Characteristic Matrices and Trellis Reduction for Tail-Biting Convolutional Codes

Masato Tajima, Senior Member, IEEE

Abstract

Basic properties of a characteristic matrix for a tail-biting convolutional code are investigated. A tail-biting convolutional code can be regarded as a linear block code. Since the corresponding scalar generator matrix $G_{tb}$ has a kind of cyclic structure, an associated characteristic matrix also has a cyclic structure, from which basic properties of a characteristic matrix are obtained. Next, using the derived results, we discuss the possibility of trellis reduction for a given tail-biting convolutional code. There are cases where we can find a scalar generator matrix $G'$ equivalent to $G_{tb}$ based on a characteristic matrix. In this case, if the polynomial generator matrix corresponding to $G'$ has been reduced, or can be reduced by using appropriate transformations, then trellis reduction for the original tail-biting convolutional code is realized. In many cases, the polynomial generator matrix corresponding to $G'$ has a monomial factor in some column and is reduced by dividing the column by the factor. Note that this transformation corresponds to cyclically shifting the associated code subsequence (a tail-biting path is regarded as a code sequence) to the left. Thus if we allow partial cyclic shifts of a tail-biting path, then trellis reduction is accomplished.

Index Terms

Tail-biting convolutional codes, tail-biting trellis, characteristic matrix, cyclic shift, trellis reduction.

M. Tajima was with the graduate School of Science and Engineering, University of Toyama, 3190 Gofuku, Toyama 930-8555, Japan (e-mail: masatotjm@kind.ocn.ne.jp).

Manuscript received April 19, 2005; revised August 26, 2015. This paper was presented in part at the IEICE Technical Committee IT Conference in March 2017.
Characteristic Matrices and Trellis Reduction for Tail-Biting Convolutional Codes

I. INTRODUCTION

From the 1980s to 1990s, trellis representations of linear block codes were studied with a great interest \[2, 7, 8, 14, 17, 18, 19\]. Subsequently, tail-biting trellises of linear block codes have received much attention. Given a linear block code, there exists a unique minimal conventional trellis. This trellis simultaneously minimizes all measures of trellis complexity. However, tail-biting trellises do not have such a property. That is, minimality of tail-biting trellises depends on the measure being used \[13\]. In general, the complexity of a tail-biting trellis may be much lower than that of the minimal conventional trellis. There have been many contributions to the subject, including \[3, 4, 9, 10, 11, 13, 15, 20, 22, 32\]. The works \[3, 13\] had a strong influence on the subsequent studies. A remarkable progress has been made by Koetter and Vardy in their paper \[13\]. They showed that for a \(k\)-dimensional linear block code of length \(n\) with full support, there exists a list of \(n\) characteristic generators (i.e., a characteristic matrix \[13\]) from which all minimal tail-biting minimal trellises can be obtained. A different method of producing tail-biting trellises was proposed by Nori and Shankar \[20\]. They used the Bahl-Cocke-Jelinek-Raviv (BCJR) construction \[2\]. These works were further investigated by Gluesing-Luerssen and Weaver \[9, 10\]. In particular, noting that a characteristic matrix for a given code is not necessarily unique, they have refined and generalized the previous works. More recent works \[4, 11\] provide further research on the subject.

On the other hand, tail-biting convolutional codes were proposed by Ma and Wolf in 1986 \[16\] (tail-biting representations of block codes were introduced by Solomon and van Tilborg \[24\]). Tail-biting (abbreviated TB) is a technique by which a convolutional code can be used to construct a block code without any loss of rate. In connection with the subject, there have been also many works, including \[11, 16, 23, 25, 28, 33\]. Since a TB convolutional code is identified with a linear block code, the results on TB trellises for linear block codes can be used. In particular, we can think of a characteristic matrix of a given TB convolutional code. In this paper, we first investigate a characteristic matrix for a TB convolutional code. And then, based on the derived results, we discuss the possibility of trellis reduction for a given TB convolutional code. An outline of the rest of the paper is as follows:

In Section II, we review the basic notions needed for this paper.

In Section III, we investigate the basic properties of a characteristic matrix for a TB convolutional code. When a TB convolutional code with generator matrix \(G(D)\) is regarded as a linear block code \(C\), a (scalar) generator matrix (denoted by \(G^{tb}\)) for \(C\) is constructed using the coefficients which appear in the polynomial expansion of \(G(D)\). We see that \(G^{tb}\) has a kind of cyclic structure. Then it is shown that the (characteristic) span list associated with a characteristic matrix for \(C\) consists of some basic spans and their right cyclic shifts, from which basic properties of a characteristic matrix are derived.

In Section IV, we deal with transformations of \(G(D)\) and discuss the relationship between these transformations and the corresponding scalar generator matrices \(G^{tb}\)'s. We see that dividing a column of \(G(D)\) by a monomial factor corresponds to cyclically shifting a column subsequence of \(G^{tb}\) to the left, whereas multiplying a column of \(G(D)\) by a monomial corresponds to cyclically shifting a column subsequence of \(G^{tb}\) to the right. These properties are essentially used for trellis reduction to be discussed in Section V.

In Section V, we discuss the possibility of trellis reduction for a given TB convolutional code (we identify the code with an \((n, k)\) block code \(C\)). As is stated above, we can think of a characteristic matrix for \(C\). Consider the case where some \(k\) characteristic generators, which consist of some basic generators and their right cyclic shifts, can generate the same code \(C\). We see that these characteristic generators form a (scalar) generator matrix associated with a (polynomial) generator matrix of another convolutional code. In this case, if the constraint length of the obtained generator matrix is smaller than that of the original one, then trellis reduction is realized. Even if this kind of reduction is not possible, there are cases where a newly obtained generator matrix contains a monomial factor in some column. Then there is a possibility that the generator matrix is reduced by sweeping the monomial factor out of the column. Note that this operation corresponds to cyclically shifting the corresponding code subsequence to the left. In this way, trellis reduction can be accomplished. We also present a trellis reduction method for high rate codes which uses a reciprocal dual encoder. We remark that the (trellis) section length is an important parameter and the proposed method is restricted to TB convolutional codes with short to moderate section length. We give an upper bound for the section length by evaluating the span lengths of characteristic generators.

Finally, conclusions are provided in Section VI.

II. PRELIMINARIES

We begin with the basic notions needed in this paper, where the underlying field is assumed to be \(F = GF(2)\). Let \(C\) be an \((n, k)\) linear block code, where the set of indices for a codeword in \(C\) is denoted by \(I \triangleq \{0, 1, \cdots, n - 1\}\). Then a codeword \(x \in C\) is expressed as \(x = (x_0, x_1, \cdots, x_{n-1})\). \(I\) is also regarded as the time axis for TB trellises for \(C\). Since TB trellises...
for $C$ are considered in this paper, it is convenient to identify $I$ with $Z_n$, the ring of integers modulo $n$. Hence, when dealing with TB trellises, all index arithmetic will be implicitly performed modulo $n$ [13].

The notion of span is fundamental in trellis theory. Given a codeword $x \in C$, a span of $x$, denoted by $[a, b]$, is a semiopen interval $(a, b) \in I$ such that the corresponding closed interval $[a, b]$ contains all the nonzero positions of $x$ [13]. Due to the cyclic structure of the time axis $I$, we adopt the following interpretation of intervals [9], [10], [13]. For $a, b \in I$, we define

$$[a, b] \triangleq \{a, a + 1, \ldots, b\},$$

if $a \leq b$

$$\{a, a + 1, \ldots, n - 1, 0, 1, \ldots, b\},$$

if $a > b$

(1)

and $(a, b) \triangleq [a, b] \setminus \{a\}$. We call the intervals $(a, b)$ and $[a, b]$ conventional if $a \leq b$ and circular otherwise.

In connection with the construction of minimal TB trellises for $C$, Koetter and Vardy [13] introduced the notion of characteristic generator for $C$. Denote by $\sigma_j(\cdot)$ a cyclic shift to the left by $j$ positions [13]. Similarly, denote by $\rho_j(\cdot)$ a cyclic shift to the right by $j$ positions. Let $X_j^*$ be a basis in minimal-span form [17] for the code $C_j \triangleq \sigma_j(C)$. A characteristic generator for $C$ is a pair consisting of a codeword $x = (x_0, x_1, \ldots, x_{n-1}) \in C$ and a span $[x] = (a, b)$ such that $x_a, x_b$ are nonzero. The set of all the characteristic generators for $C$ is given by

$$X \triangleq X_0^* \cup \rho_1(X_1^*) \cup \cdots \cup \rho_{n-1}(X_{n-1}^*).$$

(2)

Here we have an understanding that if $x^* \in X_j^*$, then $[\rho_j(x^*)] = (a + j, b + j)$, where $[x^*] = (a, b)$.

