International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

Exponential Error Bounds for Block Concatenated Codes with Tail Biting Trellis Inner Codes

Shigeichi Hirasawa† and Masao Kasahara ‡

† Department of Industrial and Management Systems Engineering,
School of Science and Engineering,
Waseda University,
3-4-1 Ohkubo Shinjuku-ku,
Tokyo, 169-8555 JAPAN
E-mail: hirasawa@hirasa.mgmt.waseda.ac.jp

‡ Faculty of Informatics,
Osaka Gakuin University
2-36-1, Kishibe-minami, Suita-shi
Osaka, 564-8511 JAPAN
E-mail: kasahara@utc.osaka-gu.ac.jp

Abstract

Tail biting trellis codes and block concatenated codes are discussed from random coding arguments. An error exponent and decoding complexity for tail biting random trellis codes are shown. Then we propose a block concatenated code constructed by a tail biting trellis inner code and derive an error exponent and decoding complexity for the proposed code. The results obtained by the proposed code show that we can attain a larger error exponent at all rates except for low rates with the same decoding complexity compared with the original concatenated code.
Keywords–block codes, concatenated codes, error exponent, tail biting convolutional codes, decoding complexity, asymptotic results

1. Introduction

A coding theorem obtained from random coding arguments mainly discussed in '70s gives us simple and elegant results on coding schemes, although it states only an existence of a code. Random coding arguments can reveal the essential mechanism with respect to the code. Since we assume maximum likelihood decoding, they can make clear the relationship between the probability of decoding error $\Pr(\epsilon)$ and the decoding complexity $G(N)$ at a given rate R, where N is the code length. It should be noted that the coding theorem can only suggest the behavior of the code, hence it is not useful enough to design an actual code.

On the other hand, turbo codes and their turbo decoding algorithms have been developed, and low density parity check codes and their decoding algorithms have been also redeveloped in '90s. It has been known that the turbo codes combined with turbo decoding have high performance such that they can almost meet the Shannon limit. It is generally difficult to show, however, the performance of them such as the probability of decoding error and the decoding complexity by simple equations without using the weight distribution of the component codes. Therefore it is important to note that we discuss coding schemes from the coding theorem aspects and lead the results obtained into practical coding problems.

Concatenated codes [1] have remarkable and important properties from both theoretical and practical viewpoints. Recently, a block code constructed by a tail biting convolutional code as a component code called parallel concatenated block codes (PCBCs) has been introduced [2]. PCBCs are evaluated by the turbo decoding techniques. Code parameters of the PCBCs that have good performance are tabulated. They are, however, regarded as just one of product type turbo codes. While, the present authors have used a terminated trellis code as an inner code of a generalized version of the block concatenated code called the code $\mathcal{C}^{(J)}$ [3] to reduce the decoding complexity. Note that the block code has an advantage in decoding delay within a constant time.

In this paper, we propose a block concatenated code with a tail biting random trellis code, which is called code $\mathcal{C}_T$. Exponential error bounds and the decoding complexity for the codes $\mathcal{C}_T$ are discussed from random coding arguments. It is shown that the codes $\mathcal{C}_T$ have larger error exponents compared to the original block concatenated codes simply called codes $\mathcal{C}$ whose inner codes are composed of ordinary block codes, or terminated trellis codes at all rates except for low rates under the same decoding complexity.

First, we derive an error exponent and the decoding complexity for tail biting random trellis codes. Next, they are applied to construct the codes $\mathcal{C}_T$, and an error exponent and the decoding complexity of the codes $\mathcal{C}_T$ are derived. Finally, they are also applied to obtain a generalized version of the concatenated code [3].

Throughout this paper, assuming a discrete memoryless channel with capacity C, we discuss the lower bound on the reliability function (usually called the error exponent) and asymptotic decoding complexity measured by the computational work [4].

