Transactions Papers

On $(n, n-1)$ Convolutional Codes With Low Trellis Complexity

Hung-Hua Tang and Mao-Chao Lin

***Abstract*—We show that the state complexity profile of a convolutional code C is the same as that of the reciprocal of the dual code of C in case that minimal encoders for both codes are used. Then, we propose an optimum permutation for any given $(n, n-1)$ binary convolutional code that will yield an equivalent code with the lowest state complexity. With this permutation, we are able to find many $(n, n-1)$ binary convolutional codes which are better than punctured convolutional codes of the same code rate and memory size by either lower decoding complexity or better weight spectra.**

***Index Terms*—Convolutional codes, decoding, trellis codes.**

I. Introduction

CONVOLUTIONAL codes are widely used in many digital communication systems for increasing the reliability of transmission due to the fact that convolutional codes have regular trellis structures and hence can be decoded by Viterbi algorithm [1]. For applications which require high coding rates, punctured convolutional codes [2] which are obtained from periodically puncturing some bits from mother codes of low coding rates are usually considered. A punctured convolutional code can be decoded by Viterbi algorithm using the decoding trellis of its mother code. Good punctured convolutional codes have been found by several researchers [2]–[7]. In particular, some of the best known punctured codes can be found in [7]. In this paper, we will show that the punctured convolutional code may not be the best choice if a rate $(n-1)/n$ convolutional code is needed.

Linear block codes can be represented by trellises [8]–[15]. In [8], Forney introduced the minimal trellis construction for the linear block code, which minimizes the number of vertices (states) at each depth, as claimed by Muder [9]. Similar concepts can be extended to the "minimal trellis" [17] of convolutional codes. Minimal trellises for convolutional codes constructed from the parity check matrices and from generator matrices have been, respectively, investigated by Sidorenko and Zyablov [16] and McEliece and Lin [17]. The state spaces of the minimal trellis of a linear block code can be described by its state complexity profile. It [8] has been shown that the state complexity profile of a linear block code and its dual code are identical. In Section III, we define the state complexity profile of the convolutional code in a manner similar to that of a linear block code. We show that the state complexity profiles of a convolutional code and the reciprocal of its dual code are identical if minimal encoders for both codes are used. We also show the relation between minimal trellises of a convolutional code and its reciprocal dual regarding nodes which have branches emanating from them and nodes which have branches entering them.

For a communication system using an error-correcting code, a permutation may be easily applied at the receiver to achieve low trellis complexity regardless of the bit ordering at the transmitter. By applying a permutation to the n bits of each word of an (n, k) convolutional code C, we have an equivalent code of C. Among the $n!$ equivalent codes, there is at least one for which the total number of vertices associated to its minimal trellis module [17] is the least and it is termed as an optimally equivalent code. In Section IV, we propose a method to find the permutation that leads to an optimally equivalent code of an $(n, 1)$ convolutional code. Hence, we are also able to find an optimally equivalent code of an $(n, n-1)$ convolutional code.

In this paper, a good convolutional code is characterized by a large free distance and thin weight spectra and low state complexity. In Section V, we show that an $(n, n-1)$ convolutional code with a trellis of low state complexity will also have small number of branches. With the method derived in Section IV, we are able to find good $(n, n-1)$ convolutional codes with the aid of computer. We provide 6 tables which contain good $(n, n-1)$ convolutional codes for $n = 3$, 4, 5, 6, 7, and 8, respectively. Many codes in these tables are better than the best punctured convolutional codes of the same code rate and memory size by either lower decoding complexity or better weight spectra.

II. Preliminaries

Let $f = \sum_{i=0}^{\infty} f_i D^i$, $f_i \in F$ be a power series over a finite field F in the indeterminate D. In this paper, we only consider the case of $F = GF(2)$. The set of all possible f is the field of Laurent series over F, denoted by $\mathcal{F}$. If v is a power series over F^n in the indeterminate D, then $v \in \mathcal{F}^n$. The coefficients $\underline{v}_i = (v_{i,0}, v_{i,1}, \ldots, v_{i,n-1})$ of the monomials D^i in v are called

Paper approved by N. C. Beaulieu, the Editor for Wireless Communication Theory of the IEEE Communications Society. Manuscript received August 21, 2000; revised July 3, 2001. This work was supported in part by the National Science Council of the R.O.C. under Grant NSC89-2213-E-002-121. This paper was presented in part at the 2000 International Symposium on Information Theory and Its Applications, Honolulu, HI, November 5-8, 2000.

The authors are with the Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: tang@eagle.ee.ntu.edu.tw; mclin@cc.ee.ntu.edu.tw).

Publisher Item Identifier S 0090-6778(01)00510-4.

0090–6778/02$17.00 © 2002 IEEE

the *words* of v. A convolutional code with components in the field of F, can be viewed as a vector space over $\mathcal{F}$. Let g and h be elements of $\mathcal{F}^n$ and let $\underline{g}_i$ and $\underline{h}_i$ be words of g and h, respectively. Let $w(g,h)=\sum_{i=0}^{\infty} w_i D^i$, where $w_i=\sum_{j=0}^{i} \underline{g}_j \cdot \underline{h}_{i-j}$ and $\underline{g}_j \cdot \underline{h}_{i-j}$ is the standard inner product over F, i.e., $\underline{g}_j \cdot \underline{h}_{i-j}=\sum_{l=0}^{n-1} g_{i,l} h_{(i-j),l}$. Then, $w(g,h)$ is an inner product of g and h over $\mathcal{F}$.

Consider an (n,k,ν) convolutional code C that is defined by a $k \times n$ polynomial generator matrix $G(D)=[g_{ij}(D)]$, where $g_{ij}(D)$ is a polynomial over F. $G(D)$ can be regarded as a convolutional encoder [18]. The memory size of the encoder is $\nu_G=\sum_{i=1}^{k} \max_j\{\deg g_{ij}(D)\}$. $G(D)$ is said to be a minimal encoder for C if it has the minimum memory size ν over all encoders realizing the code. For C, there exists a dual code $C^{\perp}$ that is the collection of all $h \in \mathcal{F}^n$ such that $w(g,h)=0$ for all $g \in C$. Let $G(D)$ and $H(D)$ be minimal encoders for C and $C^{\perp}$, respectively. The product of $G(D)$ and $H(D)$ is a zero matrix. Let ν_H be the memory size of $C^{\perp}$. It can be shown that $\nu_G=\nu_H=\nu$ [19]. Let $g_i(D)=(g_{i0}(D), g_{i1}(D), \ldots, g_{i\,n-1}(D))$, that can be viewed as a polynomial over F^n with degree $m_i=\max_j\{\deg g_{ij}(D)\}$. We may write $g_i(D)=\sum_{j=0}^{m_i} \underline{g}_i^j D^j$, where $\underline{g}_i^j$ is an n-tuple over F. A code $\widehat{C}$ is said to be the reciprocal code of C if there is a minimal encoder $\widehat{G}(D)$ for $\widehat{C}$ that is realized by the generators $\widehat{g}_i(D)=\sum_{j=0}^{m_i} \underline{g}_i^{m_i-j} D^j$, $i=1,\ldots,k$. It can be checked that $\widehat{C}^{\perp}=\widehat{C^{\perp}}$ and $\nu_G=\nu_{\widehat{G}}$. Write $G(D)$ as $G(D)=\sum_{j=0}^{m} G_j D^j$, where G_j is a $k \times n$ matrix over F and $m=\max_i\{m_i\}$. The convolutional code C with $G(D)$ can be viewed as a block code over F of semi-infinite length that has a generator matrix G_{scalar} [17] over F

$$G_{scalar}=\begin{pmatrix} G_0 & G_1 & \cdots & G_m & & & \\ & G_0 & G_1 & \cdots & G_m & & \\ & & G_0 & G_1 & \cdots & G_m & \\ & & & & \ddots & & \end{pmatrix}. \quad (1)$$

Let $x=\{x_0, x_1, x_2, \ldots\}$ be a nonzero sequence. According to [14], its *left index*, denoted $L(x)$, is the smallest index i such that $x_i \neq 0$. Similarly, the *right index* of x if exists, denoted $R(x)$ is the largest index i such that $x_i \neq 0$. Whenever $R(x)$ exists, the *span* of x, denoted $Span(x)$, is the discrete interval $[L(x), R(x)]=[L(x), L(x)+1, \ldots, R(x)]$. Otherwise $Span(x)=[L(x), \infty]$. A nonzero sequence x is said to be *active at depth* i if both $i-1$ and i are in $Span(x)$. A generator matrix of a linear block code is said to be a *minimal span generator matrix* (MSGM) [14] if for any two distinct rows x_p and x_q of it, we have $L(x_p) \neq L(x_q)$ which is termed the L-property and $R(x_p) \neq R(x_q)$ which is termed the R-property. For convolutional codes, MSGM can be similarly defined over G_{scalar} or $G(D)$ [17]. For a convolutional code, there are two matrices closely related to $G(D)$ which are very interesting. One is G_0 and the other is G_{end} which is defined as $[(\underline{g}_1^{m_1})^T \cdots (\underline{g}_k^{m_k})^T]^T$. An encoder $G(D)$ (or G_{scalar}) is an MSGM if $G(D)$ is a minimal encoder, for which the associated G_0 is with L-property and the associated G_{end} is with R-property.