Assume that $C$ has a full support. Then a characteristic matrix for $C$ is the $n \times n$ matrix having the elements of $X$ as its rows. The above definition implies that when we refer to a characteristic matrix, the associated spans are taken into account. Here note that a basis in minimal-span form is not necessarily unique. Hence, $X$ may not be uniquely determined. On the other hand, the set of spans (denoted by $T$) accompanied by $X$ is, up to ordering, uniquely determined by the code $C$ [9], [10], [15]. $T$ is called the characteristic span list of $C$ (an element of $T$ is called a characteristic span of $C$) [9], [10]. In order to clarify this fact, Gluesing-Luerssen and Weaver introduced the notion of characteristic pair $(X, T)$ of $C$ [9] Definition III.8), where $X$ is a generating set of $C$ and $T$ represents the associated spans. In this paper, we basically follow the definition of Gluesing-Luerssen and Weaver, but in order to emphasize the fact that a characteristic matrix inherently assumes the associated spans, we leave the term characteristic matrix for the definition. Thus we define as follows (cf. [9] Definition III.8)).

**Definition 2.1:** Let $C$ be an $(n, k)$ linear block code with support $I$. A characteristic matrix for $C$ with (characteristic) span list $T$ is defined to be a pair $(X, T)$, where

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

(3)

$$T = \{(a_l, b_l) : l = 1, 2, \ldots, n\}$$

(4)

have the properties:

1) $\{x_1, \ldots, x_n\}$ generates $C$.
2) $(a_l, b_l)$ is a span of $x_l$, $l = 1, \ldots, n$.
3) $a_1, \ldots, a_n$ are distinct and $b_1, \ldots, b_n$ are distinct.
4) For all $j \in I$, there exist exactly $n - k$ row indices, $l_1, \ldots, l_{n-k}$, such that $j \in (a_{l_i}, b_{l_i})$ for $i = 1, \ldots, n - k$.

Remark: Property 3) is derived from [13] Lemma 5.7 and the related remarks. Also, Property 4) is derived from the proof of [13] Theorem 5.10.

In the following, when there is no danger of confusion, we shall use the terms characteristic matrix $X$ and characteristic matrix $X$ with span list $T$ interchangeably.

**III. Characteristic Matrices for a Tail-Biting Convolutional Code**

Let $G(D)$ be a polynomial generator matrix of size $k_0 \times n_0$. Denote by $H(D)$ a corresponding polynomial check matrix. Both $G(D)$ and $H(D)$ are assumed to be canonical [13]. Consider a standard trellis of $N$ sections for a convolutional code defined by $G(D)$. Here $\max(L, M) + 1 \leq N$ is assumed, where $L$ and $M$ are the memory lengths of $G(D)$ and $H(D)$, respectively. The TB condition is a restriction that the encoder starts and ends in the same state. That is, only those paths in the trellis that start and end in the same state are admissible. We call such paths TB paths. Let $C$ be the set of all TB paths. In the following, we call a TB convolutional code of section length $N$ defined by $G(D)$ (cf. Fig.1 in Section V). When there is no danger of confusion, we will omit the phrase of section length $N$. $C$ can be regarded as a linear block code $B^{tb}$ of length $n = n_0N$. To simplify the notations, $B^{tb}$ is identified with $C$ and is denoted simply by $C$. Let

$$G(D) = G_0 + G_1D + \cdots + G_LD^L$$

(5)
be the polynomial expansion of $G(D)$, where $G_i$ $(0 \leq i \leq L)$ are $k_0 \times n_0$ matrices. Then the scalar generator matrix for $C$ $(=B^b)$ is given by

$$G_{tb}^{th} \triangleq \begin{pmatrix}
G_0 & G_1 & \ldots & G_{L-1} & G_L \\
G_0 & \ldots & \ldots & G_{L-1} & G_L \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
G_L & G_{L-1} G_L & \ldots & G_0 & \ldots \\
G_1 & G_2 & \ldots & G_L & G_0
\end{pmatrix}$$  \hspace{1cm} (6)

with size $k \times n = k_0 N \times n_0 N$ \cite{12}. Hence, we can say that a TB convolutional code $C$ is generated by $G_{tb}^{th}$. In the following, we call $G_{tb}^{th}$ the tail-biting generator matrix (abbreviated TBGM) associated with a TB convolutional code $C$ defined by $G(D)$, or simply the TBGM associated with $G(D)$.

A. Computation of Characteristic Matrices

Koetter and Vardy \cite{13} have given an algorithm which can compute a characteristic matrix for a linear block code. Consider a TB convolutional code $C$ generated by $G_{tb}^{th}$. Note that $\sigma_{n_0}(G_{tb}^{th})$ is equivalent to $G_{tb}^{th}$. That is, $G_{tb}^{th}$ has a periodic structure of period $n_0$. Using this property, a characteristic matrix $X$ for $C$ can be computed efficiently.

Let $C_j \triangleq \sigma_j(C)$. $C_j$ is the code generated by $\sigma_j(G_{tb}^{th})$. Let $X_j^n$ be a basis in minimal-span form for the code $C_j$. Then a characteristic matrix $X$ for $C$ is defined as follows \cite{14}:

$$X \triangleq X_0^* \cup \rho_1(X_1^*) \cup \ldots \cup \rho_{n_0-1}(X_{n_0-1}^*)$$
$$\cup \rho_{n_0}(X_{n_0}^*) \cup \rho_{n_0+1}(X_{n_0+1}^*) \cup \ldots \cup \rho_{2n_0-1}(X_{2n_0-1}^*)$$
$$\ldots$$
$$\cup \rho_{N(n-1)n_0}(X_{N(n-1)n_0}^*) \cup \rho_{N(n-1)n_0+1}(X_{N(n-1)n_0+1}^*) \cup \ldots \cup \rho_{Nn_0-1}(X_{Nn_0-1}^*).$$  \hspace{1cm} (7)

Since $\sigma_{n_0}(G_{tb}^{th})$ is equivalent to $G_{tb}^{th}$, we have

$$\rho_{n_0}(X_{n_0}^*) = \rho_{n_0}(X_0^*)$$
$$\rho_{n_0+1}(X_{n_0+1}^*) = \rho_{n_0}(\rho_1(X_1^*)) = \rho_{n_0}(\tilde{X}_1^*)$$
$$\ldots$$
$$\rho_{2n_0-1}(X_{2n_0-1}^*) = \rho_{n_0}(\rho_{n_0-1}(X_{n_0-1}^*)) = \rho_{n_0}(\tilde{X}_{n_0-1}^*),$$

where

$$\begin{align*}
\tilde{X}_1^* & \triangleq \rho_1(X_1^*) \\
\tilde{X}_2^* & \triangleq \rho_2(X_2^*) \\
\ldots
\end{align*}$$  \hspace{1cm} (8)

Similarly, we have

$$\rho_{2n_0}(X_{2n_0}^*) = \rho_{2n_0}(X_0^*)$$
$$\rho_{2n_0+1}(X_{2n_0+1}^*) = \rho_{2n_0}(\rho_1(X_{n_0}^*)) = \rho_{2n_0}(\tilde{X}_1^*)$$
$$\ldots$$
$$\rho_{3n_0-1}(X_{3n_0-1}^*) = \rho_{2n_0}(\rho_{n_0-1}(X_{n_0-1}^*)) = \rho_{2n_0}(\tilde{X}_{n_0-1}^*).$$

In general, for $i = 1, \ldots, N - 1$, we have

$$\rho_{in_0}(X_{in_0}^*) = \rho_{in_0}(X_0^*)$$
$$\rho_{in_0+1}(X_{in_0+1}^*) = \rho_{in_0}(\tilde{X}_1^*)$$
$$\ldots$$
$$\rho_{(i+1)n_0-1}(X_{(i+1)n_0-1}^*) = \rho_{in_0}(\tilde{X}_{n_0-1}^*).$$
Hence,

\[
X = X^*_0 \cup \rho_1(X^*_1) \cup \cdots \cup \rho_{n_0-1}(X^*_n_0-1) \\
\cup \rho_{n_0}(X^*_0) \cup \rho_{n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{n_0}(\tilde{X}^*_n_0-1) \\
\cdots \\
\cup \rho_{(N-1)n_0}(X^*_0) \cup \rho_{(N-1)n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{(N-1)n_0}(\tilde{X}^*_n_0-1)
\]

is obtained. Thus we have shown the following.