2. Preliminaries

Let an (N, K) block code over $GF(q)$ be a code of length N, number of information symbols K and rate R, where

$$R = \frac{K}{N}\ln q \quad (K \leq N) \quad \text{[nats/symbol]} \tag{1}$$

From random coding arguments for an ordinary block code, there exists a block code of length N, and rate R for which the probability of decoding error $\Pr(\epsilon)$ and the decoding complexity G satisfy:

$$\Pr(\epsilon) \leq \exp[-NE(R)] \quad (0 \leq R < C) \tag{2}$$

$$G \sim N\exp[NR] \tag{3}$$

where $E(\cdot)$ is (the lower bound on) the block code exponent [6], and the symbol $\sim$ indicates asymptotic equality.

While, let a (u, v, b) trellis code over $GF(q)$ be a code of branch length u, branch constraint length v, yielding b channel symbols per branch and rate r, where

$$r = \frac{1}{b}\ln q \quad \text{[nats/symbol]} \tag{4}$$

Hereafter, we denote $\frac{v}{u}$ by a parameter θ, i.e.,

$$\theta = \frac{v}{u} \quad (0 < \theta \leq 1) \tag{5}$$

We now have three methods for converting a trellis code into a block code [5]:

(i) Direct truncation method
(ii) Tail termination method
(iii) Tail biting method

Letting

$$N = ub \tag{6}$$

for a truncated trellis code of (i) and for a terminated trellis code of (ii), results derived are shown in Table 1, where $e(\cdot)$ is (the lower bound on) the trellis code exponent [6] (see Appendix A).

In Table 1, the rate R of the terminated trellis code is given by [6]:

$$R = \frac{u-v}{u}r = (1-\theta)r \tag{7}$$

Note that the following equation holds between $E(R)$ and $e(r)$ [6]:

$$E(R) = \max_{0<\mu\leq 1}(1-\mu)e\Big(\frac{R}{\mu}\Big) \tag{8}$$

which is called the concatenation construction [6].

3. Tail biting trellis codes

The tail biting method of (iii) is introduced as a powerful converting method for maintaining a larger error exponent with no loss in rates, although the decoding complexity increases. The tail biting method can be stated as follows [5]:

Suppose an encoder of a trellis code. First, initialize the encoder by inputting the last v information (branch) symbols of u information (branch) symbols, and ignore the output of the encoder. Next, input all u information symbols into the encoder, and output the codeword of length $N = ub$ in channel symbols and rate $r = \frac{1}{b}\ln q$. As the result, we have a (ub, u) block code over $GF(q)$ by the tail biting method.

Theorem 1 There exists a block code of length N and rate r obtained by a tail biting random trellis code with $0 < \theta \leq \frac{1}{2}$ for which the probability of decoding error $\Pr(\epsilon)$ satisfies

$$\Pr(\epsilon) \leq \exp[-N\theta e(r)] \quad (0 < \theta \leq \frac{1}{2}, 0 \leq r < C) \tag{9}$$

while the decoding complexity G of the block code is given by:

$$G \sim N^2 q^{2v} = N^2\exp[2N\theta r] \tag{10}$$

Proof: Let $\boldsymbol{w}$ be a message sequence of (branch) length u, where all messages are generated with the equal probability. Rewrite the sequence $\boldsymbol{w}$ by

$$\boldsymbol{w} = (\boldsymbol{w}_{u-v}, \boldsymbol{w}_v) \tag{11}$$

where $\boldsymbol{w}_{u-v}$ is the former part of $\boldsymbol{w}$ (length $u-v$), and $\boldsymbol{w}_v$, the latter part of $\boldsymbol{w}$ (length v). As stated in the tail biting method, first initialize the encoder by inputting $\boldsymbol{w}_v$. Next input $\boldsymbol{w}$ into the encoder. Then output the coded sequence $\boldsymbol{x}$ of length $N = ub$. Note that tail biting random trellis coding requires every channel symbols on every branch be chosen independently at random with the probability $\boldsymbol{p}$ which maximizes $E_0(\rho, \boldsymbol{p})$ on nonpathological channels [6]. Suppose the q^v Viterbi trellis diagrams, each