III. Characteristics of the Convolutional Code With Minimal Trellis

In this section, we derive characteristics of convolutional codes with minimal trellises which are similar to the counterparts of block codes. Consider an (n,k) linear block code $V \subseteq F^n$. Let $I=\{0,1,\ldots,n-1\}$ be an *index set*. Define [13] $i^-=\{0,\ldots,i-1\}$, $i^+=\{i,\ldots,n-1\}$ and $0^+=n^-=I$ and $0^-=n^+=\phi$ the empty set. The vertices at depth i of a minimal trellis for V form a state space that is isomorphic to the quotient space [13]

$$\Lambda_i(V) \equiv \frac{V}{(V_{i^-} \oplus V_{i^+})} \quad (2)$$

where V_{i^-} and V_{i^+} are the subcodes of V consisting of all the codewords of V for which the components with indices outside i^- and i^+ are zeros, respectively. If an MSGM for V is available, the dimension of $\Lambda_i(V)$, $\dim(\Lambda_i(V))=k-\dim(V_{i^-})-\dim(V_{i^+})$, is in fact the number of rows of an MSGM which are active at depth i.

For a convolutional code C, suppose that the encoder $G(D)$ is an MSGM. Associated with $G(D)$, we can obtain the minimal trellis for C [17]. In general, we are interested in the regular portion of the trellis, for which the state space at depth i is isomorphic to the state space at depth $i+pn$ for any integer p. The regular portion from depth pn to depth $(p+1)n$ is called a trellis module [17]. Now we can check the state spaces of the (n,k,ν) convolutional code C which is generated by a minimal encoder $G(D)$. The amount of $\underline{g}_\ell^j$'s, for $0 \leq j \leq m_\ell$ and $1 \leq \ell \leq k$, is $\sum_{\ell=1}^{k}(m_\ell+1)=k+\nu$. Let $G_{\rm scalar}$ be in form of MSGM. Consider a vertical slice of $G_{\rm scalar}$ that covers the depths $pn, pn+1, \ldots, pn+n$ for a given $p \geq m$. The nontrivial words in the slice are these $(k+\nu)\underline{g}_\ell^j$'s. For any given $\underline{g}_\ell^j$ in the slice, it uniquely corresponds to a row $[\cdots 0\,0\,\underline{g}_\ell^0\,\underline{g}_\ell^1 \cdots \underline{g}_\ell^{m_\ell}\,0\,0 \cdots]$ of $G_{\rm scalar}$, where $\underline{g}_\ell^j$ is the only element located in the slice. Note that $\underline{g}_\ell^j$, $0<j<m_\ell$ is active at all the depths $pn+i$, $0 \leq i \leq n$. But $\underline{g}_\ell^0$ and $\underline{g}_\ell^{m_\ell}$ may not be active at all the depths. At least $\underline{g}_\ell^0$ is inactive at depth pn and $\underline{g}_\ell^{m_\ell}$ is inactive at depth $pn+n$. Since G_0 is composed of all $\underline{g}_\ell^0$'s, there are $\dim(V_{0,i^+}^C)$ rows of G_0 which are inactive at depth $i(mod\ n)$, where V_0^C is the space generated by rows of G_0. Similarly, there are $\dim(V_{\rm end,i^-}^C)$ rows of $G_{\rm end}$ which are inactive at depth $i(mod\,n)$, where $V_{\rm end}^C$ is the space generated by rows of $G_{\rm end}$. The dimension of state space at depth $i(mod\ n)$ for a minimal trellis module of an $(n, k(C), \nu_G)$ code C is

$$s_i(C)=k(C)+\nu_G-\dim(V_{0,i^+}^C)-\dim(V_{{\rm end},i^-}^C). \quad (3)$$

Note that $s_0(C)=\nu_G$ is the dimension of state space for a conventional trellis module. Now we can define the set $\{s_0(C), s_1(C), \ldots, s_{n-1}(C)\}$ as the *state complexity profile* of the convolutional code C.

For a linear block code V and its dual code $V^{\perp}$, it has been shown [8] that $\dim(\Lambda_i(V))=k-\dim(V_{i^-})-\dim(V_{i^+})=\dim(\Lambda_i(V^{\perp}))=n-k-\dim(V_{i^-}^{\perp})-\dim(V_{i^+}^{\perp})$. In the following, we derive a similar result for the convolutional code C. Let $H(D)$ be a minimal encoder for $C^{\perp}$. Let $V_0^{C^{\perp}}$ and $V_{\rm end}^{C^{\perp}}$

be spaces generated by rows of H_0 and H_{end}, respectively. Interesting relations among V_0^C, V_{end}^C, $V_0^{C^\perp}$ and $V_{\text{end}}^{C^\perp}$ are given by the following lemma.

Lemma 1: (a) $V_0^C = \left(V_0^{C^\perp}\right)^\perp$; (b) $V_{\text{end}}^C = \left(V_{\text{end}}^{C^\perp}\right)^\perp$.

Proof: For any rows g and h of $G(D)$ and $H(D)$, respectively, we have $w(g,h) = 0$. Suppose the degrees of g and h over F^n are s and t, respectively. Then we have $\underline{g}_0 \in V_0^C$, $\underline{g}_s \in V_{\text{end}}^C$, $\underline{h}_0 \in V_0^{C^\perp}$ and $\underline{h}_t \in V_{\text{end}}^{C^\perp}$. Since $w(g,h) = 0$, then $f_0 = \underline{g}_0 \cdot \underline{h}_0 = 0$ and $f_{s+t} = \underline{g}_s \cdot \underline{h}_t = 0$. Hence, $V_0^C \subseteq (V_0^{C^\perp})^\perp$ and $V_{\text{end}}^C \subseteq (V_{\text{end}}^{C^\perp})^\perp$. In [18], it has been shown that if $G(D)$ is a minimal encoder, then $rank(G_0) = rank(G_{\text{end}}) = k$. Similarly, we have $rank(H_0) = rank(H_{\text{end}}) = n-k$. This lemma then follows from the fact that $\dim(V_0^C) = \dim((V_0^{C^\perp})^\perp) = k$ and $\dim(V_{end}^C) = \dim((V_{\text{end}}^{C^\perp})^\perp) = k$. □

Suppose V is a subspace of F^n. Let $J \subseteq I = \{0, \ldots, n-1\}$ be any subset of I for which the complementary subset is $I - J$. The *projection* $P_J(V)$ of V is the image of V under the projection operator P_J that is the mapping for which the components with indices in $I - J$ are set to zero and other components remain unchanged. The subspace V_J of V is defined as the intersection of V and $P_J(V)$. The following lemma can be easily verified from some basic concepts of linear algebra [13].

Lemma 2: If V is a k-dimensional subspace of F^n and $J \subseteq I$, then (i) $\dim(V_J) + \dim(P_{I-J}(V)) = k$ and (ii) $\dim(V_J) + \dim(P_J(V^\perp)) = |J|$. □

Theorem 1: The state complexity profiles of C and $\widehat{C^\perp}$ with minimal encoders are identical.

