Exponential Error Bounds and Decoding Complexity for Block Codes Constructed by Unit Memory Trellis Codes of Branch Length Two

Shigeichi Hirasawa*, Hideki Yagi†, Manabu Kobayashi‡, and Masao Kasahara*
* Research Institute for Science and Engineering, Waseda University, Tokyo, Japan
† Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo, Japan
‡ Center for Data Science, Waseda University, Tokyo, Japan
* Email: hira@waseda.jp

***Abstract*—Performance of block codes constructed by unit memory (UM) trellis codes is discussed from random coding arguments. There are three methods to obtain block codes from trellis codes, i.e., those of (a) Tail Termination (TT), (b) Direct Truncation (DT), and (c) Tail Biting (TB). In this paper, we derive exponential error bounds and decoding complexity for block codes constructed by the UM trellis codes of branch length two based on the above three methods to uniformly discuss their performance. For the UM trellis codes of branch length two, the error exponent of the tail biting unit memory (TB-UM) trellis codes is shown to be larger than or equal to those of the ordinary block codes, the tail termination unit memory (TT-UM) and the direct truncation unit memory (DT-UM) trellis codes for all rates less than the capacity. Decoding complexity for the TB-UM trellis codes of branch length two exhibits interesting property since their trellis diagrams become simple. Taking into account of the asymptotic decoding complexity, the TB-UM trellis codes are also shown to have a smaller upper bound on the probability of decoding error compared to the ordinary block codes for the same rate with the same decoding complexity.**

I. Introduction

It had been introduced to use unit memory (UM) convolutional codes as a byte oriented code [8]. In early 80's, bounds on free distances and error exponents of the UM trellis codes had been discussed in detail [15]. The UM trellis codes had shown to have larger error exponents compared to ordinary trellis codes especially for low rates. Since we have already obtained powerful construction methods and efficient decoding algorithms for block codes, we can effectively use them as component codes of the UM trellis codes. Therefore, the UM trellis codes have a property combining the advantages of both block codes and trellis codes. Note, however, that decoding delay for the UM trellis codes takes on a probabilistic value. Hence sometimes, it is not tolerable for practical applications.

There are three methods to obtain block codes from trellis codes, i.e., those of (a) Tail Termination (TT), (b) Direct Truncation (DT), and (c) Tail Biting (TB). Especially, (c) TB trellis codes [10] are known to be one of the most powerful codes for converting trellis codes to block codes with no loss in rates. Since the TB trellis codes require an intolerable increase in the decoding complexity, much efforts have been devoted to the studies on suboptimum decoding algorithms [1],[10] or efficient maximum-likelihood decoding algorithms [11],[14],[16]. Unfortunately, however, the decoding complexity of the latter algorithms in worst case is the same as that of the complete maximum-likelihood decoding algorithm [16].

The research leading to this paper was partially supported by MEXT Kakenhi under Grant-in-Aids for Scientific Research (C) No. 23501178.

On the other hand, a coding theorem obtained by classical random coding arguments gives us simple and elegant results on coding schemes, although it states only an existence of a code. Random coding arguments can demonstrate the essential mechanism on coding systems. Since we require complete maximum-likelihood decoding (MLD), the relationship between the probability of decoding error $\Pr(\mathcal{E})$ and the decoding complexity $G(N)$ at a given rate r, or R, can be made clear, where N is the code length. It should be noted that the coding theorem can only suggest the behavior of the code ensemble, hence it is not useful enough to design a practical code.

In this paper, we propose a new block code which is constructed by the unit memory trellis code of branch length two by using the tail biting technique. First we discuss the performance of the block codes constructed by the UM trellis codes based on the above three methods, i.e., the Tail Termination UM (TT-UM) trellis codes, the Direct Truncation UM (DT-UM) trellis codes, and the Tail Biting UM (TB-UM) trellis codes. We have shown that for a given branch length u_0 ($u_0 = 2, 3, \cdots$), the error exponent of the TB-UM trellis codes is shown to be larger than those of the TT-UM and the DT-UM trellis codes for all rates less than the capacity [6]. Taking into account of the decoding complexity $G(N)$, the former is also shown to have a smaller upper bound on the $\Pr(\mathcal{E})$ than the ordinary block codes for the same $r = R$ except for low rates with the same $G(N)$. Especially, this paper discusses for the cases of branch length two ($u_0 = 2$ in [6]), the TB-UM trellis codes performs well compared to the ordinary block codes, the TT-UM trellis coeds and DT-UM trellis codes from the view-point of the upper bound on the probability of decoding error for the same $r = R$ and the same $G(N)$.

Throughout this paper, assuming a discrete memoryless channel with capacity C, we discuss the lower bounds on the reliability function $E(\cdot)$ for block codes and $e(\cdot)$ for trellis codes, and the decoding complexity G measured by the computational work [12]. The probability of decoding error is denoted by $\Pr(\mathcal{E})$, the rate, r or R, the code length, N, and the decoding complexity, $G(N)$.