Table 1: Error exponent and decoding complexity for block codes

block code	error exponent	decoding complexity	upper bound on $\Pr(\cdot)$
ordinary block code	$E(R)$	$N\exp[NR]$	$G^{-\frac{E(R)}{R}}$
truncated trellis code	$E(r)$ [6]	N^2q^v	$G^{-\frac{E(r)}{\theta r}}$
terminated trellis code	$E(R)$ [6]	N^2q^v	$G^{-\frac{1-\theta}{\theta}\frac{E(R)}{R}}$
tail biting trellis code (Theorem 1)	$\theta e(r)$ $(0<\theta\le\frac{1}{2})$	N^2q^{2v}	$G^{-\frac{e(r)}{2r}}$

of which starts at the state $s_i(i=1,2,\cdots,q^v)$ depending on $\boldsymbol{w}_v$, and ends at the same state s_i. The Viterbi decoder generates the maximum likelihood path $\hat{\boldsymbol{w}}(i)$ in the trellis diagram for starting at s_i and ending at s_i. Computing $\max_i \hat{\boldsymbol{w}}(i)=\hat{\boldsymbol{w}}$, the decoder outputs $\hat{\boldsymbol{w}}$. The decoding error occurs when $\{\boldsymbol{w}\neq\hat{\boldsymbol{w}}\}$. Without loss of generality, let the true path $\boldsymbol{w}$ start at s_1 (and end at s_1). The probability of decoding error $\Pr(\epsilon_1)$ within a trellis diagram starting s_1 (and ending s_1) for a (u,v,b) random trellis code is given by [6]

$$\begin{aligned}\Pr(\epsilon_1) &\le uK_1\exp[-vbE_0(\rho)] \quad (0\le\rho\le 1)\\ &=\exp\big\{-N\theta[e(r)-\epsilon]\big\} \quad (0\le r<C) \qquad (12)\end{aligned}$$

where an error event begins at any time. While the probability of decoding error $\Pr(\epsilon_2)$ within trellis diagrams starting at $s_i(i\neq 1, i=2,3,\cdots,q^v)$ and ending at s_i is given by

$$\begin{aligned}\Pr(\epsilon_2) &\le |C|^{\rho}\exp[-ubE_0(\rho)]\\ &=\exp[-ubE_0(\rho)+\rho vbr]\\ &=\exp\big\{-N[E_0(\rho)-\rho\theta r]\big\}\\ &=\exp[-NE(\theta r)] \qquad (13)\end{aligned}$$

where note that the number of trellis diagrams $|C|$ which contain no true path is given by

$$|C| = q^v - 1 \lesssim \exp[vbr] \qquad (14)$$

From (12) and (13), the probability of over-all decoding error $\Pr(\epsilon)$ is bounded by the union bound:

$$\begin{aligned}\Pr(\epsilon) &\le \Pr(\epsilon_1)+\Pr(\epsilon_2)\\ &\le \exp[-N\theta e(r)]+\exp[-NE(\theta r)] \qquad (15)\end{aligned}$$

where $\epsilon=\epsilon_1\cup\epsilon_2$. If $0<\theta\le 1/2$, then $E(\theta r)=\max_\theta(1-\theta)e(r)\ge(1-\theta)e(r)\ge\theta e(r)$. Thus we have from (15)

$$\Pr(\epsilon)\le\exp\big\{-N\theta[e(r)-o(1)]\big\} \qquad (16)$$

$$o(1)=\frac{\ln 2}{N}\to 0 \ \text{ as } \ N\to\infty$$

While the maximum likelihood decoder for the tail biting trellis code requires N^2q^v comparisons for each trellis diagram and performs them in parallel for q^v trellis diagrams, we then have (10), yielding the proof. □

The result derived in Theorem 1 is also shown in Table 1.

Next, we evaluate the probability of decoding error $\Pr(\epsilon)$ by taking into account the decoding complexity G so that coding methods can be easily compared [6].

Let us assume the code length N and rate $R=r$ are the same for all conversion methods. Rewriting $\Pr(\epsilon)$ in terms of G, we have for an ordinary block code, $G\sim N\exp[NR]>\exp[NR]$ from (3)

$$N \gtrsim \frac{1}{R}\ln G \qquad (17)$$

disregarding a lower order terms than N, since we are interested in asymptotic behavior. We then have [6]

$$\Pr(\epsilon)\lesssim G^{-\frac{E(R)}{R}} \qquad (18)$$

Since $G\sim N^2q^{2v}\ge\exp[2vbr]=\exp[2N\theta r]$ holds for the tail biting trellis code, we then have the following corollary.