Proof: It follows from Lemmas 1 and 2 that

$$\begin{aligned} k - \dim(V_{0,i^+}^C) - \dim(V_{\text{end},i^-}^C) =& k \\ & - (n - i - \dim(P_{i^+}(V_0^{C^\perp})) \\ & - (i - \dim(P_{i^-}(V_{\text{end}}^{C^\perp})) \\ =& -(n-k) \\ & + \dim(P_{i^+}(V_0^{C^\perp})) \\ & + \dim(P_{i^-}(V_{\text{end}}^{C^\perp})) \\ =& -(n-k) \\ & + (n - k - \dim(V_{0,i^-}^{C^\perp})) \\ & + (n - k - \dim(V_{\text{end},i^+}^{C^\perp})) \\ =& n - k - \dim(V_{0,i^-}^{C^\perp}) \\ & - \dim(V_{\text{end},i^+}^{C^\perp}). \end{aligned}$$

From the fact that $\nu_G = \nu_H = \nu_{\widehat{H}} = \nu$, we have

$$\begin{aligned} s_i(C) =& k(C) + \nu_G - \dim(V_{0,i^+}^C) - \dim(V_{\text{end},i^-}^C) \\ =& k(C^\perp) + \nu_H - \dim(V_{0,i^-}^{C^\perp}) - \dim(V_{\text{end},i^+}^{C^\perp}) \\ =& k(\widehat{C^\perp}) + \nu_{\widehat{H}} - \dim(V_{0,i^+}^{\widehat{C^\perp}}) - \dim(V_{\text{end},i^-}^{\widehat{C^\perp}}) \\ =& s_i(\widehat{C^\perp}). \end{aligned}$$

□

Consider the minimal trellis module for a convolutional code C, which is closely related to its minimal encoder $G(D)$ that is an MSGM. Every branch of the minimal trellis module is labeled by one symbol (binary bit). Nodes of the minimal trellis module at depth i are isomorphic to states of the state space at that depth. A node at depth i from which two branches emanate implies there is an associated information bit triggering an impulse response beginning at depth i. In other words, there is one row in G_{scalar} which has left index i. A node at depth j for which two branches merge implies that there is an impulse response triggered by some information bit which is faded out at depth j. In other words, there is one row in G_{scalar} which has right index j. Hence, each node at depth i has two branches emanating from it if $\dim(V_{0,i^+}^C) = \dim(V_{0,(i+1)^+}^C) + 1$ and has only one branch emanating from it if $\dim(V_{0,i^+}^C) = \dim(V_{0,(i+1)^+}^C)$. Each node at depth i has two branches merging if $\dim(V_{\text{end},i^-}^C) = \dim(V_{\text{end},(i-1)^-}^C) + 1$ and has only one entering branch if $\dim(V_{end,i^-}^C) = \dim(V_{\text{end},(i-1)^-}^C)$. It follows from Theorem 1 that, $\dim(V_{0,i^+}^C) = \dim(V_{0,(i+1)^+}^C) + 1$ implies that $\dim(V_{\text{end},i^-}^{\widehat{C^\perp}}) = \dim(V_{\text{end},(i+1)^-}^{\widehat{C^\perp}})$. Hence, we find the following result.

1) If a convolutional code C has two branches emanating from each node at depth i, then there are no two branches merging at depth $i+1$ for its reciprocal dual code $\widehat{C}^\perp$.

   Similarly, we have the following results.

2) If C does not have two branches emanating from each node at depth i, then there are two branches merging at nodes at depth $i+1$ for $\widehat{C}^\perp$.
3) If C has two branches merging at each node at depth i, then there are no two branches emanating from each node at depth $i-1$ for $\widehat{C}^\perp$.
4) If C does not have two branches merging at nodes at depth i, then there are two branches emanating from each node at depth $i-1$ for $\widehat{C}^\perp$.

IV. Optimally Equivalent $(n,1)$ and $(n,n-1)$ Convolutional Codes

In this section, we derive the state complexity profile for the $(n,1)$ convolutional code. We also show a method to find an optimally equivalent $(n,1)$ convolutional code. From the result of the last section, we can also find an optimally equivalent $(n,n-1)$ code.

Consider a linear subspace V of F^n. Let supp$(V) = \{j : P_{\{j\}}(V) \neq \{0\}, j \in I = \{0,1,\ldots,n-1\}\}$ be the support set of V. For an $(n,1)$ convolutional code C with memory size ν and minimal encoder $g(D) = \sum_{i=0}^{\nu} \underline{g}^i D^i$, it is clear that the state complexity profile is completely determined by $\underline{g}^0$ and $\underline{g}^\nu$. Let $A = \text{supp}(V_{end}^C)$ and $B = \text{supp}(V_0^C)$. Remember that V_{end}^C and V_0^C are the spaces spanned by $\underline{g}^\nu$ and $\underline{g}^0$, respectively. Let $S_A = \{0,1,\ldots,R(\underline{g}^\nu)\}$ and $S_B = \{L(\underline{g}^0),\ldots,n-1\}$, then $A \subseteq S_A$ and $B \subseteq S_B$. Then, we have $s_i(C) = 1 + \nu - \alpha_i - \beta_i$, where

$$\alpha_i = \begin{cases} 1, & \text{if } i > R(\underline{g}^\nu) \\ 0, & \text{otherwise} \end{cases}$$
$$\beta_i = \begin{cases} 1, & \text{if } i \leq L(\underline{g}^0) \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Note that $\alpha_i \leq \alpha_j$ and $\beta_i \geq \beta_j$ for $i < j$. For clarity, we may offset the invariant quantity ν which is the memory size of the

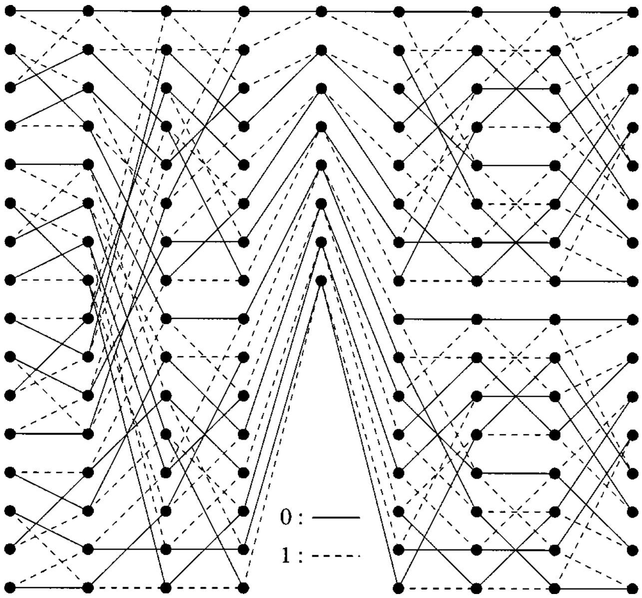

Fig. 1. The minimal trellis module for optimally equivalent code of C, for which the minimal encoder is given in (7).

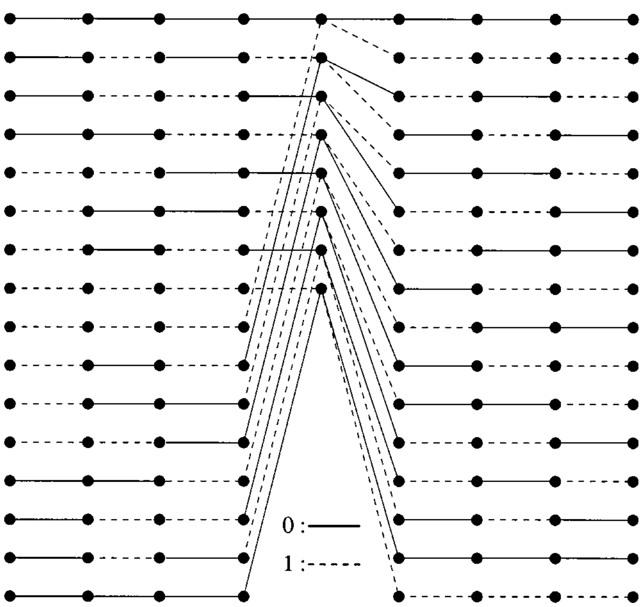

Fig. 2. The minimal trellis module for optimally equivalent code of the reciprocal dual code $\widetilde{C^{\perp}}$.

code C. We term the set $\{\tilde{s}_i\}$, the *offset state complexity profile*, where $\tilde{s}_i = s_i(C) - \nu$. Thus $1 - \alpha_i - \beta_i$ is one of -1, 0 and 1. For an $(n, 1)$ code, $\tilde{s}_i = 1 - \alpha_i - \beta_i$ is in $\{-1, 0, 1\}$. The next two lemmas are derived for the $(n, 1)$ code.

Lemma 3: For an $(n, 1)$ code, the offset $\tilde{s}_i$ is either in $\{0, 1\}$ for all i or $\{0, -1\}$ for all i.

