE_f(f_Z,R))}{K_1(N_i,N_o)}, \quad \lim_{N_i,N_o\to\infty}\frac{\log K_1(N_i,N_o)}{N_iN_o} = 0. \quad (16)$$

Consequently, we can obtain $E_f(f_Z,R)$ by assuming a particular inner codeword length vector $\boldsymbol{z}$, whose empirical inner codeword length density function is f_Z.

Assume the outer code has rate r_o, and is able to recover the source message from d outer symbol erasures and t outer symbol errors so long as $d+2t \le (1-r_o-\zeta_0)$, where $\zeta_0 > 0$ is a constant that can be made arbitrarily small. Assume, for all k, the kth inner code reports an estimate of the outer symbol $\hat{\xi}_k$ together with a reliability weight $\alpha_k \in [0,1]$. Apply Forney's GMD decoding to the outer code [7], the source message can be recovered if the following inequality holds [5, Theorem 3.1b].

$$\sum_{k=1}^{N_o}\alpha_k\mu_k > (r_o+\zeta_0)N_o, \quad (17)$$

where $\mu_k = 1$ if $\hat{\xi}_k = \xi_k$, and $\mu_k = -1$ if $\hat{\xi}_k \neq \xi_k$. Consequently, error probability conditioned on the given $\boldsymbol{z}$ vector is bounded by

$$P_e(R,r_o,\boldsymbol{z}) \le Pr\left\{\sum_{k=1}^{N_o}\alpha_k\mu_k \le (r_o+\zeta_0)N_o\right\} \le \min_{s\ge0}\frac{E\left[\exp\left(-sN_i\sum_{k=1}^{N_o}\alpha_k\mu_k\right)\right]}{\exp(-sN_i(r_o+\zeta_0)N_o)}. \quad (18)$$

where the last inequality is due to Chernoff's bound.

Given the inner codeword lengths $\boldsymbol{z}$, random variables $\alpha_k\mu_k$ for different inner codes are independent. Therefore, (18) can be further written as

$$P_e(R,r_o,\boldsymbol{z}) \le \min_{s\ge0}\frac{\prod_{k=1}^{N_o}E\left[\exp\left(-sN_i\alpha_k\mu_k\right)\right]}{\exp(-sN_i(r_o+\zeta_0)N_o)} = \min_{s\ge0}\frac{\exp\left(\sum_{k=1}^{N_o}\log E\left[\exp\left(-sN_i\alpha_k\mu_k\right)\right]\right)}{\exp(-sN_i(r_o+\zeta_0)N_o)}. \quad (19)$$

Now we will derive the expression of $\log E\left[\exp\left(-sN_i\alpha_k\mu_k\right)\right]$ for the kth inner code.

Assume the normalized effective codeword length is z_k. Given z_k, depending on the received channel symbols, the decoder generates the maximum likelihood outer code estimate $\hat{\xi}_k$, and generates α_k using Forney's algorithm presented in [5, Section 4.2]. Define an adjusted error exponent function $E_z(z)$ as follows.

$$E_z(z) = zE_{FL}\left(\frac{R}{r_oz}, p_X\right). \quad (20)$$

By following Forney's error exponent analysis presented in [5, Section 4.2], we obtain

$$-\log E\left[\exp\left(-sN_i\alpha_k\mu_k\right)\right] = \max[\min\{N_iE_z(z_k), N_i(2E_z(z_k)-s), N_is\}, 0]. \quad (21)$$

Define a function $\phi(z,s)$ as follows,

$$\phi(z,s) = \begin{cases} -s(r_o+\zeta_0) & z, E_z(z) < s/2 \\ 2E_z(z) - (1+r_o+\zeta_0)s & z, s/2 \le E_z(z) < s \\ (1-r_o-\zeta_0)s & z, E_z(z) \ge s \end{cases}. \quad (22)$$

Substitute (21) into (19), we get the expression of the conditional error exponent $E_f(f_Z, R)$ as

$$E_f(f_Z,R) = \max_{p_X, \frac{R}{C_F} \le r_o \le 1, s \ge 0} \int \phi(z,s) f_Z(z) dz. \quad (23)$$

Combining (23) with (15), fountain error exponent of the concatenated code is therefore given by

$$E_{Fc}(R) = \max_{p_X, \frac{R}{C_F} \le r_o \le 1, s \ge 0} \min_{f_Z, \int_0^\infty z f_Z(z)dz = 1} \int \left[\phi(z,s) + \frac{(1-z)^2}{2}\right] f_Z(z) dz. \quad (24)$$

Assume f_Z^* is the inner codeword length density that minimizes $E_{Fc}(R)$ in (24). Assume we can find $0 < \lambda < 1$, and two density functions $f_Z^{(1)}$, $f_Z^{(2)}$, satisfying $\int_0^\infty z f_Z^{(1)}(z)dz = 1$, $\int_0^\infty z f_Z^{(2)}(z)dz = 1$, such that

$$f_Z^* = \lambda f_Z^{(1)} + (1-\lambda) f_Z^{(2)}. \quad (25)$$

It is easily seen that $E_{Fc}(R)$ should be minimized either by $f_Z^{(1)}$ or $f_Z^{(2)}$, which contradicts the assumption that f_Z^* is optimum. In other words, if f_Z^* is indeed optimum, then a decomposition like (25) must not be possible. This implies that f_Z^* can take non-zero values on at most two different z values. Therefore, we can carry out the optimization in (24) only over the following class of f_Z functions, characterized by two variables $0 \le z_0 \le 1$ and $0 \le \gamma \le 1$.

$$f_Z(z) = \gamma\delta(z - z_0) + (1-\gamma)\delta\left(z - \frac{1-z_0\gamma}{1-\gamma}\right). \quad (26)$$

where $\delta()$ is the impulse function.