**Proposition 3.1**: A characteristic matrix \( X \) for a TB convolutional code \( C \) generated by \( G^{tb}_N \) is given by

\[
X = X^*_0 \cup \rho_1(X^*_1) \cup \cdots \cup \rho_{n_0-1}(X^*_n_0-1) \\
\cup \rho_{n_0}(X^*_0) \cup \rho_{n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{n_0}(\tilde{X}^*_n_0-1) \\
\cdots \\
\cup \rho_{(N-1)n_0}(X^*_0) \cup \rho_{(N-1)n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{(N-1)n_0}(\tilde{X}^*_n_0-1).
\]

**Corollary 3.1**: If the relation

\[
X^*_0 \cup \rho_1(X^*_1) \cup \cdots \cup \rho_{n_0-1}(X^*_n_0-1) \subseteq X^*_0 \cup \rho_{n_0}(X^*_0)
\]

holds, then a characteristic matrix \( X \) is given by

\[
X = X^*_0 \cup \rho_{n_0}(X^*_0) \cup \rho_{2n_0}(X^*_0) \cup \cdots \cup \rho_{(N-1)n_0}(X^*_0).
\]

**Proof**: From the assumption, we have

\[
\rho_{n_0}(X^*_0) \cup \rho_{n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{n_0}(\tilde{X}^*_n_0-1) \subseteq \rho_{n_0}(X^*_0) \cup \rho_{2n_0}(X^*_0).
\]

Similarly, we have

\[
\rho_{2n_0}(X^*_0) \cup \rho_{2n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{2n_0}(\tilde{X}^*_n_0-1) \subseteq \rho_{2n_0}(X^*_0) \cup \rho_{3n_0}(X^*_0)
\]

\[
\cdots \\
\rho_{(N-1)n_0}(X^*_0) \cup \rho_{(N-1)n_0}(\tilde{X}^*_1) \cup \cdots \cup \rho_{(N-1)n_0}(\tilde{X}^*_n_0-1) \subseteq \rho_{(N-1)n_0}(X^*_0) \cup \rho_{Nn_0}(X^*_0)
\]

\[
= \rho_{(N-1)n_0}(X^*_0) \cup X^*_0.
\]

Then, from Proposition 3.1, it follows that

\[
X = X^*_0 \cup \rho_{n_0}(X^*_0) \cup \rho_{2n_0}(X^*_0) \cup \cdots \cup \rho_{(N-1)n_0}(X^*_0).
\]

We remark that in many practical applications, a characteristic matrix for a TB convolutional code is obtained based on the above corollary.

**Example 1**: Consider the TB convolutional code of section length \( N = 3 \) defined by

\[
G(D) = (1 + D, D, 1 + D) \\
= (1, 0, 1) + (1, 1, 1)D \\
\triangleq G_0 + G_1 D.
\]

The associated TBGM is given by

\[
G^{tb}_3 = \begin{pmatrix}
G_0 & G_1 & 0 \\
0 & G_0 & G_1 \\
G_1 & 0 & G_0
\end{pmatrix}
\]

\[
= \begin{pmatrix}
1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1
\end{pmatrix}
\]

In this case, we have

\[
X^*_0 = \begin{pmatrix}
1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0
\end{pmatrix}
\]

(0, 5)

\[
X^*_1 = \begin{pmatrix}
1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0
\end{pmatrix}
\]

(2, 0)

(1, 7)

(1, 7)

(0, 5)

(3, 8)

(1, 7)

(3, 8)
By applying $\rho_3$ and $\rho_6$ to these matrices, a characteristic matrix $X$ is obtained as follows:

$$
X = \begin{pmatrix}
1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0
\end{pmatrix}
$$

(16)

We see that a characteristic matrix $\rho$ be a characteristic matrix for $\rho_3(X)$, whereas the spans $(0, 5]$, $(1, 7]$, $(2, 0]$, $(3, 8]$, $(4, 1]$, $(6, 2]$ are connected with $X_0^* \cup \rho_3(X_0)$. Hence, $X_0^* \cup \rho_1(X_1^*) \cup \rho_2(X_2^*) \not\subseteq X_0^* \cup \rho_3(X_0^*)$.

Note that the spans $(0, 5]$, $(1, 7]$, $(2, 0]$, $(3, 8]$, $(4, 1]$, $(6, 2]$ are connected with $X_0^* \cup \rho_3(X_0^*)$. Hence, $X_0^* \cup \rho_1(X_1^*) \cup \rho_2(X_2^*)$ is obtained as follows:

$$
X = \begin{pmatrix}
1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1
\end{pmatrix}
$$

(17)

B. Structure of the Characteristic Span List

Let $(X, T)$, where

$$
T = \begin{pmatrix}
(a_1, b_1) \\
(a_2, b_2) \\
\vdots \\
(a_n, b_n)
\end{pmatrix}
$$

be a characteristic matrix for $C$ with span list $T$, then $(\sigma_1(X), \sigma_1(T))$ is a characteristic matrix for $\sigma_1(C)$ with span list $\sigma_1(T)$ [9, Remark III.9 (b)] , where

$$
\sigma_1(X) \triangleq \begin{pmatrix}
\sigma_1(x_1) \\
\sigma_1(x_2) \\
\vdots \\
\sigma_1(x_n)
\end{pmatrix}
$$

(18)

$$
\sigma_1(T) \triangleq \begin{pmatrix}
(a_1 - 1, b_1 - 1) \\
(a_2 - 1, b_2 - 1) \\
\vdots \\
(a_n - 1, b_n - 1)
\end{pmatrix}
$$

(19)

Using repeatedly this relation, we see that $(\sigma_j(X), \sigma_j(T))$ is a characteristic matrix for $\sigma_j(C)$ with span list $\sigma_j(T)$.

Lemma 3.1: Let $C$ be a TB convolutional code generated by $G_0^N$. If $(X, T)$ is a characteristic matrix for $C$ with span list $T$, then $(\sigma_n_0(X), \sigma_n_0(T))$ is also a characteristic matrix for $C$ with span list $\sigma_n_0(T)$. Let

$$
T = \{(a_l, b_l) : l = 1, 2, \cdots, n\}
$$

(20)

Then $\sigma_n_0(T)$ is given by

$$
\sigma_n_0(T) = \{(a_l - n_0, b_l - n_0) : l = 1, 2, \cdots, n\}
$$

(21)

Since the characteristic span list is uniquely determined, $T$ and $\sigma_n_0(T)$ coincide up to ordering.

Proposition 3.2: The characteristic span list $T$ of a TB convolutional code $C$ generated by $G_0^N$ consists of the set of basic spans

$$
T_0 \triangleq \{ (0, b_0), (1, b_1), \cdots, (n_0 - 1, b_{n_0 - 1}) \}
$$

(22)
and \( \rho_{no}(T_0) \) \((i = 1, 2, \ldots, N-1)\).

Proof: Suppose that the spans in \( T \) are sorted such that

\[ T = \{(l, b_l) : l = 0, 1, \cdots, n-1\}. \]

Then we have

\[ \sigma_{no}(T) = \{(l - n_0, b_l - n_0) : l = 0, 1, \cdots, n-1\}. \]

Here take notice of the following set of spans in \( T \):

\[
\begin{cases}
(n_0, b_{n_0}), \\
(n_0 + 1, b_{n_0+1}), \\
\quad \vdots \\
(2n_0 - 1, b_{2n_0-1}).
\end{cases}
\]

In \( \sigma_{no}(T) \), it is transformed to

\[
\begin{cases}
(0, b_{n_0} - n_0), \\
(1, b_{n_0+1} - n_0), \\
\quad \vdots \\
(n_0 - 1, b_{2n_0-1} - n_0).
\end{cases}
\]

Since \( T \) and \( \sigma_{no}(T) \) coincide up to ordering,

\[
\begin{cases}
(0, b_{n_0} - n_0) = (0, b_0), \\
(1, b_{n_0+1} - n_0) = (1, b_1), \\
\quad \vdots \\
(n_0 - 1, b_{2n_0-1} - n_0) = (n_0 - 1, b_{n_0-1}),
\end{cases}
\]

holds. Hence, we have

\[
\begin{cases}
b_{n_0} = b_0 + n_0, \\
b_{n_0+1} = b_1 + n_0, \\
\quad \vdots \\
b_{2n_0-1} = b_{n_0-1} + n_0.
\end{cases}
\]

Similarly, the set of spans

\[
\begin{cases}
(2n_0, b_{2n_0}), \\
(2n_0 + 1, b_{2n_0+1}), \\
\quad \vdots \\
(3n_0 - 1, b_{3n_0-1}),
\end{cases}
\]

is transformed to

\[
\begin{cases}
(n_0, b_{2n_0} - n_0), \\
(n_0 + 1, b_{2n_0+1} - n_0), \\
\quad \vdots \\
(2n_0 - 1, b_{3n_0-1} - n_0),
\end{cases}
\]

Then for the same reason,

\[
\begin{cases}
(n_0, b_{2n_0} - n_0) = (n_0, b_{n_0}), \\
(n_0 + 1, b_{2n_0+1} - n_0) = (n_0 + 1, b_{n_0+1}), \\
\quad \vdots \\
(2n_0 - 1, b_{3n_0-1} - n_0) = (2n_0 - 1, b_{2n_0-1}),
\end{cases}
\]

holds. Hence, we have

\[
\begin{cases}
b_{2n_0} = b_{n_0} + n_0 = b_0 + 2n_0, \\
b_{2n_0+1} = b_{n_0+1} + n_0 = b_1 + 2n_0, \\
\quad \vdots \\
b_{3n_0-1} = b_{2n_0-1} + n_0 = b_{n_0-1} + 2n_0.
\end{cases}
\]