Proof: Suppose that $\tilde{s}_m \neq 0$ for an $m \in I$. In case that $\tilde{s}_m = 1$, then $\alpha_m + \beta_m = 0$ which implies $\alpha_m = \beta_m = 0$. By (4), we have $0 = \alpha_i \leq \alpha_m = 0$ and $1 \geq \beta_i \geq \beta_m = 0$ for $i < m$. Hence, $\alpha_i + \beta_i \leq 1$. Similarly, we also have $\alpha_i + \beta_i \leq 1$ for $i > m$. Therefore $\tilde{s}_i \geq 0$ for all i. In case that $\tilde{s}_m = -1$, then $\alpha_m + \beta_m = 2$ which implies $\alpha_m = \beta_m = 1$. Since $0 \leq \alpha_i \leq \alpha_m = 1$ and $1 = \beta_i \geq \beta_m = 1$ for $i < m$, we have $\alpha_i + \beta_i \geq 1$. Similarly, we have $\alpha_i + \beta_i \geq 1$ for $i > m$. Therefore $\tilde{s}_i \leq 0$ for all i. □

TABLE I
GOOD (3,2) CONVOLUTIONAL CODES

ν	$G(D)$	d_∞	χ	Spectra $\frac{t_1, t_2, \cdots}{f_1, f_2, \cdots}$
3	$\begin{pmatrix} 3 & 1 & 2 \\ 6 & 4 & 3 \end{pmatrix}$	4	-1	$\frac{2, 11, 34, 109, 366, 1244, \cdots}{5, 41, 193, 808, 3299, 13191, \cdots}$
3	$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 7 \end{pmatrix}$	4	0	$\frac{10, 0, 86, 0, 760, 0, \cdots}{26, 0, 432, 0, 5592, 0, \cdots}$
4	$\begin{pmatrix} 1 & 5 & 7 \\ 4 & 6 & 3 \end{pmatrix}$	5	-1	$\frac{4, 17, 54, 192, 681, 2481, \cdots}{8, 69, 313, 1458, 6204, 26581, \cdots}$
4	$\begin{pmatrix} 2 & 7 & 6 \\ 6 & 6 & 1 \end{pmatrix}$	5	0	$\frac{8, 25, 66, 248, 917, 3153, \cdots}{25, 126, 471, 2046, 8872, 35842, \cdots}$
5	$\begin{pmatrix} 5 & 7 & 2 \\ 2 & 10 & 17 \end{pmatrix}$	6	-1	$\frac{13, 0, 180, 0, 2519, 0, \cdots}{43, 0, 1288, 0, 25946, 0, \cdots}$
5	$\begin{pmatrix} 6 & 2 & 7 \\ 16 & 17 & 2 \end{pmatrix}$	6	0	$\frac{18, 0, 219, 0, 3097, 0, \cdots}{60, 0, 1446, 0, 30442, 0, \cdots}$
6	$\begin{pmatrix} 13 & 17 & 5 \\ 12 & 4 & 13 \end{pmatrix}$	6	-1	$\frac{1, 17, 59, 175, 668, 2638, \cdots}{1, 81, 402, 1487, 6793, 31018, \cdots}$
6	$\begin{pmatrix} 10 & 17 & 7 \\ 6 & 14 & 15 \end{pmatrix}$	6	0	$\frac{2, 19, 61, 205, 802, 3019, \cdots}{5, 77, 385, 1733, 7955, 34855, \cdots}$
6	$\begin{pmatrix} 16 & 3 & 6 \\ 16 & 14 & 13 \end{pmatrix}$	6	1	$\frac{6, 27, 70, 285, 1103, 4063, \cdots}{26, 129, 494, 2446, 10878, 46500, \cdots}$
6	$\begin{pmatrix} 2 & 17 & 4 \\ 16 & 2 & 5 \end{pmatrix}$	6	2	$\frac{13, 0, 180, 0, 2519, 0, \cdots}{43, 0, 1288, 0, 25946, 0, \cdots}$
7	$\begin{pmatrix} 13 & 17 & 7 \\ 10 & 32 & 33 \end{pmatrix}$	8	-1	$\frac{60, 0, 649, 0, 10075, 0, \cdots}{314, 0, 5899, 0, 120705, 0, \cdots}$
7	$\begin{pmatrix} 14 & 15 & 7 \\ 2 & 34 & 33 \end{pmatrix}$	8	0	$\frac{75, 0, 810, 0, 12246, 0, \cdots}{422, 0, 7558, 0, 149728, 0, \cdots}$
8	$\begin{pmatrix} 37 & 31 & 10 \\ 16 & 24 & 27 \end{pmatrix}$	8	-1	$\frac{9, 58, 161, 566, 2251, 8668, \cdots}{38, 416, 1404, 5994, 27194, 118184, \cdots}$
8	$\begin{pmatrix} 20 & 23 & 17 \\ 36 & 10 & 31 \end{pmatrix}$	8	0	$\frac{27, 0, 432, 0, 5859, 0, \cdots}{119, 0, 3698, 0, 67041, 0, \cdots}$
8	$\begin{pmatrix} 37 & 31 & 10 \\ 16 & 24 & 27 \end{pmatrix}$	8	1	$\frac{41, 0, 528, 0, 7497, 0, \cdots}{234, 0, 4854, 0, 93342, 0, \cdots}$
8	$\begin{pmatrix} 4 & 33 & 13 \\ 32 & 16 & 17 \end{pmatrix}$	8	2	$\frac{60, 0, 649, 0, 10075, 0, \cdots}{314, 0, 5899, 0, 120705, 0, \cdots}$

Lemma 4: For an $(n, 1)$ code, the nonzero $\tilde{s}_i'$s occur consecutively.

Proof: The offset $\tilde{s}_i$'s are either in $\{0, 1\}$ or $\{0, -1\}$. For $i < m < j$, suppose $\tilde{s}_i = \tilde{s}_j \neq 0$. Then, $\alpha_i + \beta_i = \alpha_j + \beta_j \in \{0, 2\}$. Thus α_i, α_j, β_i and β_j are all equal to γ, where γ is either 0 or 1. Also we have $\alpha_i \leq \alpha_m \leq \alpha_j$, $\beta_i \geq \beta_m \geq \beta_j$. Thus $\alpha_m = \beta_m = \gamma$ and hence $\tilde{s}_m = \tilde{s}_i = \tilde{s}_j$. □

By Theorem 1, we have the following result.

Corollary 1: For an $(n, n-1)$ convolutional code, the nonzero offset $\tilde{s}_i$'s are equal and occur consecutively. □

By applying the same permutation to the n bits of each word of an (n, k) convolutional code C, we have an equivalent code of C. Among all the possible equivalent codes, there is at least one for which the number of vertices associated to its minimal trellis

TABLE II
GOOD (4,3) CONVOLUTIONAL CODES

ν	$G(D)$	d_∞	χ	Spectra $\frac{t_1,t_2,\cdots}{f_1,f_2,\cdots}$
3	$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 3 & 1 \\ 4 & 4 & 0 & 3 \end{pmatrix}$	4	-1	$\frac{29,0,532,0,10146,0,\cdots}{119,0,4404,0,124747,0,\cdots}$
4	$\begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 1 & 1 & 1 \\ 0 & 4 & 6 & 3 \end{pmatrix}$	4	-1	$\frac{3,44,160,638,3558,17222,\cdots}{6,296,1354,6891,47098,263917,\cdots}$
4	$\begin{pmatrix} 2 & 3 & 2 & 2 \\ 0 & 2 & 3 & 1 \\ 6 & 6 & 0 & 1 \end{pmatrix}$	4	0	$\frac{5,39,151,690,3548,16976,\cdots}{12,232,1258,7132,45287,257459,\cdots}$
4	$\begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 2 & 3 & 1 \\ 2 & 0 & 6 & 5 \end{pmatrix}$	4	1	$\frac{5,36,152,708,3439,16510,\cdots}{9,136,1012,6380,38193,220131,\cdots}$
4	$\begin{pmatrix} 2 & 3 & 1 & 1 \\ 2 & 0 & 2 & 3 \\ 4 & 2 & 1 & 3 \end{pmatrix}$	4	2	$\frac{10,46,202,949,4464,21072,\cdots}{31,237,1565,9389,53863,299704,\cdots}$
4	$\begin{pmatrix} 2 & 3 & 1 & 0 \\ 2 & 2 & 2 & 1 \\ 4 & 0 & 3 & 2 \end{pmatrix}$	4	3	$\frac{29,0,532,0,10233,0,\cdots}{119,0,4404,0,125278,0,\cdots}$
5	$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 2 & 2 & 3 & 3 \\ 10 & 12 & 6 & 1 \end{pmatrix}$	5	-1	$\frac{13,67,318,1587,8115,41657,\cdots}{48,449,3012,18634,114018,682461,\cdots}$
5	$\begin{pmatrix} 2 & 2 & 2 & 3 \\ 6 & 7 & 0 & 2 \\ 2 & 4 & 7 & 2 \end{pmatrix}$	5	0	$\frac{15,81,354,1766,9233,46606,\cdots}{59,528,3232,20199,126284,745331,\cdots}$
6	$\begin{pmatrix} 5 & 5 & 3 & 2 \\ 2 & 6 & 1 & 7 \\ 4 & 2 & 6 & 3 \end{pmatrix}$	6	-1	$\frac{45,109,844,3444,20880,100121,\cdots}{259,865,9268,43727,319237,1738085,\cdots}$
6	$\begin{pmatrix} 6 & 5 & 2 & 2 \\ 6 & 2 & 7 & 0 \\ 2 & 2 & 4 & 7 \end{pmatrix}$	6	0	$\frac{65,0,1712,0,45173,0,\cdots}{333,0,17333,0,664262,0,\cdots}$

module is the least and it is termed as an optimally equivalent code. The sum of dimensions of state spaces in the minimal trellis module for an $(n,1)$ code is