Corollary 1 For the tail biting trellis code, we have

$$\Pr(\epsilon)\lesssim G^{-\frac{e(r)}{2r}} \qquad (19)$$

Proof: See Appendix B. □

A similar derivation gives the evaluations for truncated trellis code of (i) and for terminated trellis code of (ii) as shown in Table 1 after a little manipulation, where $q^v=\exp[vbr]=\exp[N\theta r]$ holds (See Appendix C).

Example 1 On a very noisy channel, the negative exponent $\frac{e(r)}{2r}$ of G in (19) for a tail biting trellis code is larger than that $\frac{E(R)}{R}$ of G in (18) for an ordinary block code, independent of θ except for rates $0\le R=r\le\frac{C}{4}$. The two negative exponents are shown in Fig. 1 (See Appendix D). □

4. Concatenated codes with tail biting trellis inner codes

First, consider the original concatenated code $\mathcal{C}$ [1] over $GF(q)$ with an ordinary block inner code and a Reed Solomon (RS) outer code.

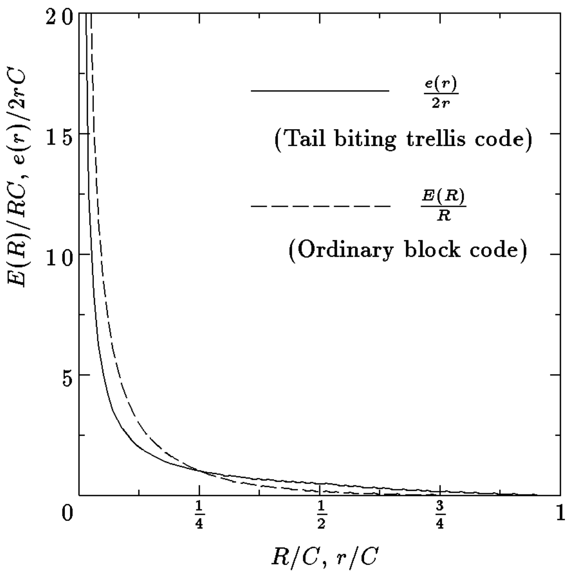

Figure 1: Negative exponents in terms of G for very noisy cannel.

Lemma 1 ([1]) There exists an original concatenated code $\mathcal{C}$ of overall length N_0 and overall rate R_0 in nats/symbol whose probability of decoding error $\Pr(\epsilon)$ is given by

$$\Pr(\epsilon) \le \exp[-N_0 E_C(R_0)] \quad (0 \le R_0 < C) \tag{20}$$

where

$$E_C(R_0) = \max_{0<R<C} \Big(1 - \frac{R_0}{R}\Big) E(R) \tag{21}$$

which is called the concatenation exponent [1]. While the overall decoding complexity G_0 for the code $\mathcal{C}$ is given by at most

$$G_0 = O(N_0^2 \log^2 N_0) \tag{22}$$

where the outer decoder of the RS code performs generalized minimum distance (GMD) decoding. □

Next, let us suppose a block concatenated code $\mathcal{C}_T$ over $GF(q)$ constructed by a (u, v, b) tail biting trellis inner code and an (n, k) RS outer code, where

$$n = q^u \tag{23}$$

holds. From (9), we have the following theorem.