Continuing the same argument, we have

\[
\begin{cases}
b_{in_0} = b_0 + in_0, \\
b_{in_0+1} = b_1 + in_0, \\
\quad \vdots \\
b_{in_0} = b_{(i+1)n_0-1} + in_0,
\end{cases}
\]

for \( i = 1, \cdots, N - 1. \)
**Example 2:** Consider the TB convolutional code of section length \( N = 3 \) defined by the rate \( R = 2/3 \) encoder

\[
G(D) = \begin{pmatrix} 1 + D & 1 + D \\ D & 1 \end{pmatrix}
\]

\[
= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} D
\]

\[
\triangleq G_0 + G_1 D.
\]

Using the associated TBGM, i.e.,

\[
G^\text{th}_3 = \begin{pmatrix} G_0 & G_1 & 0 \\ 0 & G_0 & G_1 \\ G_1 & 0 & G_0 \end{pmatrix},
\]

a characteristic matrix \( X \) is computed as follows:

\[
X = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
\]

\[
\text{(0, 4)} \quad \text{(1, 3)} \quad \text{(2, 5)} \quad \text{(3, 7)} \quad \text{(4, 6)} \quad \text{(5, 8)} \quad \text{(6, 1)} \quad \text{(7, 0)} \quad \text{(8, 2)}.
\]

We see that the characteristic span list \( T \) consists of the set of basic spans

\[
T_0 = \{ (0, 4), (1, 3), (2, 5) \}
\]

and its right cyclic shifts by 3 and 6 positions.

**C. Counting Characteristic Matrices**

Recall the definition of a characteristic matrix \( X \) for a given code \( C \), i.e.,

\[
X \triangleq X_0^* \cup \rho_1(X_1^*) \cup \cdots \cup \rho_{n-1}(X_{n-1}^*),
\]

where \( X_j^* \) is a basis in minimal-span form for the code \( C_j = \sigma_j(C) \). Note that \( X_j^* \) is not necessarily unique. Hence, \( X \) is not uniquely determined \([9]\). With respect to this subject, Weaver \([32]\) discussed the relationship between the characteristic span list of \( C \) and the number of characteristic matrices for \( C \).

Let \( T = \{ (a_l, b_l) : l = 1, 2, \ldots, n \} \) be the characteristic span list of \( C \). Define the set \( \Theta_l \) as follows \([32]\):

\[
\Theta_l \triangleq \{ r : (a_r, b_r) \subset (a_l, b_l) \}.
\]

\(|\Theta_l|\) represents the number of spans (in \( T \)) included in a specified span \( (a_l, b_l) \). Weaver \([32]\) proved the following.

**Lemma 3.2 (Weaver \([32]\))**: Let \((a_l, b_l)\) be a characteristic span of \( C \). Then there exist \( 2^{|\Theta_l|} \) characteristic generators for \( C \) having this span.

This fact is derived from the next observation:

Let \((a_r, b_r) \subset (a_l, b_l)\) and consider two characteristic generators \( x_r \) and \( x_l \) with spans \((a_r, b_r)\) and \((a_l, b_l)\), respectively. Then \( x_l + x_r \) is also a characteristic generator with span \((a_l, b_l)\).

Consider a TB convolutional code \( C \) generated by \( G^H_N \). We have already shown that the characteristic span list \( T \) of \( C \) consists of the set of basic spans

\[
T_0 = \{ (0, b_0), (1, b_1), \ldots, (n_0 - 1, b_{n_0 - 1}) \}
\]

and \( \rho_{i_0}(T_0) \) \((i = 1, 2, \ldots, N - 1)\). Hence, it suffices to consider the spans in \( T_0 \) for the purpose of counting the number of characteristic matrices. Define \( \Theta_i \) \((1 \leq i \leq n_0)\) as follows:

\[
\Theta_1 \triangleq \{ r : (a_r, b_r) \subset (0, b_0) \}
\]

\[
\Theta_2 \triangleq \{ r : (a_r, b_r) \subset (1, b_1) \}
\]

\[
\Theta_i \triangleq \{ r : (a_r, b_r) \subset (n_0 - 1, b_{n_0 - 1}) \}.
\]
Also, let
\[ \theta_i = |\Theta_i| \quad (i = 1, 2, \cdots, n_0) \]  
\[ \theta \triangleq \theta_1 + \theta_2 + \cdots + \theta_{n_0}. \] (29) (30)

Then we have the following:
- There exist \( 2^{\theta_1} \) characteristic generators having span \( (0, b_0] \).
- There exist \( 2^{\theta_2} \) characteristic generators having span \( (1, b_1] \).
- There exist \( 2^{\theta_{n_0}} \) characteristic generators having span \( (n_0 - 1, b_{n_0 - 1}] \).

As a result, the degree of freedom related to the spans in \( T_0 \) is given by
\[ 2^{\theta_1} \times 2^{\theta_2} \times \cdots \times 2^{\theta_{n_0}} = 2^{\theta_1 + \theta_2 + \cdots + \theta_{n_0}} = 2^\theta. \] (31)

Since this degree of freedom is common to other \((N - 1)\) blocks of spans in \( T \), the overall degree of freedom related to \( T \) becomes
\[ (2^\theta)^N = 2^{\theta N}. \] (32)

Thus we have shown the following.

Proposition 3.3: Let \( C \) be a TB convolutional code generated by \( G^t_N \). Let \( \theta_i \) \((1 \leq i \leq n_0)\) and \( \theta \) be as above. Then there exist \( 2^{\theta N} \) characteristic matrices for \( C \).

Example 2 (Continued): Take notice of the first three rows of the characteristic matrix \( X \). We have
\[ \begin{cases} 
\Theta_1 = \{2\} \\
\Theta_2 = \emptyset \\
\Theta_3 = \emptyset. 
\end{cases} \]

Hence, \( \theta = 1 + 0 + 0 = 1 \) and there exist \( 2^{1 \times 3} = 8 \) characteristic matrices.

D. Span Lengths of Characteristic Generators

Let \( |x| = (a, b) \) be a span of a codeword \( x \). Then the span length of \( x \) is defined by \(|[a, b]|\), i.e., the number of elements in the closed interval \([a, b]\). When a span alone is referred to without specifying the accompanied codeword, we use the term the span length of a span \((a, b)\). Let \( T \) be the characteristic span list of a TB convolutional code \( C \) generated by \( G^t_N \). Suppose that the spans in \( T \) are sorted such that
\[ T = \{(l, b_l) : l = 0, 1, \cdots, n - 1 \}. \]

Then by [13] Theorem 5.10,
\[ |(0, b_0)| + |(1, b_1)| + \cdots + |(n - 1, b_{n_0 - 1})| = n(n - k) \] (33)
holds. Due to the structure of \( T \) (see Proposition 3.2), the left-hand side of the above equality becomes
\[ N(|(0, b_0)| + |(1, b_1)| + \cdots + |(n_0 - 1, b_{n_0 - 1})|) = N((\ell_1 - 1) + (\ell_2 - 1) + \cdots + (\ell_{n_0} - 1)) = N(\ell_1 + \ell_2 + \cdots + \ell_{n_0} - n_0) = N(\ell - n_0), \]
where
\[ \begin{cases} 
\ell_1 \triangleq |[0, b_0]| \\
\ell_2 \triangleq |[1, b_1]| \\
\cdots \\
\ell_{n_0} \triangleq |[n_0 - 1, b_{n_0 - 1}]| \\
\ell \triangleq \ell_1 + \ell_2 + \cdots + \ell_{n_0}. 
\end{cases} \] (34)

In the derivation, we also used the relation
\[ \ell_i = |[i - 1, b_{i-1}]| = |(i - 1, b_{i-1})| + 1 \quad (i = 1, 2, \cdots, n_0). \] (35)
Replacing \(n\) and \(k\) by \(n_0 N\) and \(k_0 N\), respectively, the above equality reduces to
\[
N(\ell - n_0) = n_0 N(n_0 N - k_0 N)
\]
\[
\ell - n_0 = n_0((n_0 - k_0) N)
\]
\[
\ell = n_0((n_0 - k_0) N + 1).
\]

Thus we have shown the following.

**Proposition 3.4:** Let \(T\) be the characteristic span list of a TB convolutional code \(C\) generated by \(G_N^b\). Denote by \(T_0\) the set of basic spans in \(T\). Then the sum \(\ell\) of span lengths of spans in \(T_0\) is given by
\[
\ell = n_0((n_0 - k_0) N + 1).
\]

**Example 2 (Continued):** We have
\[
\begin{align*}
\ell_1 &= [0, 4] = 5 \\
\ell_2 &= [1, 3] = 3 \\
\ell_3 &= [2, 5] = 4
\end{align*}
\]
\[
\ell = \ell_1 + \ell_2 + \ell_3 = 5 + 3 + 4 = 12.
\]

Also, we have
\[
n_0((n_0 - k_0) N + 1) = 3((3 - 2) \times 3 + 1) = 12.
\]

**IV. Transformations of \(G(D)\) and the Corresponding TBGM’s**

In this section, we discuss the relationship between transformations of a generator matrix \(G(D)\) and the corresponding TBGM’s \((G_N^b)'\)’s. We consider the following transformations of \(G(D)\):

a) Dividing the \(j\)th column by \(D^p\).

b) Multiplying the \(j\)th column by \(D^q\).

c) Adding the \(i\)th row multiplied by \(D^q\) to the \(j\)th row.

d) Implicit transformations.