$$\sum_{i=0}^{n-1} s_i = n\nu + \sum_{i=0}^{n-1} \tilde{s}_i = n\nu - \chi \tag{5}$$

where $\chi = -\sum_{i=0}^{n-1} \tilde{s}_i$. By Lemma 3, we can easily check that the condition of a smaller total number of vertices associated to a minimal trellis is equivalent to a larger χ. For an $(n,1)$ code, the distribution of $\tilde{s}_i$ is completely determined by the indices $R(\underline{g}^\nu)$ and $L(\underline{g}^0)$. In case $R(\underline{g}^\nu) \leq L(\underline{g}^0)$, the nonzero $\tilde{s}_i$ occur at depths from $R(\underline{g}^\nu)+1$ to $L(\underline{g}^0)$ and hence there are $n + R(\underline{g}^\nu) - L(\underline{g}^0)$ 0's and $L(\underline{g}^0) - R(\underline{g}^\nu)$ -1's in the offset state complexity profile. In case $R(\underline{g}^\nu) \geq L(\underline{g}^0)$, the nonzero $\tilde{s}_i$ occur at depths from $L(\underline{g}^0)+1$ to $R(\underline{g}^\nu)$ and hence there are $R(\underline{g}^\nu) - L(\underline{g}^0)$ 1's and $n - R(\underline{g}^\nu) + L(\underline{g}^0)$ 0's in the offset state complexity profile. Thus, we have $\chi = L(\underline{g}^0) - R(\underline{g}^\nu)$.

In the following, we will show a method to find an optimally equivalent $(n,1)$ convolutional code. From Theorem 1, we can also find an optimally equivalent $(n, n-1)$ convolutional code.

Theorem 2: For an $(n,1)$ convolutional code C, the permutation

$$\pi = \begin{cases} (A-B) \wedge (A \cap B) \wedge (I - A), & \text{if } A \cap B \neq \phi \\ A \wedge (I - (A \cup B)) \wedge B, & \text{if } A \cap B = \phi \end{cases} \tag{6}$$

will result in an optimally equivalent code of C, where $X \wedge Y$ means the concatenation of two ordered sets X and Y.

TABLE III
GOOD (5,4) CONVOLUTIONAL CODES

ν	$G(D)$	d_∞	χ	Spectra $\frac{t_1, t_2, \cdots}{f_1, f_2, \cdots}$
3	$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 2 & 2 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 & 1 \\ 0 & 0 & 2 & 0 & 3 \end{pmatrix}$	3	-1	$\frac{5, 36, 200, 1065, 5893, 32633, \cdots}{12, 192, 1576, 11350, 78542, 520904, \cdots}$
3	$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 \\ 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 & 1 \end{pmatrix}$	3	0	$\frac{10, 51, 256, 1362, 7188, 37783, \cdots}{28, 244, 1776, 12234, 78996, 491738, \cdots}$
4	$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 2 & 2 & 1 & 1 \\ 4 & 4 & 2 & 0 & 1 \end{pmatrix}$	4	-1	$\frac{31, 0, 1254, 0, 46870, 0, \cdots}{118, 0, 11826, 0, 698403, 0, \cdots}$
4	$\begin{pmatrix} 0 & 3 & 1 & 2 & 1 \\ 0 & 2 & 3 & 1 & 0 \\ 2 & 2 & 0 & 1 & 1 \\ 0 & 2 & 2 & 2 & 3 \end{pmatrix}$	4	0	$\frac{30, 126, 815, 4822, 28896, 173230, \cdots}{147, 870, 8024, 58169, 421058, 2939081, \cdots}$
5	$\begin{pmatrix} 1 & 3 & 0 & 2 & 0 \\ 0 & 2 & 3 & 2 & 2 \\ 2 & 2 & 0 & 1 & 1 \\ 2 & 4 & 4 & 2 & 1 \end{pmatrix}$	4	-1	$\frac{3, 55, 305, 1828, 12099, 78349, \cdots}{4, 329, 2785, 23071, 188049, 1454754, \cdots}$
5	$\begin{pmatrix} 0 & 1 & 3 & 3 & 1 \\ 2 & 2 & 1 & 1 & 1 \\ 0 & 2 & 2 & 3 & 3 \\ 6 & 2 & 6 & 0 & 3 \end{pmatrix}$	4	0	$\frac{4, 52, 338, 2022, 12930, 84149, \cdots}{11, 288, 2987, 24232, 192882, 1498975, \cdots}$
5	$\begin{pmatrix} 2 & 1 & 0 & 1 & 1 \\ 2 & 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 3 & 1 \\ 2 & 4 & 6 & 2 & 3 \end{pmatrix}$	4	1	$\frac{7, 56, 376, 2236, 14399, 92556, \cdots}{12, 300, 3364, 25668, 209396, 1604431, \cdots}$
5	$\begin{pmatrix} 3 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 2 & 2 & 0 & 1 & 1 \\ 0 & 2 & 4 & 2 & 1 \end{pmatrix}$	4	2	$\frac{7, 73, 432, 2657, 16795, 106203, \cdots}{15, 449, 4215, 34170, 265852, 1999914, \cdots}$
5	$\begin{pmatrix} 2 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & 3 & 3 \\ 2 & 2 & 2 & 0 & 1 \\ 0 & 4 & 3 & 3 & 0 \end{pmatrix}$	4	3	$\frac{12, 80, 487, 3066, 19308, 121939, \cdots}{48, 562, 5168, 42418, 330042, 2476369, \cdots}$
5	$\begin{pmatrix} 0 & 3 & 1 & 0 & 1 \\ 2 & 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 3 & 1 \\ 0 & 2 & 4 & 2 & 3 \end{pmatrix}$	4	4	$\frac{30, 126, 815, 4822, 29076, 174826, \cdots}{147, 870, 8024, 58169, 422690, 2956367, \cdots}$
6	$\begin{pmatrix} 2 & 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 3 & 1 \\ 7 & 1 & 2 & 0 & 1 \\ 0 & 2 & 4 & 2 & 5 \end{pmatrix}$	5	-1	$\frac{20, 144, 896, 5841, 38536, 254252, \cdots}{89, 1144, 10040, 84059, 671340, 5204314, \cdots}$
6	$\begin{pmatrix} 3 & 1 & 1 & 1 & 0 \\ 0 & 2 & 3 & 1 & 1 \\ 0 & 0 & 2 & 7 & 7 \\ 2 & 0 & 4 & 6 & 3 \end{pmatrix}$	5	0	$\frac{23, 172, 1030, 6661, 44049, 290819, \cdots}{108, 1430, 11873, 97399, 774682, 5984578, \cdot\cdot}$

Proof: In case that $A \cap B \neq \phi$, then $R(\sigma \underline{g}^\nu) \geq L(\sigma \underline{g}^0)$ for any n-column permutation σ. We now need to find a permutation σ which leads to the largest χ or equivalently the least $R(\sigma \underline{g}^\nu) - L(\sigma \underline{g}^0)$. For $i \in A \cap B$, we have $L(\sigma \underline{g}^0) \leq i \leq R(\sigma \underline{g}^\nu)$. Thus for any σ, $R(\sigma \underline{g}^\nu) - L(\sigma \underline{g}^0) + 1 \geq |A \cap B|$, where the equality holds if σ is the permutation π given in (6).

Consider the case of $A \cap B = \phi$. Suppose that $S_A \cap S_B \neq \phi$, that implies $0 \leq p = L(\underline{g}^0) < q = R(\underline{g}^\nu) < n$. Clearly,