Theorem 2 There exits a block concatenated code $\mathcal{C}_T$ of length $N_0 (= nN)$ and rate R_0 which satisfies

$$\Pr(\epsilon) \le \exp[-N_0 \theta e_C(R_0)] \quad (0 < \theta \le \frac{1}{2}, 0 \le R_0 < C) \tag{24}$$

where

$$e_C(R_0) = \max_{0<r<C} \Big(1 - \frac{R_0}{r}\Big) e(r) \tag{25}$$

Proof: Let the block inner code be the tail biting trellis code of length N and rate r whose average probability of decoding error $\overline{p_\epsilon}$ satisfies

$$\overline{p_\epsilon} \le \exp[-N\theta e(r)] \tag{26}$$

from Theorem 1. Then the over-all probability of decoding error $\Pr(\epsilon)$ is given by

$$Pr(\epsilon) \le \exp\Big[- N_0 \theta \Big(1 - \frac{R_0}{r}\Big) e(r) \Big] \tag{27}$$

where we assume GMD decoding of the RS outer code of length $n (N_0 = nN)$, completing the proof. □

Substitution of $\frac{R_0}{r}$ and r in (25) into μ and $\frac{R}{\mu}$ in (8), respectively, gives the following Corollary.

Corollary 2 From (8) and (25), the relation

$$e_C(R_0) = E(R_0) \tag{28}$$

holds. □

Although the exponent in (24) of the code $\mathcal{C}_T$ is essentially coincides with that of the concatenated code with a convolutional inner code and a block outer code [7], note that the latter is a member of a class of convolutional codes.

Let the decoding complexity of the inner code be denoted by G_I, that of the outer code, by g_O, and that of overall concatenated code $\mathcal{C}_T$, by G_0. Since the maximum likelihood decoder for the tail biting inner code of length N and rate r requires $N^2 \exp[2N\theta r]$ comparisons for each received word of the inner code, and repeats them n times, we have

$$G_I = O(nN^2 \exp[2N\theta r]) \tag{29}$$

On the other hand, for the GMD decoder for the (n, k) RS outer code, we have [3]

$$g_O = O(n^2 \log^4 n) \tag{30}$$

Substituting (23) into (29) and (30), and letting $\max[G_I, g_O] = G_0$, the overall decoding complexity G_0 for the code $\mathcal{C}_T$ is calculated as follows:

Theorem 3 The overall decoding complexity for a block concatenated code $\mathcal{C}_T$ of length N_0 is give by

$$G_0 = O(N_0^2 \log^2 N_0) \quad (0 < \theta \le \frac{1}{2}) \tag{31}$$

Proof: From (29) and (30), we have

$$\begin{aligned} G_0 &= \max[G_I, g_O] \\ &= \max[O(nN^2 \exp[2N\theta r]), O(n^2 \log^4 n)] \\ &= \max[O(n \log^2 n \cdot n^{2\theta}), O(n^2 \log^4 n)] \\ &= \max[O(n^{1+2\theta} \log^2 n), O(n^2 \log^4 n)] \\ &= O(n^2 \log^4 n) \quad (0 < \theta \le \frac{1}{2}) \end{aligned} \tag{32}$$

where we have used (23) and $n = \exp[Nr]$ or $N = O(\log n)$. Since

$$N_0 = nN = O(n \log n) \tag{33}$$

or

$$n = O\Big(\frac{N_0}{\log N_0}\Big) \tag{34}$$

we have (31) from (32) by disregarding the lower order terms than or equal to $\log\log N_0$. □

From Theorem 2, the error exponent $\theta e_c(R_0)$ for the code $\mathcal{C}_T$ is larger than that $E_C(R_0)$ for the code $\mathcal{C}$ at high rates with the same decoding complexity[1] from Lemma 1 and Theorem 3. Especially, the former approaches a half of the block code exponent $\frac{1}{2}E(\cdot)$ as $\theta \to \frac{1}{2}$.

Example 2 The case of $\theta = \frac{1}{2}$ gives the largest error exponent for the code $\mathcal{C}_T$ with the same overall decoding complexity for the code $\mathcal{C}_T$ and for the code $\mathcal{C}$. On a very noisy channel, the error exponent for the code $\mathcal{C}_T$ is larger than that for the code $\mathcal{C}$, except for $0 \le R_0 \le 0.06\,C$. Substitution of (D.1) and (D.2) into (25) and (21), respectively, gives Fig. 2. □

We easily see that the error exponent for the code $\mathcal{C}_T$ is larger than that for code $\mathcal{C}$ at high rates with the same decoding complexity over binary symmetric channels.