Now let us fix p_X, r_o, γ, and consider the following optimization of $E_{Fc}(R, p_X, r_o, \gamma)$ over z_0 and s.

$$E_{Fc}(R, p_X, r_o, \gamma) = \min_{0 \le z_0 \le 1} \max_{s \ge 0} \gamma\phi(z_0, s) + (1-\gamma)\phi\left(\frac{1-z_0\gamma}{1-\gamma}, s\right) + \frac{\gamma}{1-\gamma}\frac{(1-z_0)^2}{2}. \quad (27)$$

Since given z_0, $\gamma\phi(z_0,s) + (1-\gamma)\phi\left(\frac{1-z_0\gamma}{1-\gamma}, s\right)$ is a linear function of s, depending on the value of γ, the optimum s^* that maximizes (27) should either satisfy $s^* = E_z(z_0)$ or $s^* = E_z\left(\frac{1-z_0\gamma}{1-\gamma}\right)$.

When $\gamma \ge \frac{1-r_o-\zeta_0}{2}$, we have $s^* = E_z(z_0)$. This yields

$$E_{Fc}(R, p_X, r_o) = \min_{0 \le z_0, \gamma \le 1} \left[\frac{\gamma}{1-\gamma}\frac{(1-z_0)^2}{2} + (1-r_o-\zeta_0)E_z(z_0)\right]. \quad (28)$$

When $\gamma \le \frac{1-r_o-\zeta_0}{2}$, we have $s^* = E_z\left(\frac{1-z_0\gamma}{1-\gamma}\right)$. It can be shown that

$$E_{Fc}(R, p_X, r_o) \ge \min_{0 \le z_0, \gamma \le 1} \left[\frac{\gamma}{1-\gamma}\frac{(1-z_0)^2}{2} + (1-r_o-\zeta_0)E_z\left(1 - \frac{\gamma}{1-\gamma}\frac{1+r_o+\zeta_0}{1-r_o-\zeta_0}(1-z_0)\right)\right]. \quad (29)$$

Both (28) and (29) are minimized at $\gamma = \frac{1-r_o-\zeta_0}{2}$.

Consequently, substitute $\gamma = \frac{1-r_o-\zeta_0}{2}$ into (28), we get

$$E_{Fc}(R, p_X, r_o) = \min_{0 \le z_0 \le 1} \left[\frac{1-r_o-\zeta_0}{1+r_o+\zeta_0}\frac{(1-z_0)^2}{2} + (1-r_o-\zeta_0)E_z(z_0)\right] \quad (30)$$

Minimize (30) over z_0 gives the desired result.

Because the complexity of encoding and decoding the outer code is linear in N_o, if we fix N_i at a large constant and only take N_o to infinity, the overall decoding complexity of the concatenated code is linear in $N = N_iN_o$. The overall encoding complexity is linear in the number of transmitted symbols (given that N_i is fixed). Since fixing N_i causes a reduction of $\zeta_1 > 0$ in the achievable error exponent [7], and both ζ_0, ζ_1 can be made arbitrarily small as we increase N_i, we conclude that fountain error exponent $E_{Fc}(R)$ given in (5) can be *arbitrarily approached* by one-level concatenated fountain codes with linear complexity.

■

REFERENCES

[1] J. Byers, M. Luby, and A. Rege, "A Digital Fountain Approach to Reliable Distribution of Bulk Data," *ACM SIGCOMM,* Vancouver, Canada, Sep. 1998.

[2] M. Luby, "LT codes," *IEEE FOCS,* Vancouver, Canada, Nov. 2002.

[3] A. Shokrollahi, "Raptor Codes," *IEEE Trans. Inform. Theory,* Vol. 52, pp. 2551-2567, Jun. 2006.

[4] S. Shamai, I. Teletar, and S. Verdú, "Fountain Capacity," *IEEE Trans. Inform. Theory,* Vol. 53, pp. 4327-4376, Nov. 2007.

[5] G. Forney, "Concatenated Codes," *The MIT Press,* 1966.

[6] E. Blokh and V. Zyablov, "Linear Concatenated Codes," Nauka, Moscow, 1982 (In Russian).

[7] Z. Wang and J. Luo, "Approaching Blokh-Zyablov Error Exponent with Linear-Time Encodable/Decodable Codes," to appear in *IEEE Comm. Letters.*

[8] V. Guruswami and P. Indyk, "Linear-Time Encodable/Decodable Codes With Near-Optimal Rate," *IEEE Trans. Inform. Theory,* Vol. 51, pp. 3393-3400, Oct. 2005.

[9] R. Gallager, "A Simple Derivation of The Coding Theorem and Some Applications," *IEEE Trans. Inform. Theory,* Vol.11, pp.3-18, Jan. 1965.

[10] A. Barg and G. Forney, "Random Codes: Minimum Distances and Error Exponents," *IEEE Trans. Inform. Theory,* Vol. 48, pp. 2568-2573, Sep. 2002.