In Section II, we briefly review on the error exponents of the block codes and the trellis code as preliminaries. Section III describes the results of the UM trellis codes [15]. We derive error exponents and decoding complexity for the TT-UM trellis codes, the DT-UM trellis codes, and the TB-UM trellis codes in Section IV. Section V discusses the results on the exponential error bounds and the asymptotic decoding complexity. Section VI is concluding remarks of this paper.

II. Preliminaries

A. Block Codes

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

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

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 $\Pr(\mathcal{E})$ and the decoding complexity $G(N)$ satisfy

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

and

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

where $E(\cdot)$ is (a lower bound on) the block code exponent [5], and the symbol "$\sim$" indicates asymptotic equality.

B. Trellis Codes

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, which satisfies

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

The probability of decoding error $\Pr(\mathcal{E})$ and the decoding complexity $G(v)$ satisfy [5]

$$\begin{aligned} \Pr(\mathcal{E}) &\leq uK_1\exp[-vbE_0(\rho)] && (0 \leq \rho \leq 1) \quad (5)\\ &\leq \exp\{-vb[e(r) - o(1)]\} && (0 \leq r < C), \quad (6)\end{aligned}$$

and

$$G(v) \sim u^2 q^v = u^2\exp[vbr], \tag{7}$$

where K_1 is a constant independent of v, $o(1) \to 0$ $(v \to \infty)$, and $e(\cdot)$ is (a lower bound on) the trellis code exponent [5] given by

$$e(r) = \begin{cases} E_0(1) & (0 \leq r \leq R_{\text{comp}}) \\ E_0(\rho_r) & (R_{\text{comp}} < r = E_0(\rho_r)/\rho_r < C), \end{cases} \tag{8}$$

where $E_0(\rho)$ is the Gallager's function, and $R_{\text{comp}} = E_0(1)$ is the computational cut-off rate of the channel. Note that the following relation holds between $E(R)$ and $e(r)$:

$$E(R) = \max_{0<\mu\leq 1}(1-\mu)e(R/\mu), \tag{9}$$

which is called the concatenation construction [5]. Similarly, letting

$$\theta = v/u, \tag{10}$$

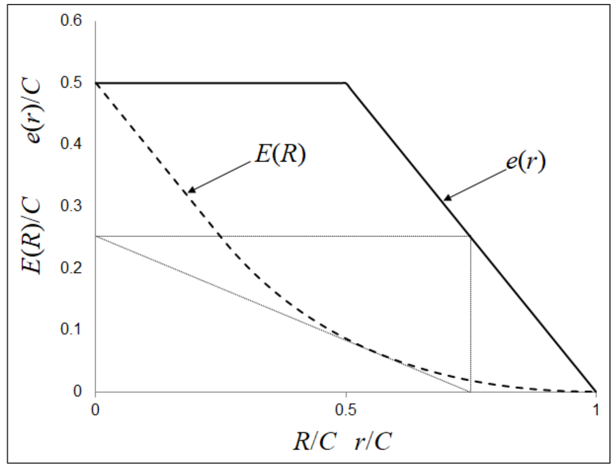


Fig. 1. Concatenation construction of $E(R)$ from $e(r)$ for very noisy channel.

the following equation also holds [5]:

$$e(r) = \min_{0<\theta\leq 1}(1/\theta)E[(1-\theta)r], \tag{11}$$

which is called the inverse concatenation construction [5]. Letting $\mu = R/r$, and $\theta = 1 - \mu$ in (9), we have $E(R) = \max_{r,\theta: R=r(1-\theta)} \theta e(r)$. By this equation, we can construct $E(R)$ from $e(r)$, i.e., $E(R)$ curve is given by the upper envelope of straight lines from $(0, e(r))$ to $(r, 0)$ for all r (See Fig. 1).

III. Unit Memory (UM) Trellis Codes

The UM trellis codes was discussed as the convolutional codes which have largest free distance among all codes of the same rates and a table on free distance obtained by the UM trellis codes was given at short block lengths [8]. And also the UM trellis codes were expected to have superior properties [15].

For the (u, v, b) conventional trellis code discussed in the previous section, letting $v = 1$ and introducing an (n_0, k_0) block code to each branch as a component code, we have a (u, n_0, k_0) unit memory (UM) trellis code. where $u = 2, 3, \cdots$ [15]. Note that we let $n_0 \to \infty$ for the UM trellis codes, while $v \to \infty$ for the conventional trellis codes.