5. Generalized version of concatenated codes with tail biting trellis inner codes

A detailed discussion is omitted here, it is obvious that the code $\mathcal{C}_T$ can be applicable to construct a generalized version of concatenated code [3] called a code $\mathcal{C}_T^{(J)}$. A larger error exponent can be obtained by the code $\mathcal{C}_T^{(J)}$. The decoding complexity, however, increases as J increases, although it is still kept in algebraic order of overall length N_0.

[1] It is difficult to clearly state the superiority of the code $\mathcal{C}_T$ in contrast to the discussion given in such as (18) and (19), since we cannot show $\Pr(\epsilon)$ as a function of G in this section. This is because G appears exponential part, and hence asymptotic arguments have no meaning.

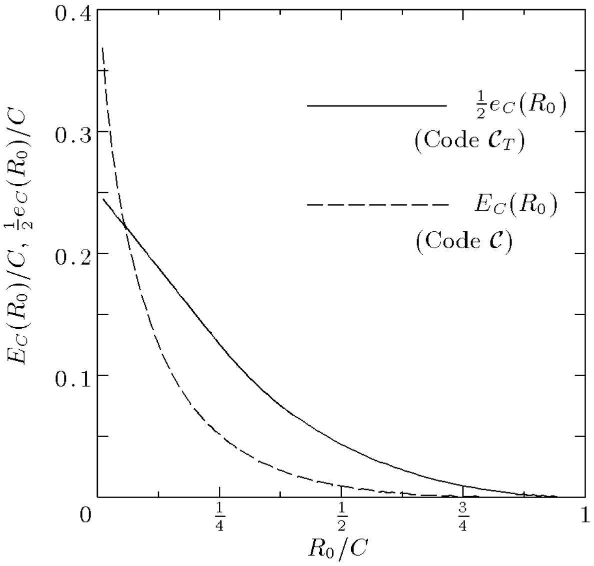

Figure 2: Error exponents for code $\mathcal{C}$ and code $\mathcal{C}_T$ for very noisy cannel.

6. Concluding remarks

We have shown that the error exponents of block codes and block concatenated codes are improved by using tail biting trellis codes at high rates without increasing the decoding complexity. Improvements in both error exponents and the decoding complexity at low rates will be further investigation.

We prefer to discuss the performance obtainable with the proposed code rather than compute in detail that with a particular code. As stated earlier, since the random coding arguments suggest some useful aspects to construct the code, we should note to make them applicable to a practical code, which is also a future work.

Acknowledgement

One of the authors, S. H. would like to thank Mr. T. Sato for his comment.

Appendix A: Derivations of error exponents and decoding complexity for a truncated trellis code and a terminated trellis code in Table 1.

(a) For a truncated trellis code, we have [6]

$$\begin{aligned} \Pr(\epsilon) &\le q^{\rho u} \exp[-ubE_0(\rho)] \\ &= \exp\big\{-N[E_0(\rho) - \rho r]\big\} \quad (0 \le \rho \le 1) \\ &= \exp[-NE(r)] \end{aligned} \tag{A.1}$$

where $q^{\rho u} = \exp[N\rho r]$, since $r = (\frac{1}{b})\ln q$ ($q = \exp[rb]$, $q^{\rho u} = \exp[rb\rho u] = \exp[N\rho r]$), and $E_0(\rho)$ is the Gallager's function. Obviously, the decoder requires q^v comparisons for each step, and repeats them u times. Since these operations are carried out u units logic, we have N^2q^v computational work as the decoding complexity.
(b) For a terminated trellis code, we have [6]

$$\begin{aligned}\Pr(\epsilon) &\le (u-v)K_1 \exp\{-vb[e(r)-\epsilon]\}\\ &\le NK_1 \exp\{-N\theta[e(r)-\epsilon]\}\\ &= NK_1 \exp\{-N[E(R)-\epsilon]\}\end{aligned} \tag{A.2}$$

where

$$\begin{aligned}E(R) &= \max_{0\le\rho\le 1}[E_0(\rho)-\rho R]\\ R &= (1-\theta)r\end{aligned} \tag{A.3}$$

and K_1 is a constant independent of u. Substituting $\theta = 1-\mu$ in (8) and disregarding ϵ in (A.2), we have an error exponent $E(R)$. Similar derivations to (a) gives N^2q^v computational work for the terminated trellis code.