TABLE IV
GOOD (6,5) CONVOLUTIONAL CODES

ν	$G(D)$	d_∞	χ	Spectra $\frac{t_1,t_2,\cdots}{f_1,f_2,\cdots}$
3	$\begin{pmatrix} 1&1&0&1&0&0\\ 0&1&1&0&1&0\\ 2&0&0&1&0&1\\ 2&2&2&0&1&1\\ 0&0&0&0&2&3 \end{pmatrix}$	3	-1	$\frac{15, 96, 601, 3903, 25325, 164252, \cdots}{57, 642, 5875, 49414, 393375, 3025294, \cdots}$
4	$\begin{pmatrix} 1&1&1&1&0&0\\ 0&0&3&1&3&1\\ 2&2&0&1&0&1\\ 2&0&2&0&1&1\\ 0&0&2&0&2&3 \end{pmatrix}$	4	-1	$\frac{111, 0, 5628, 0, 291251, 0, \cdots}{742, 0, 72998, 0, 5560874, 0, \cdots}$
5	$\begin{pmatrix} 3&1&0&0&1&1\\ 0&1&3&2&1&1\\ 0&0&2&3&2&2\\ 2&2&2&0&1&0\\ 2&0&0&2&2&1 \end{pmatrix}$	4	-1	$\frac{13, 149, 1064, 8175, 64293, 503549, \cdots}{51, 1215, 13272, 133688, 1291188, 12025722, \cdots}$
5	$\begin{pmatrix} 3&1&2&3&1&1\\ 2&0&1&1&1&0\\ 2&2&2&1&0&1\\ 2&0&0&2&3&0\\ 2&0&0&0&2&3 \end{pmatrix}$	4	0	$\frac{14, 162, 1156, 8770, 68534, 532965, \cdots}{47, 1285, 13598, 134004, 1286677, 11876579, \cdots}$
5	$\begin{pmatrix} 3&1&2&2&1&1\\ 2&0&1&1&1&0\\ 0&0&0&1&3&2\\ 2&2&0&0&1&1\\ 2&0&0&2&2&1 \end{pmatrix}$	4	1	$\frac{19, 160, 1186, 9132, 70707, 548046, \cdots}{48, 1122, 13012, 130414, 1256010, 11642215, \cdots}$
5	$\begin{pmatrix} 0&3&1&1&1&0\\ 2&0&1&1&0&1\\ 0&2&2&3&3&2\\ 0&0&2&2&1&1\\ 2&0&0&2&2&1 \end{pmatrix}$	4	2	$\frac{20, 197, 1372, 10457, 81152, 624719, \cdots}{65, 1540, 16069, 158685, 1511817, 13836949, \cdots}$
5	$\begin{pmatrix} 1&1&1&0&1&1\\ 2&0&1&1&0&1\\ 2&2&0&0&1&1\\ 0&2&2&2&0&1\\ 6&4&4&3&2&0 \end{pmatrix}$	4	3	$\frac{25, 225, 1576, 11964, 91773, 701498, \cdots}{135, 2026, 20913, 203992, 1902278, 17158181, \cdots}$
5	$\begin{pmatrix} 2&1&0&1&1&0\\ 0&2&1&1&1&0\\ 0&0&0&1&3&2\\ 0&2&0&2&1&1\\ 2&2&2&0&0&1 \end{pmatrix}$	4	4	$\frac{64, 0, 3724, 0, 211636, 0, \cdots}{304, 0, 47699, 0, 4364981, 0, \cdots}$

$R(\underline{g}^\nu)$ is in A but not in B and $L(\underline{g}^0)$ is in B but not in A. Let σ' be the permutation that switches the $L(\underline{g}^0)$-th and the $R(\underline{g}^\nu)$-th columns. Then, $R(\sigma'\underline{g}^\nu) < R(\underline{g}^\nu)$ and $L(\sigma'\underline{g}^0) > L(\underline{g}^0)$. With this σ', the resultant offset state complexity profile $(\tilde{s}'_0, \tilde{s}'_1, \ldots, \tilde{s}'_{n-1})$ satisfies $\tilde{s}'_q \leq \tilde{s}_q - 1$, $\tilde{s}'_{p+1} \leq \tilde{s}_{p+1} - 1$ and $\tilde{s}'_m \leq \tilde{s}_m$ for m other than q and $p+1$. As long as $L(\sigma'\underline{g}^0) < R(\sigma'\underline{g}^\nu)$, similar processes can be continued. Since n is finite, these processes can not last forever. Therefore we may assume $R(\sigma\underline{g}^\nu) < L(\sigma\underline{g}^0)$. A permutation σ which leads to the largest χ or equivalently the largest number of $L(\sigma\underline{g}^0) - R(\sigma\underline{g}^\nu) \; - 1$'s is preferred. Since $R(\sigma\underline{g}^\nu) \geq |A| - 1$ and $L(\sigma\underline{g}^0) \leq n - |B|$, then $L(\sigma\underline{g}^0) - R(\sigma\underline{g}^\nu) \leq n + 1 - |A| - |B| = |I - (A \cup B)| + 1$, where the equality holds if σ is the permutation π given in (6). □

Example: Consider an (8,7,4) convolutional code C with minimal encoder

$$G(D) = \begin{pmatrix} 1&1&1&1&0&0&0&0\\ 0&0&0&0&1&1&1&1\\ 0&0&1&1&1&1&0&0\\ 1&0&0&1&0&1&1&0\\ 0&0&2&0&2&0&1&1\\ 0&1&2&0&0&0&0&1\\ 2&0&4&4&0&0&0&1 \end{pmatrix}. \quad (7)$$

The minimal encoder $H(D)$ for $C^\perp$ is (26,32,3,17,5,11,20,34). The components in both $G(D)$ and $H(D)$ are represented in octal form. The state complexity profile of either C or

TABLE V
GOOD (7,6) CONVOLUTIONAL CODES

ν	$G(D)$	d_∞	χ	Spectra $\frac{t_1,t_2,\cdots}{f_1,f_2,\cdots}$
4	$\begin{pmatrix} 1&1&0&0&1&1&0\\ 0&0&1&0&1&1&1\\ 2&2&0&1&0&1&1\\ 2&0&2&0&1&0&0\\ 0&2&0&2&0&1&0\\ 2&0&0&0&2&0&1 \end{pmatrix}$	3	-1	$\frac{6,77,696,5884,51037,443875,\cdots}{11,592,8326,96478,1059653,11139839,\cdots}$
4	$\begin{pmatrix} 0&1&0&1&1&1&0\\ 0&0&1&0&1&1&1\\ 2&2&0&1&1&0&0\\ 0&0&2&2&1&1&0\\ 0&0&0&2&2&1&0\\ 2&0&2&0&0&0&1 \end{pmatrix}$	3	0	$\frac{8,93,809,6790,57905,494013,\cdots}{18,700,9452,110428,1198626,12398355,\cdots}$
4	$\begin{pmatrix} 0&1&1&0&1&1&0\\ 0&0&1&1&1&0&1\\ 0&0&2&3&1&0&1\\ 2&0&0&0&1&1&1\\ 0&2&2&0&0&1&0\\ 0&0&0&2&2&0&1 \end{pmatrix}$	3	1	$\frac{12,124,1024,8307,68623,567653,\cdots}{33,869,11490,127116,1322381,13192157,\cdots}$
4	$\begin{pmatrix} 0&1&1&1&0&0&0\\ 0&0&0&1&1&1&1\\ 2&0&1&0&1&1&0\\ 0&0&0&2&3&1&1\\ 2&2&0&0&0&1&1\\ 0&0&0&0&2&2&1 \end{pmatrix}$	3	2	$\frac{12,124,1024,8295,68643,569209,\cdots}{35,909,12344,138419,1449338,14533952,\cdots}$

$\widehat{C^\perp}$ is $\{4,5,5,5,5,5,4,4\}$. The associated $V_{\rm end}^{\widehat{C^\perp}}$ and $V_0^{\widehat{C^\perp}}$ are $\mathrm{span}\{(00\,111\,100)\}$ and $\mathrm{span}\{(11\,000\,011)\}$, respectively. For $\widehat{C^\perp}$, we have $A = \mathrm{supp}(V_{\rm end}^{\widehat{C^\perp}}) = \{2,3,4,5\}$ and $B = \mathrm{supp}(V_0^{\widehat{C^\perp}}) = \{0,1,6,7\}$. Hence an optimal permutation for the reciprocal dual code $\widehat{C^\perp}$ is $\pi = (2,3,4,5,0,1,6,7)$. According to Theorem 1, π is also an optimal permutation for C. The state complexity profile for the equivalent code of either C or $\widehat{C^\perp}$ is $\{4,4,4,4,3,4,4,4\}$ that is better than that of the original code. The minimal trellis modules for such optimally equivalent codes of C and $\widehat{C^\perp}$ are shown in Figs. 1 and 2, respectively.