Let a (u, n_0, k_0) unit memory (UM) trellis code over $GF(q)$ be a code of branch length u, length of component code n_0, number of information symbols of component code k_0, and rate r, which satisfies

$$r = (k_0/n_0)\ln q. \quad [\text{nats/symbol}] \tag{12}$$

Lemma 1 (Thommesen & Justesen [15]): There exists a (u, n_0, k_0) unit memory (UM) trellis code for which the probability of decoding error $\Pr(\mathcal{E})$ and the decoding complexity $G(n_0)$ satisfy

$$\Pr(\mathcal{E}) \leq \exp[-n_0 e_{\text{UM}}(r)] \quad (0 \leq r < C), \tag{13}$$

and[1]

$$G(n_0) \sim u^2 n_0 q^{2k_0} = u^2 n_0 \exp[2n_0 r], \tag{14}$$

where

$$e_{\mathrm{UM}}(r) = \begin{cases} 2E(r/2) & (0 \le r \le R_{\mathrm{comp}}) \\ e(r) & (R_{\mathrm{comp}} < r < C), \end{cases} \tag{15}$$

holds. □

We then easily have [15]

$$e_{\mathrm{UM}}(r) = 2E(r/2) > e(r) \quad (0 \le r < R_{\mathrm{comp}}). \tag{16}$$

IV. Block Codes Constructed by Unit Memory (UM) Trellis Codes of Branch Length Two

Let us consider a (u, n_0, k_0) unit memory (UM) trellis code, and we let this code be converted to a block code by using the methods described above. Since we shall compare the conventional (N, K) block code to the block codes constructed by the UM trellis codes, the code length N is chosen as:

$$N = u_0 n_0, \tag{17}$$

denoting

$$0 \le \theta_0 = 1/u_0 \le 1/2, \tag{18}$$

where u_0 is the branch length and is an integer such that $u_0 = 2, 3, \cdots$. The exponents of the above three methods for the general cases of $u_0 = 2, 3, \cdots$ have been derived in [6]. The exponents for the case of branch length two, i.e. $u_0 = 2$ discussed in this paper are easily given by substitution of $u_0 = 2$ into (21), (25), and (29) in [6]. In contrast to the exponents, the decoding complexity of the above three methods for the case of $u_0 = 2$ is not simply given by substitution of it. With careful derivations, we have quite interesting results for the TB-UM trellis codes, since the decoding complexity can be drastically decreased[2].

A. Tail Termination Unit Memory (TT-UM) Trellis Codes

Letting a (u, n_0, k_0) unit memory code be terminated at the 2nd branch level, we have a $(2, n_0, k_0)$ tail termination unit memory (TT-UM) trellis code over $GF(q)$ of length $N = 2n_0$, and rate R. We then easily have the following lemma [5], where actual rate $R = r/2$ and $r = (k_0/n_0)\ln q$.

Lemma 2: The probability of decoding error $\Pr(\mathcal{E})$ and the decoding complexity $G(N)$ for the TT-UM trellis code are given by

$$\Pr(\mathcal{E}) \le \exp[-NE_{\text{TT-UM}}(R)] \quad (0 \le R < C/2), \tag{19}$$

where

$$E_{\text{TT-UM}}(R) = \begin{cases} E_0(1) - R & (0 \le R < R_{\mathrm{comp}}) \\ 0 & (R_{\mathrm{comp}} \le R < C) \end{cases} \tag{20}$$

and

$$G(N) \sim 2Nq^{k_0} = 2N \exp[NR]. \tag{21}$$

□

Note that the inverse concatenation construction of $e_{\mathrm{UM}}(r)$ for $0 < \theta_0 \le 1/2$ gives $E(R)$, while (20) is a straight line from $(0, E_0(1))$ to $(R_{\mathrm{comp}}, 0)$, since $\theta_0 = 1/2$.

B. Direct Truncation Unit Memory (DT-UM) Trellis Codes

Similarly, truncation at the 2nd branch level of a (u, n_0, k_0) unit memory trellis code gives a $(2, n_0, k_0)$ direct truncation unit memory (DT-UM) trellis code over $GF(q)$ of length $N = 2n_0$, and rate r. We then have the following lemma, where there is no loss in rate.

Lemma 3: For the DT-UM trellis code, $\Pr(\mathcal{E})$, and $G(N)$ are given by

$$\Pr(\mathcal{E}) \le \exp[-NE_{\text{DT-UM}}(r)] \quad (0 \le r < C), \tag{22}$$

where

$$E_{\text{DT-UM}}(r) = E(r), \tag{23}$$

and

$$G(N) \sim 2Nq^{2k_0} = 2N \exp[Nr], \tag{24}$$

where $N = 2n_0$, and $r = (k_0/n_0)\ln q$. □

(Proof) See Appendix A.

C. Tail Biting Unit Memory (TB-UM) Trellis Codes

Again consider a $(2, n_0, k_0)$ unit memory (UM) trellis code, and we let this code be converted to a block code by using the tail biting technique.