Appendix B: Proof of Corollary 1

From (10), we have

$$G \sim \exp[2N\theta r + 2\ln N] \tag{B.1}$$

Substitution of (B.1) into (9) gives

$$\begin{aligned}\Pr(\epsilon) &\le \exp[-N\theta e(r)]\\ &= G^{\frac{-e(r)}{2r+o(1)}}\end{aligned} \tag{B.2}$$

where

$$o(1) = \frac{2\ln N}{N\theta} \to 0 \quad \text{as} \quad N\to\infty \tag{B.3}$$

Disregarding the term $o(1)$ in (B.2), we complete the proof.

Appendix C: Derivations of the exponents of $\Pr(\epsilon)$ in terms of G

As similar to Appendix B,
(a) For a truncated trellis code, we have

$$\begin{aligned}\Pr(\epsilon) &\le \exp[-NE(r)]\\ &\sim G^{\frac{-E(r)}{\theta r+o(1)}}\end{aligned} \tag{C.1}$$

where

$$o(1) = \frac{2\ln N}{N} \to 0 \quad \text{as} \quad N\to\infty$$

(b) For a terminated trellis code, we have

$$\begin{aligned}\Pr(\epsilon) &\le \exp[-NE(R)]\\ &\sim G^{\frac{-E(R)}{\theta r+o(1)}}\\ &= G^{\frac{-E(R)}{\theta R/(1-\theta)+o(1)}}\end{aligned} \tag{C.2}$$

where

$$o(1) = \frac{2\ln N}{N} \to 0 \quad \text{as} \quad N\to\infty$$

yielding the proof.

Appendix D: The exponents $\frac{e(r)}{2r}$ and $\frac{E(R)}{R}$ of $\Pr(\epsilon)$ in terms of G for a very noisy channel

The error exponent for a very noisy channel is given by [1]

$$e(r) = \begin{cases} \frac{C}{2}, & 0 \le r < \frac{C}{2};\\ C - r, & \frac{C}{2} \le r < C\end{cases} \tag{D.1}$$

$$E(R) = \begin{cases} \frac{C}{2} - R, & 0 \le R < \frac{C}{4};\\ (C^{\frac{1}{2}} - R^{\frac{1}{2}})^2, & \frac{C}{4} \le R < C\end{cases} \tag{D.2}$$

Substitution of (D.1) and (D.2) into (19) and (18), respectively, gives Fig. 1.

References

[1] G. D. Forney, Jr., *Concatenated Codes*, Cambridge, MA. MIT, 1966.

[2] C. Weiß, C. Bettstetter, and S. Riedel, "Code construction and decoding of parallel concatenated tail biting codes," *IEEE Trans. Inform. Theory*, vol.47, no.1, pp.366–386, Jan.2001.

[3] S. Hirasawa, M. Kasahara, Y. Sugiyama, and T. Namekawa, "Certain generalizations of concatenated codes–Exponential error bounds and decoding complexity," *IEEE Trans. Inform. Theory*, vol.TI-26, no.5, pp.527–534, Sept.1980.

[4] J. E. Savage, "The complexity of decoders–Part II: Computational work and decoding time," *IEEE Trans. Inform. Theory*, vol.IT-17, no.1, pp.77–85, Jan.1971.

[5] H. H. Ma, and J. K. Wolf, "On tail biting convolutional codes," *IEEE Trans. Communications*, vol.COM-34, no.2, pp.104–111, Feb.1986.

[6] G. D. Forney, Jr., "Convolutional codes II. Maximum-likelihood decoding," *Inform. Contr.* vol.25, pp.222–266, 1976.

[7] J. L. Ramsey, "Cascaded Tree Codes," *MIT Res. Lab. of Electron.*, Cambridge, MA. Tech. Rep. 478, Sept.1970.