In the next section, we will see that these transformations play an essential role in trellis reduction for TB convolutional codes.

**A. Dividing a Column of \(G(D)\) by \(D^p\)**

Suppose that the \(j\)th column of \(G(D)\) has a monomial factor \(D^p\) \((1 \leq p \leq L)\). We can assume without loss of generality that \(j = 1\) and \(p = 1\). Hence, \(G(D)\) has the form
\[
G(D) = \begin{pmatrix}
g_{1,1}^{(1)}(D) & g_{1,2}^{(1)}(D) & \cdots & g_{1,n_0}^{(1)}(D) \\
g_{2,1}^{(1)}(D) & g_{2,2}^{(1)}(D) & \cdots & g_{2,n_0}^{(1)}(D) \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(1)}(D) & g_{k_0,2}^{(1)}(D) & \cdots & g_{k_0,n_0}^{(1)}(D)
\end{pmatrix}.
\]

Let
\[
G(D) = \begin{pmatrix}
g_{1,1}^{(0)} & g_{1,2}^{(0)} & \cdots & g_{1,n_0}^{(0)} \\
g_{2,1}^{(0)} & g_{2,2}^{(0)} & \cdots & g_{2,n_0}^{(0)} \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(0)} & g_{k_0,2}^{(0)} & \cdots & g_{k_0,n_0}^{(0)}
\end{pmatrix} + \begin{pmatrix}
g_{1,1}^{(1)} & g_{1,2}^{(1)} & \cdots & g_{1,n_0}^{(1)} \\
g_{2,1}^{(1)} & g_{2,2}^{(1)} & \cdots & g_{2,n_0}^{(1)} \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(1)} & g_{k_0,2}^{(1)} & \cdots & g_{k_0,n_0}^{(1)}
\end{pmatrix} D + \begin{pmatrix}
g_{1,1}^{(2)} & g_{1,2}^{(2)} & \cdots & g_{1,n_0}^{(2)} \\
g_{2,1}^{(2)} & g_{2,2}^{(2)} & \cdots & g_{2,n_0}^{(2)} \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(2)} & g_{k_0,2}^{(2)} & \cdots & g_{k_0,n_0}^{(2)}
\end{pmatrix} D^2 + \cdots + \begin{pmatrix}
g_{1,1}^{(L)} & g_{1,2}^{(L)} & \cdots & g_{1,n_0}^{(L)} \\
g_{2,1}^{(L)} & g_{2,2}^{(L)} & \cdots & g_{2,n_0}^{(L)} \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(L)} & g_{k_0,2}^{(L)} & \cdots & g_{k_0,n_0}^{(L)}
\end{pmatrix} D^L
\]
\[
\triangleq G_0 + G_1 D + \cdots + G_{L-1} D^{L-1} + G_L D^L
\]
be the polynomial expansion of $G(D)$. Comparing the $(i,1)$ $(1 \leq i \leq k_0)$ entries, we have

$$
\begin{align*}
g_{1,1}(D)D &= g_{1,1}^{(0)} + g_{1,1}^{(1)}D + g_{1,1}^{(2)}D^2 + \cdots + g_{1,1}^{(L)}D^L \\
g_{2,1}(D)D &= g_{2,1}^{(0)} + g_{2,1}^{(1)}D + g_{2,1}^{(2)}D^2 + \cdots + g_{2,1}^{(L)}D^L \\
& \vdots \\
g_{k_0,1}(D)D &= g_{k_0,1}^{(0)} + g_{k_0,1}^{(1)}D + g_{k_0,1}^{(2)}D^2 + \cdots + g_{k_0,1}^{(L)}D^L.
\end{align*}
$$

By these equations, we have

$$
\begin{align*}
g_{1,1}'(D) &= g_{1,1}^{(0)} + g_{1,1}^{(1)}D + \cdots + g_{1,1}^{(L)}D^{L-1} + 0 \cdot D^L \\
g_{2,1}'(D) &= g_{2,1}^{(0)} + g_{2,1}^{(1)}D + \cdots + g_{2,1}^{(L)}D^{L-1} + 0 \cdot D^L \\
& \vdots \\
g_{k_0,1}'(D) &= g_{k_0,1}^{(0)} + g_{k_0,1}^{(1)}D + \cdots + g_{k_0,1}^{(L)}D^{L-1} + 0 \cdot D^L.
\end{align*}
$$

Dividing the first column of $G(D)$ by $D$, let the resulting matrix be $G'(D)$. Then $G'(D)$ has the polynomial expansion:

$$
G'(D) = \begin{pmatrix} g_{1,1}'(D) & g_{1,2}'(D) & \cdots & g_{1,n_0}'(D) \\
g_{2,1}'(D) & g_{2,2}'(D) & \cdots & g_{2,n_0}'(D) \\
& \vdots & \ddots & \vdots \\
g_{k_0,1}'(D) & g_{k_0,2}'(D) & \cdots & g_{k_0,n_0}'(D) \end{pmatrix} + \begin{pmatrix} g_{1,1}^{(2)} & g_{1,2}^{(2)} & \cdots & g_{1,n_0}^{(2)} \\
g_{2,1}^{(2)} & g_{2,2}^{(2)} & \cdots & g_{2,n_0}^{(2)} \\
& \vdots & \ddots & \vdots \\
g_{k_0,1}^{(2)} & g_{k_0,2}^{(2)} & \cdots & g_{k_0,n_0}^{(2)} \end{pmatrix} D + \begin{pmatrix} g_{1,1}^{(3)} & g_{1,2}^{(3)} & \cdots & g_{1,n_0}^{(3)} \\
g_{2,1}^{(3)} & g_{2,2}^{(3)} & \cdots & g_{2,n_0}^{(3)} \\
& \vdots & \ddots & \vdots \\
g_{k_0,1}^{(3)} & g_{k_0,2}^{(3)} & \cdots & g_{k_0,n_0}^{(3)} \end{pmatrix} D^2 + \cdots + \begin{pmatrix} g_{1,1}^{(L)} & g_{1,2}^{(L)} & \cdots & g_{1,n_0}^{(L)} \\
g_{2,1}^{(L)} & g_{2,2}^{(L)} & \cdots & g_{2,n_0}^{(L)} \\
& \vdots & \ddots & \vdots \\
g_{k_0,1}^{(L)} & g_{k_0,2}^{(L)} & \cdots & g_{k_0,n_0}^{(L)} \end{pmatrix} D^{L-1} + \begin{pmatrix} 0 & g_{1,2}^{(L)} & \cdots & g_{1,n_0}^{(L)} \\
0 & g_{2,2}^{(L)} & \cdots & g_{2,n_0}^{(L)} \\
& \vdots & \ddots & \vdots \\
0 & g_{k_0,2}^{(L)} & \cdots & g_{k_0,n_0}^{(L)} \end{pmatrix} D^L
\end{align*}
$$

\begin{align*}
\triangleq G_0' + G_1' D + \cdots + G_{L-1}' D^{L-1} + G_L' D^L.
\end{align*}

Consider the TBGM associated with $G'(D)$, denoted by $G_{N}^{tb}$, where

$$
G_{N}^{tb} = \begin{pmatrix} G_0' & G_1' & \cdots & G'_{L-1} & G_L' \\
G_0' & G_1' & \cdots & G'_{L-1} & G_L' \\
& \ddots & \ddots & \ddots & \ddots \\
& G_0' & G_1' & \cdots & G'_{L-1} \\
G_{L-1}' & G_L' & \cdots & G'_{L-1} & \end{pmatrix}.
$$

Note that both $G_N^{tb}$ and $G_N^{tb}$ can be regarded as matrices having $N$ blocks of $n_0$ columns. Then in view of the entries of $G'_i$ $(0 \leq i \leq L)$ and the relation

$$
\begin{align*}
g_{1,1}^{(0)} &= 0 \\
g_{2,1}^{(0)} &= 0 \\
& \vdots \\
g_{k_0,1}^{(0)} &= 0,
\end{align*}
$$
Accordingly, the polynomial expansion of each block to the left by \( n_0 \) positions. Thus we have the following.

**Proposition 4.1:** Regard \( G_N^{th} \) as a matrix having \( N \) blocks of \( n_0 \) columns. Suppose that the \( j \)th column of \( G(D) \) has a monomial factor \( D^p \) \((1 \leq p \leq L)\). Then dividing the \( j \)th column of \( G(D) \) by \( D^p \) is equivalent to cyclically shifting the \( j \)th column of each block of \( G_N^{th} \) to the left by \( p n_0 \) positions.