Theorem 1: There exists a $(2, n_0, k_0)$ tail biting unit memory (TB-UM) trellis code for which the probability of decoding error $\Pr(\mathcal{E})$ and the decoding complexity $G(N)$ satisfy

$$\Pr(\mathcal{E}) \le \exp[-NE_{\text{TB-UM}}(r)] \quad (0 \le r < C), \tag{25}$$

where

$$E_{\text{TB-UM}}(r) = (1/2)e_{\mathrm{UM}}(r) \tag{26}$$

and

$$G(N) \sim 2Nq^{2k_0} = 2N \exp[Nr], \tag{27}$$

where $N = 2n_0$ and $r = (k_0/n_0)\ln q$. □

(Proof) See Appendix B.

V. Discussions

A. Exponential Error Bounds

We have derived error exponents as given by (20), (23), and (26) for the TT-UM, the DT-UM, and the BT-UM trellis codes, respectively. Finally, we give computational results for a very noisy channel.

Example 1: Over a very noisy channel (VNC), the error exponent $E_{\text{TB-UM}}(r)$ for the TB-UM trellis codes is depicted in Fig. 2 together with those $E_{\text{TT-UM}}(R)$ for the TT-UM trellis code and $E_{\text{DT-UM}}(r)$ for the DT-UM trellis code. Note that $E_{\text{DT-UM}}(r)$ coincides with the ordinary block code exponent $E(R)$ by proper choice of the rate of the component code as $r = R$, which gives comparison between block code. □

Over a VNC, we see that the following equation holds:

$$E_{\text{TT-UM}}(R) \le E_{\text{DT-UM}}(r) \le E_{\text{TB-UM}}(r) \quad (0 \le R, r < C). \tag{28}$$

It should be noted that TT-UM trellis codes with $u_0 = 2, 3, \cdots$ give $E(R)$ curve which coincides with $E_{\text{DT-UM}}(r)$

[1]The number of the states of the Viterbi decoder is q^{k_0}, and at each state it compares q^{k_0} survivors. The decoder repeats it u times, where the length of the register is un_0. Then the over-all decoding complexity of the UM trellis code is given by (14).

[2]Strictly speaking, (22) and (30) in [6] hold for the case of $u_0 = 3, 4, \cdots$.

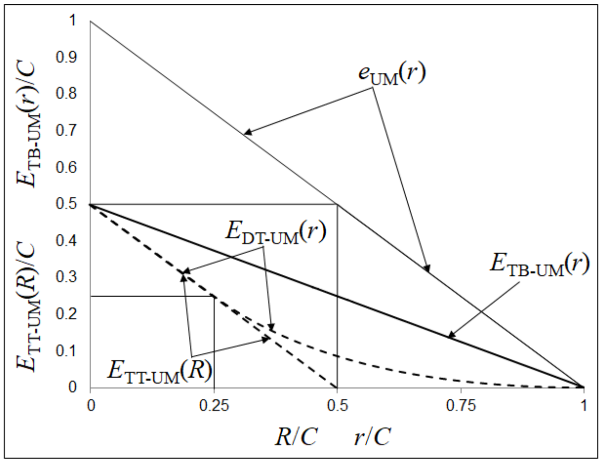


Fig. 2. Error exponents for very noisy channel.

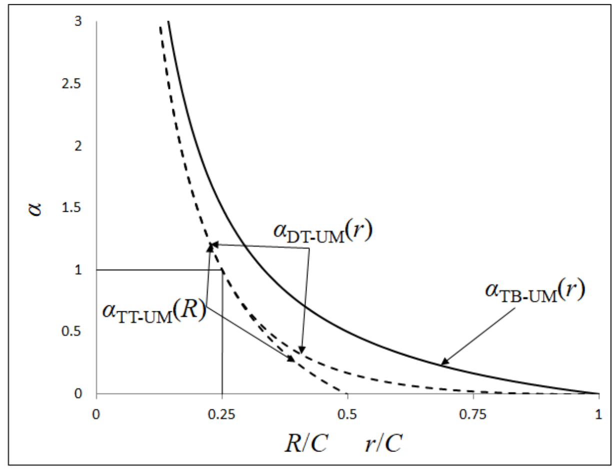


Fig. 3. Decoding complexity for very noisy channel.

for $R = r$ in Fig. 2. Since we are discussing the case of $u_0 = 2$ $(\theta = 1/2)$, $E_{\text{TT-UM}}(R)$ takes on positive values only in the range $0 \leq R < C/2$ for $0 \leq r < C$.