V. GOOD $(n, n-1)$ CONVOLUTIONAL CODES

In applying the Viterbi algorithm (VA) to decoding a convolutional code, the decoding complexity is sometimes measured by the number of vertices per minimal trellis module or sometimes by the number of branches per minimal trellis module, where each branch corresponds to one code bit. In Section IV, we have already shown that for an $(n, n-1)$ convolutional code C in case the decoding complexity is measured by the number of vertices, then a large χ is desired. In the following, we consider the case that the decoding complexity is measured by the number of branches. For an $(n, n-1)$ convolutional code C with memory size ν, it has a reciprocal dual $\widehat{C}^\perp$, which is an $(n,1)$ code with indices $L(\widehat{\underline{h}}^0)$ and $R(\widehat{\underline{h}}^\nu)$. For $\widehat{C}^\perp$, there is only a single branch emanating from each state at depth i for $i \neq L(\widehat{\underline{h}}^0)$ and there are two branches emanating from each state at depth i for $i = L(\widehat{\underline{h}}^0)$, while there are two branches merging at each state at depth j for $j = R(\widehat{h}^\nu) + 1$ and there is only a single branch entering each state at depth j for $j \neq R(\widehat{h}^\nu) + 1$. According to the observation given at the end of Section III, for C, there are two branches emananating from each state at depth i for $i \neq R(\widehat{h}^\nu)$ and there is only a single branch emanating from each state at depth i for $i = R(\widehat{h}^\nu)$. Let $\tilde{s}_i$, $i = 0,1,\ldots,n-1$ be the offset state complexity profile of $\widehat{C}^\perp$. Suppose that $\chi = L(\widehat{h}^0) - R(\widehat{h}^\nu) \geq 0$. Then, $\tilde{s}_i = 0$ for $i \leq R(\widehat{h}^0)$ and $i > L(\widehat{h}^0)$ while $\tilde{s}_i = -1$ for $R(\widehat{h}^\nu) < i \leq L(\widehat{h}^0)$. Hence, the decoding complexity can be calculated to be

$$(n-\chi-1)\cdot 2^\nu \cdot 2 + 2^\nu + \chi \cdot 2^{\nu-1} \cdot 2 = (n-\chi)\cdot 2^{\nu+1} + (\chi-1)\cdot 2^\nu. \tag{8}$$

TABLE VI
GOOD (8,7) CONVOLUTIONAL CODES

ν	$G(D)$	d_∞	χ	Spectra $\frac{t_1,t_2,\cdots}{f_1,f_2,\cdots}$
4	$\begin{pmatrix} 1&1&1&1&1&1&0&0\\ 0&0&1&0&1&1&1&0\\ 0&0&0&1&1&0&1&1\\ 2&0&2&0&1&0&1&1\\ 2&2&0&0&0&1&0&1\\ 2&0&2&2&2&0&1&1\\ 2&0&0&2&0&0&0&1 \end{pmatrix}$	3	-1	$\frac{12, 154, 1546, 14954, 148172, 1466039, \cdots}{36, 1505, 22441, 294546, 3677773, 43811679, \cdots}$
4	$\begin{pmatrix} 1&1&1&0&0&0&0&0\\ 0&0&1&1&0&0&0&1\\ 0&0&0&1&1&1&1&0\\ 0&0&2&2&1&0&0&1\\ 2&0&0&0&0&1&0&1\\ 0&0&0&0&2&2&1&0\\ 0&0&0&2&2&0&0&1 \end{pmatrix}$	3	0	$\frac{18, 191, 1839, 17686, 171339, 1658803, \cdots}{62, 1505, 22201, 289290, 3525673, 41093580, \cdots}$
4	$\begin{pmatrix} 0&1&1&1&1&0&0&0\\ 0&0&1&1&0&0&1&1\\ 0&0&0&1&1&1&1&0\\ 0&2&2&0&1&1&0&0\\ 2&0&0&0&0&1&1&1\\ 0&0&0&0&2&2&1&1\\ 0&0&2&2&0&0&0&1 \end{pmatrix}$	3	1	$\frac{28, 274, 2456, 22664, 209734, 1939381, \cdots}{138, 2224, 29024, 346352, 3928042, 43028988, \cdots}$
4	$\begin{pmatrix} 0&1&1&0&1&0&0&0\\ 0&0&1&0&1&0&1&1\\ 0&0&0&1&1&1&1&0\\ 0&2&0&0&1&1&0&0\\ 2&0&0&0&0&1&1&1\\ 0&0&2&2&0&0&1&1\\ 0&0&2&0&0&2&0&1 \end{pmatrix}$	3	2	$\frac{28, 274, 2456, 22524, 208336, 1925595, \cdots}{135, 2384, 31320, 374521, 4259847, 46749804, \cdots}$
4	$\begin{pmatrix} 0&1&0&1&0&1&0&0\\ 0&0&1&1&1&1&1&0\\ 0&0&0&1&1&0&1&1\\ 0&2&0&0&1&1&0&0\\ 2&0&0&0&0&1&1&1\\ 0&0&0&2&2&0&1&0\\ 0&2&2&0&0&0&0&1 \end{pmatrix}$	3	3	$\frac{28, 274, 2456, 22524, 208588, 1930273, \cdots}{134, 2408, 31736, 380074, 4329590, 47586300, \cdots}$
4	$\begin{pmatrix} 1&1&0&1&1&0&0&0\\ 0&0&0&1&0&1&1&0\\ 0&0&0&0&1&0&1&1\\ 2&0&1&0&0&1&0&0\\ 2&2&0&0&0&1&1&0\\ 2&0&2&0&0&0&1&0\\ 2&2&2&2&0&0&0&1 \end{pmatrix}$	3	4	$\frac{28, 274, 2456, 22440, 207066, 1910623, \cdots}{146, 2648, 35040, 419448, 4767794, 52283292, \cdots}$
4	$\begin{pmatrix} 0&1&1&0&0&1&1&0\\ 0&0&1&1&1&0&0&0\\ 0&0&0&1&0&0&1&1\\ 2&2&2&0&1&1&1&0\\ 2&0&0&0&0&1&1&1\\ 2&2&0&0&0&0&1&0\\ 2&0&0&2&0&0&0&1 \end{pmatrix}$	3	5	$\frac{28, 274, 2456, 22356, 205012, 1880307, \cdots}{147, 2616, 34552, 412127, 4660995, 50837934, \cdots}$

Suppose that $\chi = L(\widehat{h}^0) - R(\widehat{h}^\nu) < 0$. Then, $\tilde{s}_i = 0$ for $i > R(\widehat{h}^\nu)$ and $i \leq L(\widehat{h}^0)$ while $\tilde{s}_i = 1$ for $L(\widehat{h}^0) < i \leq R(\widehat{h}^\nu)$. Hence, the decoding complexity can be calculated to be

$$(n+\chi)\cdot 2^\nu \cdot 2 + (-\chi-1)\cdot 2^{\nu+1}\cdot 2 + 2^{\nu+1} = (n-\chi-1)\cdot 2^{\nu+1}. \quad (9)$$

We see that a large χ is desired for both measures of decoding complexity. Hence, χ is a good measure of decoding complexity for an $(n, n-1)$ code.

If we decode an $(n, n-1)$ binary convolutional code with memory size ν by applying VA to the conventional trellis, the decoding complexity measured by the number of branches

TABLE VII
SOME OF THE BEST KNOWN $(n, n-1)$ PUNCTURED CODES SHOWN IN [7]

n	ν	generators*	d_∞	Spectra $(t_1, t_2, \cdots)/(f_1, f_2, \cdots)$
3	3^l	15,(11,17)	4	$(2, 11, 34, 109, 366, 1244, \cdots)/(5, 41, 193, 808, 3299, 13191, \cdots)$
3	4^l	25,(23,35)	5	$(5, 18, 54, 193, 714, 2578, \cdots)/(15, 88, 370, 1640, 7116, 29942, \cdots)$
3	5^l	65,(57,75)	6	$(16, 0, 182, 0, 2700, 0, \cdots)/(56, 0, 1301, 0, 27620, 0, \cdots)$
3	6^c	147,(135,147)	6	$(1, 17, 59, 175, 668, 2638, \cdots)/(1, 81, 402, 1487, 6793, 31018, \cdots)$
3	7^c	337,(251,337)	8	$(66, 0, 706, 0, 10727, 0, \cdots)/(395, 0, 6695, 0, 135288, 0, \cdots)$
3	8^b	625,(577,711)	8	$(9, 58, 161, 566, 2251, 8668, \cdots)/(38, 416, 1404, 5994, 27194, 118184, \cdots)$
4	3^c	15,17,(17,15)	4	$(29, 0, 532, 0, 10059, 0, \cdots)/(124, 0, 4504, 0, 125991, 0, \cdots)$
4	4^l	31,31,(35,23)	4	$(5, 42, 134, 662, 3643, 16585, \cdots)/(10, 290, 1188, 7174, 48976, 262074, \cdots)$
4	5^l	65,65,(47,61)	5	$(13, 71, 326, 1626, 8320, 42351, \cdots)/(51, 474, 2978, 18918, 116366, 690938, \cdots)$
4	6^l	117,173,(165,127)	6	$(45, 109, 844, 3444, 20880, 100121, \cdots)/(276, 843, 9588, 44046, 326876, 1756787, \cdots)$
5	3^l	15,11,11,(11,17)	3	$(6, 32, 185, 1030, 5745, 32204, \cdots)/(11, 184, 1627, 12094, 85568, 578261, \cdots)$
5	4^b	27,33,27,(37,05)	4	$(30, 126, 815, 4822, 29046, 174460, \cdots)/(159, 990, 9076, 66149, 482470, 3378525, \cdots)$
5	5^l	75,75,71,(67,41)	4	$(4, 46, 295, 1832, 11910, 76572, \cdots)/(11, 297, 2876, 23759, 192413, 1475175, \cdots)$
5	6^l	145,113,153,(113,145)	5	$(22, 146, 920, 5983, 39409, 260246, \cdots)/(99, 1184, 10987, 89453, 708470, 5487822, \cdots)$
6	3^b	17,13,13,15,(13,17)	3	$(15, 96, 601, 3963, 26039, 170868, \cdots)/(61, 686, 6257, 53004, 426069, 3309892, \cdots)$
6	4^l	25,27,31,37,(31,35)	4	$(111, 0, 5628, 0, 291695, 0, \cdots)/(754, 0, 74393, 0, 5682302, 0, \cdots)$
6	5^l	73,47,75,67,(73,51)	4	$(15, 138, 993, 7841, 61322, \cdots)/(58, 1090, 11475, 119518, 1158329, \cdots)$
7	4^l	31,35,35,21,21,(33,25)	3	$(6, 77, 696, 5884, 51187, \cdots)/(11, 592, 8326, 96478, 1060924, \cdots)$
8	4^b	21,27,37,23,27,21,(35,23)	3	$(12, 154, 1546, 14978, 149512, \cdots)/(37, 1480, 22199, 290717, 3640770, \cdots)$

[c] Code found by Cain [2].
[l] Code found by Lee [3].
[b] Code found by Bocharova [7].
* The numbers in parentheses correspond to the generator polynomials providing outputs transmitted on the same branch of the trellis.