Let \( C \) be a TB convolutional code of section length \( N \) defined by \( G(D) \). Note that each codeword in \( C \) consists of \( N \) blocks of \( n_0 \) components. Here let us cyclically shift the \( j \)th component of each block to the left by \( p n_0 \) positions. Denote by \( C' \) the set of resulting (modified) codewords. We have already shown that \( G_N^{th} \) is obtained from \( G_N^{th} \) by cyclically shifting the \( j \)th column of each block to the left by \( p n_0 \) positions. Hence, \( C' \) is generated by \( G_N^{th} \). In words, \( C' \) is represented as a TB convolutional code defined by \( G'(D) \).

**B. Multiplying a Column of** \( G(D) \) **by** \( D^q \)

Consider multiplication of the \( j \)th column of \( G(D) \) by \( D^q \), where \( q + L + 1 \leq N \). In the following, we assume without loss of generality that \( j = 1 \times q = 1 \ (L + 2 \leq N) \). Hence, the resulting matrix \( G'(D) \) has the form

\[
G'(D) = \begin{pmatrix}
g_{1,1}(D) & g_{1,2}(D) & \cdots & g_{1,n_0}(D) \\
g_{2,1}(D) & g_{2,2}(D) & \cdots & g_{2,n_0}(D) \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}(D) & g_{k_0,2}(D) & \cdots & g_{k_0,n_0}(D)
\end{pmatrix}
\]

\[
\triangleq \begin{pmatrix}
g_{1,1}(D) & g_{1,2}(D) & \cdots & g_{1,n_0}(D) \\
g_{2,1}(D) & g_{2,2}(D) & \cdots & g_{2,n_0}(D) \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}(D) & g_{k_0,2}(D) & \cdots & g_{k_0,n_0}(D)
\end{pmatrix} D^q
\]

Then we have

\[
g'_{1,1}(D) = g_{1,1}(D) D
\]

\[
= (g_{1,1}^{(0)} + g_{1,1}^{(1)} D + \cdots + g_{1,1}^{(L)} D^L) D
\]

\[
= g_{1,1} D + g_{1,1}^2 D^2 + \cdots + g_{1,1}^{(L-1)} D^L + g_{1,1}^{(L)} D^{L+1}
\]

\[
\vdots
\]

\[
g'_{k_0,1}(D) = g_{k_0,1}(D) D
\]

\[
= (g_{k_0,1}^{(0)} + g_{k_0,1}^{(1)} D + \cdots + g_{k_0,1}^{(L)} D^L) D
\]

\[
= g_{k_0,1} D + g_{k_0,1}^2 D^2 + \cdots + g_{k_0,1}^{(L-1)} D^L + g_{k_0,1}^{(L)} D^{L+1}.
\]

Accordingly, the polynomial expansion of \( G'(D) \) becomes

\[
G'(D) = \begin{pmatrix}
0 & g_{1,2}^{(0)} & \cdots & g_{1,n_0}^{(0)} \\
0 & g_{2,2}^{(0)} & \cdots & g_{2,n_0}^{(0)} \\
\vdots & \vdots & \ddots & \vdots \\
0 & g_{k_0,2}^{(0)} & \cdots & g_{k_0,n_0}^{(0)}
\end{pmatrix}
+ \begin{pmatrix}
g_{1,1}^{(1)} & g_{1,2}^{(1)} & \cdots & g_{1,n_0}^{(1)} \\
g_{2,1}^{(1)} & g_{2,2}^{(1)} & \cdots & g_{2,n_0}^{(1)} \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(1)} & g_{k_0,2}^{(1)} & \cdots & g_{k_0,n_0}^{(1)}
\end{pmatrix} D
\]

\[
+ \begin{pmatrix}
g_{1,1}^{(L-1)} & g_{1,2}^{(L-1)} & \cdots & g_{1,n_0}^{(L-1)} \\
g_{2,1}^{(L-1)} & g_{2,2}^{(L-1)} & \cdots & g_{2,n_0}^{(L-1)} \\
\vdots & \vdots & \ddots & \vdots \\
g_{k_0,1}^{(L-1)} & g_{k_0,2}^{(L-1)} & \cdots & g_{k_0,n_0}^{(L-1)}
\end{pmatrix} D^L
\]

\[\triangleq G'_0 + G'_1 D + \cdots + G'_L D^L + G'_{L+1} D^{L+1}.
\]
Consider the TBGM associated with $G'(D)$, denoted by $G'^{tb}_N$, where

$$
G'^{tb}_N = \begin{pmatrix}
G'_0 & G'_1 & \cdots & G'_{L-1} & G'_L & G'_{L+1} \\
G'_0 & \cdots & G'_{L-1} & G'_L & \cdots & G'_{L+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
G'_{L-1} & G'_L & \cdots & G'_1 & G'_{L+1} & G'_0 \\
G'_0 & G'_1 & \cdots & G'_L & \cdots & G'_{L+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
G'_1 & G'_2 & \cdots & G'_L & \cdots & G'_{L+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
G'_0 & & & & & \\
G'_0 & & & & & \\
\end{pmatrix}.
$$

(49)

Note that $G'^{tb}_N$ and $G'^{tb}_N$ consist of $N$ blocks of $n_0$ columns as above. In view of the entries of $G'_i$ ($0 \leq i \leq L + 1$), we see that $G'^{tb}_N$ is obtained from $G'^{tb}_N$ by cyclically shifting the first column of each block to the right by $n_0$ positions. Thus we have the following.

**Proposition 4.2:** Regard $G'^{tb}_N$ as a matrix having $N$ blocks of $n_0$ columns. Suppose that $q + L + 1 \leq N$. Then multiplying the $j$th column of $G(D)$ by $D^q$ is equivalent to cyclically shifting the $j$th column of each block of $G'^{tb}_N$ to the right by $qn_0$ positions.

**Remark:** In order for $G'^{tb}_N$ to be defined, the condition $q + L + 1 \leq N$ is required.

Let $C$ be a TB convolutional code of section length $N$ with generator matrix $G(D)$. Let $C'$ be as in the previous section. In this case, however, the $j$th component of each block is cyclically shifted to the right by $qn_0$ positions. We have shown that $G'^{tb}_N$ is obtained from $G'^{tb}_N$ by cyclically shifting the $j$th column of each block to the right by $qn_0$ positions. Hence, $C'$ is generated by $G'^{tb}_N$. In words, $C'$ is represented as a TB convolutional code defined by $G'(D)$.

C. $g_j(D) \leftrightarrow g_j(D) + D^q g_i(D)$

Consider addition of the $i$th row $g_i(D)$ multiplied by $D^q$ to the $j$th row $g_j(D)$, denoted by $g_j(D) \leftrightarrow g_j(D) + D^q g_i(D)$, where $q + L + 1 \leq N$. In the following, we assume without loss of generality that $i = 1$ and $j = 2$. Let the first row of $G(D)$ be

$$
\begin{align*}
g_1(D) &= (g_{1,1}(D), g_{1,2}(D), \ldots, g_{1,n_0}(D)).
\end{align*}
$$

Also, let

$$
\begin{align*}
g_1(D) &= (g_{1,1}^{(0)}(D), g_{1,2}^{(0)}(D), \ldots, g_{1,n_0}^{(0)}(D)) + (g_{1,1}^{(1)}(D), g_{1,2}^{(1)}(D), \ldots, g_{1,n_0}^{(1)}(D)) + \cdots \\
&\quad + (g_{1,1}^{(L-1)}(D), g_{1,2}^{(L-1)}(D), \ldots, g_{1,n_0}^{(L-1)}(D))D^{L-1} + (g_{1,1}^{(L)}(D), g_{1,2}^{(L)}(D), \ldots, g_{1,n_0}^{(L)}(D))D^L \\
&\triangleq g_0 + g_1 D + \cdots + g_{L-1} D^{L-1} + g_L D^L
\end{align*}
$$

be the polynomial expansion, where the size of $g_i$ ($0 \leq i \leq L$) is $1 \times n_0$. Then the polynomial expansion of $g_1(D)D^q$ becomes

$$
\begin{align*}
g_1(D)D^q &= g_0 D^q + g_1 D^{q+1} + \cdots + g_{L-1} D^{q+L-1} + g_L D^{q+L}.
\end{align*}
$$

Note that the first row of $G'^{tb}_N$ is expressed as

$$
\begin{pmatrix}
g_{0}, g_{1}, \ldots, g_{L-1}, g_{L}, 0, \ldots, 0 \\
0, \ldots, 0, g_{0}, g_{1}, \ldots, g_{L-1}, g_{L}, 0, \ldots, 0
\end{pmatrix}_N
$$

Hence, its right cyclic shift by $qn_0$ positions, i.e.,

$$
\begin{pmatrix}
g_{0}, g_{1}, \ldots, g_{L-1}, g_{L}, 0, \ldots, 0 \\
0, \ldots, 0, g_{0}, g_{1}, \ldots, g_{L-1}, g_{L}, 0, \ldots, 0
\end{pmatrix}_N
$$

coincides with the $(qk_0 + 1)$th row of $G'^{tb}_N$. That is, $g_2(D) \leftrightarrow g_2(D) + D^q g_1(D)$ corresponds to addition of the $(qk_0 + 1)$th row to the second row within the matrix $G'^{tb}_N$. Note that this is an elementary row operation. Thus we have the following.