B. Asymptotic Decoding Complexity

Using (21), (24), and (27), we shall discuss the decoding complexity of the block codes constructed by the UM trellis code based on the three methods. Taking into account of (20), (23), and (26), the probability of decoding error $\Pr(\mathcal{E})$ can be rewritten in terms of the decoding complexity G as shown in the following, where we have derived asymptotically N by using (21), (24), and (27), and have substituted them into (19), (22), and (25), respectively, which asymptotically leads a form as $n \to \infty$ such that:

$$\Pr(\mathcal{E}) \lesssim G^{-\alpha(r)}. \tag{29}$$

Theorem 2: The following equations hold asymptotically at the same G:

(a) For the TT-UM trellis codes:

$$\alpha_{\text{TT-UM}}(R) = \begin{cases} [E_0(1) - R]/R & (0 \leq R < R_{\text{comp}}) \\ 0 & (R_{\text{comp}} \leq R < C), \end{cases} \tag{30}$$

(b) for the DT-UM trellis codes:

$$\alpha_{\text{DT-UM}}(r) = E(r)/r \quad (0 \leq r < C), \tag{31}$$

(c) for the TB-UM trellis codes:

$$\alpha_{\text{TB-UM}}(r) = (1/2)e_{\text{UM}}(r)/r \quad (0 \leq r < C). \tag{32}$$

□

(Proof) See Appendix C .

Comparing the ordinary block code with the TB-UM trellis code, we have the following interesting remark.

Remark 1: For the same decoding complexity $G(N)$ and the same rate $r = R$, for all rates less than the capacity over a very noisy channel, the upper bound on the probability of decoding error $\Pr(\mathcal{E})$ for the TB-UM trellis code is asymptotically smaller than or equal to that for the ordinary block code as shown in Fig. 3. □

Since $\alpha_{\text{TB-UM}}(r)$ is larger than $\alpha_{\text{TT-UM}}(R)$ and $\alpha_{\text{DT-UM}}(r)$, the TB-UM trellis codes can achieve a smaller upper bound on the probability of decoding error than the others. However it decreases only algebraically in decoding complexity G.

VI. Concluding Remarks

We have shown that the block codes constructed by the tail biting unit memory trellis codes of branch length two have remarkable properties from both the error exponent and the decoding complexity compared to the ordinary block codes.

In this paper, we have discussed only random coding exponents, a derivation of expurgated exponents is remained as a further research. The minimum distance and the asymptotic distance ratio of the codes discussed here are also remained as further works.

Acknowledgement

One of the authors, S. Hirasawa would like to thank Professor S. Oishi of Waseda University for giving an opportunity to write this paper. He would also thank to all the member of Goto Lab. and Matsushima Lab. at Waseda University.

References

[1] R. V. Cox, and C.-E. W. Sundberg, "An efficient adaptive circular Viterbi algorithm for for decoding generalized tailbiting convolutional codes," *IEEE Trans. Vehicular Technology*, vol.43, no.1, pp.57-68, Feb. 1994.

[2] P. M. Ebert, *Error Bounds for Parallel Communication Channels*, M.I.T. Research Laboratory of Electronics, Cambridge, Mass., Tech. Rept448, 1966.

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

[4] G. D. Forney, Jr., "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inform. Theory*, vol. IT-14, no. 2, pp.206-220, Mar. 1968.

[5] G. D. Forney, Jr., "Convolutional codes II, Maximum-likelihood decoding," *Inform. Contr.*, vol. 25, pp.222-266, 1974.

[6] S. Hirasawa, H. Yagi, M. Kobayashi, and M. Kasahara, "Exponential error bounds for block bodes constructed by unit memory trellis codes," *IEICE, Tech. Rep.* IT2017-33, Chiba, Japan, July 2017.

[7] S. Hirasawa, and M. Kasahara, "Exponential error bounds for tail biting trellis codes," *Proc. 2018 IEICE General Conference*, A-2-2, Tokyo, Japan, March 2018.

[8] L. N. Lee, "Short unit-memory byte-oriented binary convolutional codes have maximal free distance," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 349-352, May 1976.
[9] S. Lin and R. Y. Shao, "General structure and construction of tail biting trellis for linear block codes," *Proc. 2000 IEEE Int. Symp. on Inform. Theory (ISIT)*, p.117, Italy, June 25-30, 2000.
[10] H. H. Ma, and J. K. Wolf, "On tail biting convolutional codes," *IEEE Trans. Comm.*, vol. COM-34, no. 2, pp.104-111, Feb. 1986.
[11] J. Ortin, P. Gaicia, F. Gutierrez, and A. Valdovinos, "A* based algorithm for reduced complexity ML decoding of tailbiting codes, " *IEEE Comm. Letters*, vol. 14, no. 9, pp.854-856, Sept. 2010.
[12] J. E. Savage, "Three measure of decoder complexity," *IBM J. Res. Devlop.*, pp.417-425, July 1970.
[13] C. E. Shannon, R. G. Gallager, and Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Inform. Contr.*, vol. 10, pp.65-103, Jan. 1967.
[14] P. Shanker, P. N. Kumar, K. Sasidharan, and B. S. Rajan, "ML decoding of block codes on their tailbiting trellis," *Proc. 2001 IEEE Int. Symp. on Inform. Theory (ISIT)*, p.291, U.S.A. June 24-29, 2001.
[15] C. Thommesen, and J. Justesen, "Boundes on distances and error exponents of unit memory codes," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 5, pp.637-649, Sept. 1983.
[16] Q. Wang, and V. K. Bhargava, "An efficient maximum likelihood decoding for generalized tail biting convolutional codes including quasicyclic coeds," *IEEE Trans. Comm.*, vol.Com-37, no.8, pp.875-879, Aug. 1989.
[17] C. Weiss, C. Bettstetter, and S. Riedel, "Code construction and decoding of parallel concatenated tailbiting codes," *IEEE Trans. Inform. Theory*, vol.47, no.1, pp.366-386, Jan. 2001.
[18] V. Zyablov, J. Justesen, C. Thommesen and S. Shavgulidze, "Bounds on distances for unit memory concatenated codes," *Problems of Information Transmission*, vol. 32, no. 1, pp.58-69, Jan.-Mar., 1996.