(each branch corresponds to 1 code bit) per trellis module is $n \cdot 2^{n+\nu-1}$. For an $(n, n-1)$ binary punctured code with memory size ν obtained from a certain convolutional code, the decoding complexity measured by the number of branches per trellis module (each branch corresponds to 1 code bit) is $n \cdot 2^{1+\nu}$. The advantage of decoding for $(n, n-1)$ punctured convolutional code over $(n, n-1)$ convolutional code using conventional trellis is clear. However, we note that compared to an $(n, n-1)$ punctured convolutional code for $\chi = -1$, an $(n, n-1)$ convolutional code with minimal trellis module has the same decoding complexity for $\chi = -1$ and has lower decoding complexity for $\chi > -1$ in case the decoding complexity measured by the number of branches per trellis module and each branch corresponds to one code bit.

For a coding system, low decoding complexity as well as low error rate is desired. The error performance of an (n, k) convolutional code with free distance d_∞ can be estimated by its code weight spectrum and information weight spectrum which are, respectively, represented by t_i and f_i, where t_i is the total number of code sequences with weight $d_\infty + i - 1$ and f_i is the total number of information bits associated to the code sequences with weight $d_\infty + i - 1$. In case the code is applied over a symmetric and memoryless channel and maximum-likelihood decoding is used, the first event error probability of the coding system can be estimated by $P_f \leq \sum_{i=1}^{\infty} t_i P_e(d_\infty + i - 1)$ and the symbol error probability can be estimated by $P_s \leq 1/k \sum_{i=1}^{\infty} f_i P_e(d_\infty + i - 1)$, where $P_e(d)$ is the probability of erroneously decoding a code sequence into a give code sequence which is separated by a distance of d. We see that for achieving low error rate, small t_i and f_i or "thin" weight spectra are desired. In [20], an algorithm called FAST algorithm is proposed to efficiently compute the weight spectra of convolutional codes.

With the aid of Theorem 1, 2, FAST algorithm and computer, we are able to search for good $(n, n-1)$ convolutional codes. For the given n, ν and $\chi \geq -1$, we search for $(n, n-1)$ codes which have the currently best weight spectra. For the given n, ν and χ, we exhaustively check all the possible $(n, 1)$ codes. For each $(n, 1)$ code to be checked, we randomly choose a generator matrix with LR-property for the associated reciprocal dual $(n, n-1)$ code and compute the associated weight spectra. The codes with the currently best weight spectra found in this search are listed in Tables I–VI for $n = 3, 4, 5, 6, 7$, and 8, respectively. For comparison, we also list some of the best known punctured codes in Table VII. We can see many codes in these tables are better than punctured convolutional codes of the same code rate and memory size [2], [3], [7] by either lower decoding complexity or better weight spectra. Note that we only randomly choose one of the many possible generator matrices with LR-property for a given $(n, n-1)$ code. It is likely that there exist better $(n, n-1)$ codes if we exhaustively checking all the possible generator matrices.

VI. Concluding Remarks

In this paper, we show that the state complexity profiles of a convolutional code and the reciprocal of its dual code are identical if minimal encoders for both codes are used. We also propose an optimum permutation for any given $(n, n-1)$ binary convolutional code that will yield an equivalent code with the lowest state complexity. Moreover, we find many good $(n, n-1)$ convolutional codes which are superior to the popular punctured convolutional codes by either lower decoding complexity or better weight spectra. The code search used here is not complete. Hence, it is likely that there exist codes better than those found in this code search. In fact, how to design a method to efficiently check the possible encoders under the restriction of LR-property for the reciprocal of the dual code of an $(n, 1)$ code is an interesting problem.

References

[1] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 260–269, Apr. 1967.

[2] J. B. Cain, G. C. Clark Jr, and J. M. Geist, "Punctured convolutional codes of rate $(n, n-1)$ and simplified maximum likelihood decoding," *IEEE Trans. Inform. Theory*, vol. 25, pp. 97–100, Jan. 1979.

[3] P. J. Lee, "Constructions of rate $(n-1)/n$ punctured convolutional codes with minimal required SNR criterion," *IEEE Trans. Commun.*, vol. 36, pp. 1171–1173, Oct. 1988.

[4] D. Haccoun and G. Begin, "High rate punctured convolutional codes for Viterbi and sequential decoding," *IEEE Trans. Commun.*, vol. 37, pp. 1113–1125, Nov. 1989.

[5] G. Begin and D. Haccoun, "High rate punctured convolutional codes: Structure properties and construction construction technique," *IEEE Trans. Commun.*, vol. 37, pp. 1381–1385, Dec. 1989.

[6] M.-G. Kim, "On systematic punctured convolutional codes," *IEEE Trans. Commun.*, vol. 45, pp. 133–139, Feb. 1997.

[7] I. E. Bocharova and B. D. Kudryashov, "Rational rate punctured convolutional codes for soft-decision Viterbi decoding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1305–1313, July 1997.

[8] G. D. Forney Jr, "Coset codes—Part II: Binary lattices and related codes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1152–1187, Sept. 1988.

[9] D. J. Muder, "Minimal trellises for block codes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1049–1053, Sept. 1988.

[10] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[11] J. L. Massey, "Foundation and methods of channel encoding," in *Proc. Int. Conf. Information Theory and Systems*, vol. 65, Berlin, Germany, 1978, pp. 148–157.

[12] A. D. Kot and C. Leung, "On the construction and dimensionality of linear block code trellises," in *IEEE Int. Symp. Inform. Theory*, San Antonio, TX, 1993.

[13] G. D. Forney Jr, "Dimension/length profiles and trellis complexity of linear block codes," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1741–1752, Nov. 1994.

[14] R. J. McEliece, "On the BCJR trellis for linear block codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1072–1092, July 1996.

[15] A. Vardy and F. R. Kschischang, "Proof of a conjecture of McEliece regarding the expansion index of the minimal trellis," *IEEE Trans. Inform. Theory*, vol. 42, pp. 2027–2034, Nov. 1996.

[16] V. Sidorenko and V. Zyablov, "Decoding of convolutional codes using a syndrome trellis," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1663–1666, Sept. 1994.

[17] R. J. McEliece and W. Lin, "The trellis complexity of convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1855–1864, Nov. 1996.

[18] P. Piret, *Convolutional Codes*. Cambridge, MA: MIT Press, 1988.

[19] G. D. Forney Jr, "Convolutional codes I: Algebraic structure," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 720–738, Nov. 1970.

[20] M. Cedervall and R. Johannesson, "A fast algorithm for computing distance spectrum of convolutional codes," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1146–1159, Nov. 1989.

Hung-Hua Tang was born in Taipei, Taiwan, R.O.C., in 1962. He received the B.S. degree in control engineering from Chiao Tung University, R.O.C., in 1993 and the M.S. and Ph.D. degrees, both in electrical engineering, from National Taiwan University,R.O.C., in 1993 and 2001, respectively. His research interests include coding theory and wireless communication.

Mao-Chao Lin was born in Taipei, Taiwan, R.O.C., on December 24, 1954. He recieved the Bachelor and Masters degrees, both in electrical engineering, from National Taiwan University,R.O.C., in 1977 and 1979, respectively, and the Ph.D. degree in electrical engineering from University of Hawaii in 1986.

From 1979 to 1982, he was an Assistant Scientist of Chung-Shan Institute of Science and Technology at Lung-Tan, Taiwan, R.O.C. He is currently with National Taiwan University, as a Professor with the Department of Electrical Engineering. Since 2000, he has been the chairman of Graduate Institute of Communication Engineering, National Taiwan University. His research interests include coding theory and ARQ techniques.