**Proposition 4.3:** Suppose that $q + L + 1 \leq N$. Consider the operation $g_j(D) \leftrightarrow g_j(D) + D^q g_i(D)$. Let the resulting matrix be $G'(D)$ and the associated TBGM be $G'^{tb}_N$. Then $G'^{tb}_N$ is equivalent to $G'^{tb}_N$.

Taking into consideration Proposition 4.3, let us introduce a useful notion. Let $C$ and $C'$ be TB convolutional codes of section length $N$ defined by $G(D)$ and $G'(D)$, respectively. Denote by $L$ and $L'$ the memory lengths of $G(D)$ and $G'(D)$, respectively, where $\max(L, L') + 1 \leq N$. Let $G'_N$ and $G'^{tb}_N$ be the TBGM’s associated with $C$ and $C'$, respectively. We see that if $G'_N$ and $G'^{tb}_N$ are equivalent, then $C = C'$. All of this leads to the following definition.

**Definition 4.1:** When $G'^{tb}_N$ and $G'^{tb}_N$ are equivalent, we say that $G(D)$ and $G'(D)$ are “TB-equivalent”. 
Thus we have the following.

**Proposition 4.4:** If $G(D)$ and $G'(D)$ are TB-equivalent, then a TB convolutional code defined by $G(D)$ is represented as a TB convolutional code defined by $G'(D)$, and vice versa.

**Proof:** A direct consequence of the definition of TB-equivalent.



## V. Trellis Reduction for TB Convolutional Codes

In this section, we will show that for a TB convolutional code of short to moderate section length, the associated TB trellis can be reduced. We begin with an example.

### A. An Example of Trellis Reduction

![Fig. 1. The tail-biting convolutional code defined by $G(D) = (1 + D + D^2, 1 + D^2)$ ($N = 5$).](image)

Consider the TB convolutional code $C$ defined by $G(D) = (1 + D + D^2, 1 + D^2)$, where the section length $N$ is set to 5. The corresponding TB trellis is shown in Fig. 1, where the paths which start and end in the same state are TB paths (i.e., valid codewords). Then $C$ is the set of all TB paths. Since $G(D)$ has the polynomial expansion

$$G(D) \equiv (1, 1) + (1, 0)D + (1, 1)D^2 \quad \text{(50)}$$

the TBGM associated with $C$ is given by

$$G_{tb}^5 = \begin{pmatrix} G_0 & G_1 & G_2 & 0 & 0 \\
0 & G_0 & G_1 & G_2 & 0 \\
0 & 0 & G_0 & G_1 & G_2 \\
G_2 & 0 & 0 & G_0 & G_1 \\
G_1 & G_2 & 0 & 0 & G_0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \quad \text{(52)}$$

Based on $G_{tb}^5$, a characteristic matrix $X$ for $C$ is computed as follows:

$$X = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \ \end{pmatrix} \quad \text{(54)}$$
Choosing 5 rows from \( X \), let

\[
G' = \begin{pmatrix}
0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

(55)

We see that the rows of \( G' \) are linearly independent and thus generate \( C \), i.e., \( G' \) is equivalent to \( G'^{tb} \).

Here note that \( G' \) consists of the first row and its right cyclic shifts by \( i \times 2 \) (\( 1 \leq i \leq 4 \)) positions. Accordingly, \( G' \) can be regarded as the TBGM associated with

\[
G'(D) = (D^3, 1 + D)
\]

(56)

\[
= (0, 1) + (0, 1)D + (1, 0)D^3.
\]

(57)

Hence, \( C \) is equally represented as a TB convolutional code defined by \( G'(D) = (D^3, 1 + D) \). We remark that the constraint length of \( G'(D) \) is \( \nu' = 3 \) and is greater than that of \( G(D) \).

On the other hand, observe that the first column of \( G'(D) \) has a factor \( D^2 \). Then dividing the first column by \( D^2 \), we have

\[
G'(D) = (D^3, 1 + D) \rightarrow G''(D) = (D, 1 + D).
\]

(58)

Note that this transformation corresponds to cyclically shifting the first component of each branch (of a TB path) to the left by two branches (cf. Proposition 4.1). By this transformation, the original TB convolutional code is represented using a trellis associated with \( G''(D) \) as well. The trellis for \( G''(D) = (D, 1 + D) \) is shown in Fig.2. For example, take notice of the TB path in Fig.1 which starts and ends in state \((00)\):

\[
w = 11 \ 01 \ 10 \ 01 \ 11.
\]

Cyclically shifts the first component of each branch to the left by two branches, it becomes

\[
w_{tb} = 11 \ 01 \ 10 \ 11 \ 01.
\]

We see that the modified path is represented as a path which starts and ends in state \((1)\) in Fig.2.

This example shows that there are cases where a given TB convolutional code is represented using a reduced trellis with less state complexity, if we allow partial cyclic shifts of a TB path.

![Fig. 2. The tail-biting convolutional code defined by \( G''(D) = (D, 1 + D) \) (\( N = 5 \)).](image)

**B. Trellis Reduction for TB Convolutional Codes**

The argument in the previous section, though it was presented in terms of a specific example, is entirely general. Then the method can be directly extended to a general case. Let \( G(D) \) be as in Section III. Denote by \( \nu \) the constraint length of \( G(D) \). Consider a TB convolutional code \( C \) of section length \( N \) defined by \( G(D) \). The trellis reduction procedure becomes as follows.

**Procedure for trellis reduction:**

i) Compute a characteristic matrix \( X \) for \( C \) based on the TBGM \( G'^{tb} \), where \( X \) consists of \( n_0 \) rows and their right cyclic shifts by integer multiple of \( n_0 \).

ii) Choosing \( k \) rows from \( X \), form \( G' \), where \( G' \) has the properties:

1) The rows of \( G' \) are linearly independent and thus generate \( C \).

2) \( G' \) consists of \( k_0 \) rows and their right cyclic shifts by integer multiple of \( n_0 \).

iii) (Direct reduction) \( G' \) is regarded as the TBGM associated with another generator matrix \( G'(D) \). Let \( \nu' \) be the constraint length of \( G'(D) \). If \( \nu' < \nu \), then trellis reduction for \( C \) is realized.

iv) (Indirect reduction) Even if \( \nu' \geq \nu \), there are cases where \( G'(D) \) has a monomial factor \( D^p \) in some \((j)\)th column. Then there is a possibility that \( \nu' \) is reduced by dividing the \( j \)th column of \( G'(D) \) by \( D^p \) (the resulting matrix is denoted by \( G''(D) \)). Let \( \nu'' \) be the constraint length of \( G''(D) \). If \( \nu'' < \nu \), then the original TB trellis can be reduced. That is, by
cyclically shifting the $j$th component of each branch of a TB path ($\in C$) to the left by $p$ branches, the set of modified paths are equally represented as a TB convolutional code defined by $G''(D)$ (this is justified by Proposition 4.1). Thus trellis reduction is accomplished.

v) $X$ is not necessarily unique \([9]\). Hence, if necessary, try i) $\sim$ iv) using another characteristic matrix $X'$ for $C$.

Remark: For row rate codes, it is rather easy to find $G'$ which is equivalent to $G''_N$. Also, row rate codes make it easy to determine whether $G'(D)$ can be reduced or not.

As is stated above, there are some restrictions on the selection of $X$ and $G'$. We have the following.

Proposition 5.1: The number of characteristic matrices $X'$s in i) is given by $2^p$, where $\theta$ is defined in Section III-C. For a fixed $X$, the number of $G''_N'$s which satisfy the condition 2) in ii) is given by $m,C_k$.

Proof: $G'$ is a candidate for a TBGM associated with an encoder. Hence, the above is a consequence of the structure of TBGM.

Example 3: Consider the TB convolutional code $C$ of section length $N = 6$ defined by

$$G(D) = (1 + D + D^2 + D^3, 1 + D^2 + D^3).$$

Using the associated TBGM, i.e., $G_{6}^{tb}$, a characteristic matrix $X$ for $C$ is computed as follows:

$$X = \begin{pmatrix}
1000010010000 \\
011001100000 \\
001000100100 \\
000110011000 \\
000010000101 \\
000001100010 \\
010000111000 \\
100100010000 \\
011000000110 \\
001001000010 \\
100010010001 \\
100111000001
\end{pmatrix}$$

Choosing 6 rows from $X$, define $G'$ as

$$G' = \begin{pmatrix}
1000010010000 \\
0010000100100 \\
0001000100100 \\
0100000100010 \\
0010010000100 \\
0001001000010 \\
0100100000100 \\
0010010000010 \\
0100011000001 \\
1000011000010 \\
1001110000010 \\
1001110000010
\end{pmatrix}$$

We see that $G'$ is equivalent to $G_{6}^{tb}$. Also, we see that $G'$ is the TBGM associated with

$$G'(D) = (1 + D^2, D^3).$$

Note that the constraint length $n' = 3$ of $G'(D)$ is not reduced compared to that of $G(D)$. On the other hand, observe that the second column of $G'(D)$ has a factor $D$. Then dividing the column by $D$, we have

$$G'(D) = (1 + D^2, D^3) \rightarrow G''(D) = (1 + D^2, D^2).$$

This transformation corresponds to cyclically shifting the second component of each branch of a TB path to the left by one branch (cf. Proposition 4.1). As a result, the modified paths are represented using the trellis for $G''(D)$. Thus trellis reduction for $C$ is accomplished.