APPENDIX A
PROOF OF LEMMA 3

Followed by (126) of [5], we have

$$\begin{aligned}\Pr(\mathcal{E}) &\leq q^{2k_0\rho}\exp[-2n_0E_0(\rho)]\\ &\leq \exp[-NE(r)]. \end{aligned} \tag{A.1}$$

APPENDIX B
PROOF OF THEOREM 1

Instead of showing the encoding and decoding procedure of (c) the TB-UM trellis code, we give the trellis diagrams of it together with (a) the TT-UM and (b) the DT-UM trellis codes as illustrated in Fig. A.1. In this figure, the states $\boldsymbol{s}_i$ $(i = 1, 2, \cdots, q^{k_0})$ are shown by the nodes. The true path is denoted by[3] $\boldsymbol{x}_1^*(= 0^N)$ with bold solid lines, the maximum likelihood (ML) codeword (path) for the i-th sub-trellis diagram, by $\hat{\boldsymbol{x}}_i$ with solid lines, and possible paths, with dotted lines. Note that the over-all decoded codeword $\hat{\boldsymbol{x}}$ of the TB-UM trellis code is given by:

$$\hat{\boldsymbol{x}} = \arg\max_i \Pr(\boldsymbol{y}|\hat{\boldsymbol{x}}_i) \quad (i = 1, 2, \cdots, q^{k_0}) \tag{A.2}$$

where

$$\hat{\boldsymbol{x}}_i = \arg\max_j \Pr(\boldsymbol{y}|\boldsymbol{x}_{i,j}) \quad (j = 1, 2, \cdots, q^{k_0}), \tag{A.3}$$

$\boldsymbol{y}$ is a received sequence, and $\boldsymbol{x}_{i,j}$ is the j-th path of the i-th sub-trellis diagram.

Let us define the error events $\mathcal{E}_1, \mathcal{E}_2$, and $\mathcal{E}_3$ as follows:

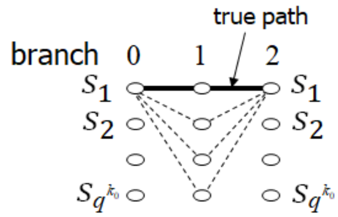
$\mathcal{E}_1$: The error events of the 1st sub-trellis diagram which contain the true path $\boldsymbol{x}_1^*$, where the all possible paths $\boldsymbol{x}_{1,j}$ $(j = 2, 3, \cdots, q^{k_0})$ remerge with $\boldsymbol{x}_1^*$ at the 2nd branch level (which does not contain the event $\mathcal{E}_3$).

$\mathcal{E}_2$: The error events of the all sub-trellis diagrams except for the 1st sub-trellis diagrams (i.e., 2nd, 3rd, $\cdots$, q^{k_0}-th sub-trellis diagrams) which remerge with the true path $\boldsymbol{x}_1^*$ at the 1st branch level (which do not contain the true path, and also do not contain the error event $\mathcal{E}_3$).

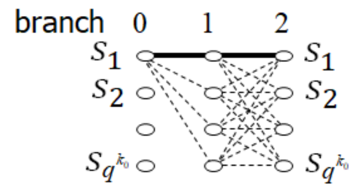
$\mathcal{E}_3$: The error events for which the all possible paths diverge at the 0-th branch level, and never remerge with $\boldsymbol{x}_1^*$ in the all sub-trellis diagrams except for the 1st sub-trellis diagram.

The error events $\mathcal{E}_1$, $\mathcal{E}_2$, and[4] $\mathcal{E}_3$ for the TB-UM trellis code

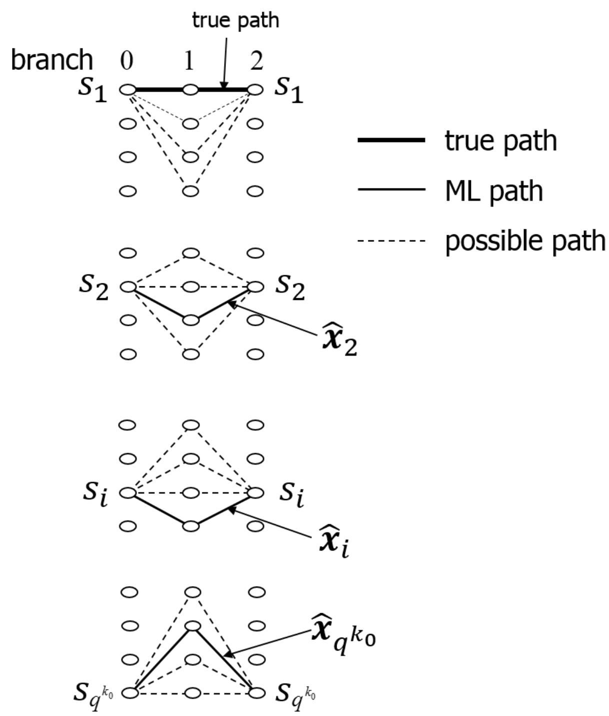
[3]Without loss of generality, we assume the true path is given by all 0 of length $N = 2n_0$, since codes are linear.



(a) TT-UM trellis code.



(b) DT-UM trellis code.



(c) TB-UM trellis code.

Fig. A.1. Trellis diagrams for TT-UM, DT-UM, and TB-UM trellis codes $(q = 2, k_0 = 2)$.

[4]Since the tail biting condition must hold, the state $\boldsymbol{s}_i$ at the 0-th branch level is the same $\boldsymbol{s}_i$ at the 2nd branch level, and the number of possible paths is only $q^{k_0} - 1$ $(= 3$, in the figure) for each subtrellis diagram, where $i = 2, 3, \cdots, q^{k_0}$ for the TB-UM trellis code of branch length two.

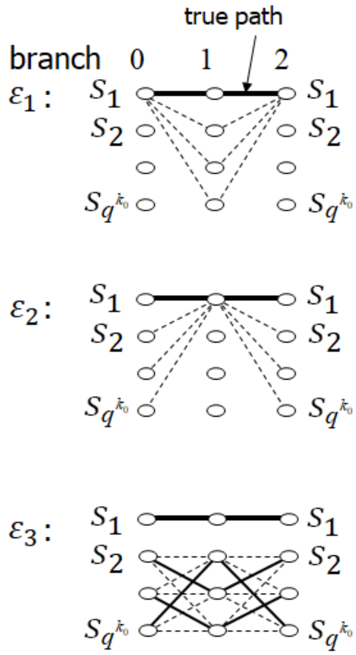


Fig. A.2. Error events $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ for the TB-UM trellis code.

are illustrated in Fig. A.2. Then the probability of decoding error $\Pr(\mathcal{E}_1)$ for the 1st sub-trellis diagram is upper bounded by:

$$\Pr(\mathcal{E}_1) \leq \exp[-N(1/2)e_{\mathrm{UM}}(r)]. \tag{A.4}$$

While assuming the (-1)-th branch level, and cyclically shifting the branch level by -1, we easily see that the probability of decoding error $\Pr(\mathcal{E}_2)$ within the sub-trellis diagrams starting at $\boldsymbol{s}_i$ and ending at $\boldsymbol{s}_i$ $(i \neq 1, i = 2, \cdots, q^{k_0})$ is also bounded by $\Pr(\mathcal{E}_1)$. At the decoding step of computing (A.3), if the event $\mathcal{E}_3$ occurs, the true path $\boldsymbol{x}_1^*$ is never on the list $\mathcal{L}$, which includes only the $q^{k_0} - 1$ $\hat{\boldsymbol{x}}_i$'s as shown in Fig. A.1 (c) and Fig. A.2 $\mathcal{E}_3$. As a result, $\mathcal{E}_3$ can be regarded as list decoding error $\mathcal{L}$ with the list size $|\mathcal{L}| = q^{k_0} - 1$. The probability of list decoding error $\Pr(\mathcal{L})$ with the list size $|\mathcal{L}|$ can be derived as follows [2], [4], [5], [13]:

Lemma A.1 (Ebert [2]): Consider a block code of length N and rate r. Letting $\boldsymbol{x}_m$ be a transmitted codeword, and $\boldsymbol{y}$, a received sequence over a discrete memoryless channel , if

$$\Pr(\boldsymbol{y}|\boldsymbol{x}_m) \leq \Pr(\boldsymbol{y}|\boldsymbol{x}_{m'}) \quad \text{for at least } |\mathcal{L}| \text{ distinct } \boldsymbol{x}_{m'} \quad (m \neq m', m, m' = 1, 2, \cdots, M) \tag{A.5}$$

hold, then the probability of list decoding error $\Pr(\mathcal{L})$ is given by

$$\Pr(\mathcal{L}) \leq \exp[-NE(r')], \tag{A.6}$$

where

$$r' = (1/N)\ln(M/|\mathcal{L}|) \tag{A.7}$$

and $M = \exp[Nr]$. □

Lemma A.2: Let $\boldsymbol{y}$ be a received sequence which never remerges with the true path $\boldsymbol{x}_1^*$. Then there are $\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{x}}_{i'}$'s which satisfy (A.5), by substitution of $\boldsymbol{x}_1^*$ into $\boldsymbol{x}_m$, and $\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{x}}_{i'}$'s into $\boldsymbol{x}_{m'}$, where $i \neq i', i, i' = 2, 3, \cdots, q^{k_0}$, and we have assumed that $\Pr(\boldsymbol{x}_{i,j})$ are uniformly distributed. □

(Proof) There is $\hat{\boldsymbol{x}}_i$ such that $\boldsymbol{y}$ satisfies $\boldsymbol{y} \in \mathcal{R}(\hat{\boldsymbol{x}}_i)$ $(i \neq 1)$, where $\mathcal{R}(\hat{\boldsymbol{x}}_i)$ denotes the decoding region of $\hat{\boldsymbol{x}}_i$ (i.e., $\boldsymbol{y}$ is decoded into $\hat{\boldsymbol{x}}_i$ with maximum-likelihood decoding). Next, assume the best component code whose average probability of decoding error $p_{\mathcal{E}}$ is least and satisfies [3]

$$p_{\mathcal{E}} = \exp\{-N_0[E(r) + o(1)]\}, \tag{A.8}$$

and also assume the equi-error channel between codewords of the component code[5]. The transition probability between distinct codewords is given by

$$p_e = p_{\mathcal{E}}/(q^{k_0} - 1). \tag{A.9}$$

Note that asymptotically as $n_0 \to \infty$,

$$\Pr(\boldsymbol{y} \in \mathcal{R}(\hat{\boldsymbol{x}}_i)|\hat{\boldsymbol{x}}_i) = (1 - p_{\mathcal{E}})^2 \tag{A.10}$$

$$\Pr(\boldsymbol{y} \in \mathcal{R}(\hat{\boldsymbol{x}}_i)|\hat{\boldsymbol{x}}_{i'}) \geq (1 - p_{\mathcal{E}})p_e, \tag{A.11}$$

and

$$\Pr(\boldsymbol{y} \in \mathcal{R}(\hat{\boldsymbol{x}}_i)|\boldsymbol{x}_1^*) = {p_e}^2 \tag{A.12}$$

hold. Since $1 - p_{\mathcal{E}} \simeq 1$, and $p_e \ll 1$, we complete the proof. □

After the above preparations, by letting $|\mathcal{L}| = q^{k_0} - 1 < \exp[Nr/2]$ in (A.7), and hence $r' = r/2$, we have over-all probability of decoding error $\Pr(\mathcal{E})$ by taking the union bound:

$$\begin{aligned}\Pr(\mathcal{E}) &\leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2) + \Pr(\mathcal{E}_3)\\ &\leq 3\exp\{-N\min[(1/2)e_{\mathrm{UM}}(r), E(r/2)]\}\\ &\leq \exp\{-N[(1/2)e_{\mathrm{UM}}(r) - o(1)]\},\end{aligned} \tag{A.13}$$

where $o(1) = (1/N)\ln 3$. From (11), (15), and (16), since $e(r) = \min_\theta(1/\theta)E[(1-\theta)r] \leq (1/\theta)E[(1-\theta)r]$ $(0 < \theta \leq 1)$ is satisfied, we have (A.13). Finally from Fig. A.1 (c), since the paths diverge at 0-th branch level with starting the state $\boldsymbol{s}_i$ and all the paths remerge at the 2nd branch level with the (same ending) state $\boldsymbol{s}_i$ for the i-th sub-trellis diagram $(i = 1, 2, \cdots, q^{k_0})$, there is no need to compare q^{k_0} survivors at the 2nd branch level. The number of the sub-trellis diagrams is q^{k_0}, then the total complexity is $2Nq^{2k_0}$, which leads (27).

Appendix C
Proof of Theorem 2

For the TB-UM trellis codes, we have asymptotically from (27)

$$N \sim (\ln G)/r. \tag{A.14}$$

Substitution of (A.14) into (25) and (27) gives (32). By quite similar manipulation, we have (30) and (31).

[5]The equi-error channel gives the maximum value of $E_0(\cdot)$ function for the super channel with q^{k_0} inputs and q^{k_0} outputs [3].