Remark: As is stated above, $X$ is not necessarily unique. For example, if a characteristic matrix

$$X' = \begin{pmatrix}
1110111100000 \\
1110011000000 \\
0111101110000 \\
0111100111000 \\
0111100111000 \\
0111100111000 \\
0111100111000 \\
0111100111000 \\
0111100111000 \\
0111100111000 \\
0111100111000 \\
0111100111000
\end{pmatrix}$$

is defined in Section III-C. For a
is used, then trellis reduction cannot be realized using the above procedure.

Using appropriate characteristic matrices, the above reduction method can also be applied to the following cases:

1. \( R = 1/2, \nu = 4, N = 6 \):
   
   \[
   G(D) = (1 + D + D^4, 1 + D^2 + D^3 + D^4)
   \]
   
   \[
   G'(D) = (D^2, 1 + D)
   \]
   
   \[
   G''(D) = (D, 1 + D).
   \]

2. \( R = 1/2, \nu = 5, N = 10 \):
   
   \[
   G(D) = (1 + D + D^2 + D^3 + D^4 + D^5, 1 + D^3 + D^5)
   \]
   
   \[
   G'(D) = (D^4 + D^5, 1 + D + D^4)
   \]
   
   \[
   G''(D) = (D^3 + D^5, 1 + D + D^4).
   \]

3. \( R = 1/2, \nu = 6, N = 8 \):
   
   \[
   G(D) = (1 + D + D^4 + D^5 + D^6, 1 + D^2 + D^3 + D^4 + D^6)
   \]
   
   \[
   G'(D) = (D + D^2 + D^3, 1 + D^2 + D^3).
   \]

4. \( R = 1/2, \nu = 6, N = 8 \):
   
   \[
   G(D) = (1 + D + D^2 + D^3 + D^4 + D^6)
   \]
   
   \[
   G'(D) = (1 + D^2 + D^3, D^2 + D^3 + D^4)
   \]
   
   \[
   G''(D) = (1 + D^2 + D^3, D + D^2 + D^3).
   \]

5. \( R = 1/3, \nu = 3, N = 5 \):
   
   \[
   G(D) = (1 + D + D^2 + D^3, 1 + D + D^3, 1 + D^2 + D^3)
   \]
   
   \[
   G'(D) = (D + D^2 + D^3, 1)
   \]
   
   \[
   G''(D) = (1 + D, D, 1).
   \]

6. \( R = 2/3, \nu = 4, N = 6 \):
   
   \[
   G(D) = \begin{pmatrix}
   D + D^2 & 1 + D & 1 + D + D^2 \\
   1 & 1 + D^2 & 1 + D^2
   \end{pmatrix}
   \]
   
   \[
   G'(D) = \begin{pmatrix}
   D^2 & 1 + D & D \\
   D & 0 & 1 + D^2
   \end{pmatrix}
   \]
   
   \[
   G''(D) = \begin{pmatrix}
   D & 1 + D & D \\
   1 & 0 & 1 + D^2
   \end{pmatrix}.
   \]

C. Trellis Reduction Using a Reciprocal Dual Encoder

For high rate codes, \( G'(D) \) may not have a monomial factor in any columns. Then it is not easily determined whether \( G'(D) \) can be reduced or not. In such cases, it is useful to consider a reciprocal dual encoder \( \tilde{H}'(D) \) associated with \( G'(D) \).

A reciprocal dual encoder \( [21] \) is defined as follows: Let \( G(D) \) be as in Section III. Also, let \( H(D) \) be a corresponding check matrix with size \((n_0 - k_0) \times n_0\). A reciprocal dual encoder \( \tilde{H}(D) \) is obtained by substituting \( D^{-1} \) for \( D \) in \( H(D) \) and by multiplying the \( i \)th \((1 \leq i \leq n_0 - k_0)\) row of the resulting matrix by \( D^{\nu_i} \), where \( \nu_i \) is the degree of the \( i \)th row of \( H(D) \).

Definition 5.1 (McEliece and Lin [18]): Let \( G_{\text{scalar}} \) be a scalar generator matrix for a terminated convolutional code defined by \( G(D) \) [18, 21]. \( G_{\text{scalar}} \) is given by

\[
G_{\text{scalar}} = \begin{pmatrix}
G_0 G_1 \ldots G_L \\
\vdots \\
G_0 \ldots G_{L-1} G_L \\
G_0 G_1 \ldots G_L \\
\vdots \\
G_0 \ldots G_{L-1} G_L \\
\end{pmatrix}.
\]  \hspace{1cm} (65)

The \((L + 1)k_0 \times n_0\) matrix

\[
\hat{G} \triangleq \begin{pmatrix}
G_L \\
G_{L-1} \\
\vdots \\
G_0
\end{pmatrix},
\]  \hspace{1cm} (66)
which repeatedly appears as a vertical slice in \( G_{\text{scalar}} \) except initial and final transient sections, is called the matrix module. Then the trellis module \( T_0 \) for the trellis associated with \( G_{\text{scalar}} \) corresponds to \( \tilde{G} \). If \( G_{\text{scalar}} \) is in minimal-span form, then \( T_0 \) is minimal. The state complexity profile of \( T_0 \) is an \( n_0 \)-tuple consisting of the dimensions of state spaces \( V_i \) \((0 \leq i \leq n_0 - 1)\) of \( T_0 \).

The meaning of obtaining a reciprocal dual encoder is based on the following result \[26, 30, 31\].

**Proposition 5.2 (Tang and Lin \[30\]):** Consider a minimal trellis module of \( G(D) \) and that of an associated reciprocal dual encoder \( \tilde{H}(D) \). Then their state complexity profiles are identical.

Hence, in order to determine whether \( G'(D) \) is reduced or not, we can compute a reciprocal dual encoder \( \tilde{H}'(D) \) associated with \( G'(D) \). In connection with an encoder \( G(D) \) and an associated reciprocal dual encoder \( H(D) \), we have the following.

**Proposition 5.3:** Let \( G_{tb}^N \) be the TBGM associated with \( G(D) \). Then a check matrix corresponding to \( G_{tb}^N \) is obtained as the TBGM (denoted by \( \tilde{H}_{tb}^N \)) associated with a reciprocal dual encoder \( \tilde{H}(D) \).

**Proof:** Let the polynomial expansion of \( H(D) \) be

\[
H(D) = H_0 + H_1 D + \cdots + H_{M-1} D^{M-1} + H_M D^M, \tag{67}
\]

where \( M \) is the memory length of \( H(D) \) and \( H_i \) \((0 \leq i \leq M)\) are \((n_0 - k_0) \times n_0\) matrices. It is known (e.g., \[28\]) that a check matrix corresponding to \( G_{tb}^N \) is given by

\[
H_{tb} = \begin{pmatrix}
H_0 & H_1 & H_2 & \ldots & H_M \\
H_1 & H_0 & H_2 & \ldots & H_M \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
H_{M-1} & \ldots & \ldots & \ldots & H_M \\
H_M & H_{M-1} & \ldots & \ldots & H_M \\
\end{pmatrix}, \tag{68}
\]

with size \((n_0 - k_0)N \times n_0N\). On the other hand, let the polynomial expansion of \( \tilde{H}(D) \) be

\[
\tilde{H}(D) = \tilde{H}_0 + \tilde{H}_1 D + \cdots + \tilde{H}_{M-1} D^{M-1} + \tilde{H}_M D^M. \tag{69}
\]

Then the TBGM associated with \( \tilde{H}(D) \) (denoted by \( \tilde{H}_{tb}^N \)) is defined by

\[
\tilde{H}_{tb}^N = \begin{pmatrix}
\tilde{H}_0 & \tilde{H}_1 & \tilde{H}_2 & \cdots & \tilde{H}_M \\
\tilde{H}_1 & \tilde{H}_0 & \tilde{H}_2 & \cdots & \tilde{H}_M \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\tilde{H}_{M-1} & \ldots & \ldots & \ldots & \tilde{H}_M \\
\tilde{H}_M & \tilde{H}_{M-1} & \ldots & \ldots & \tilde{H}_M \\
\end{pmatrix}, \tag{70}
\]

with size \((n_0 - k_0)N \times n_0N\).

Here take notice of the \( i \)-th \((1 \leq i \leq n_0 - k_0)\) row of

\[
(\tilde{H}_0, \tilde{H}_1, \cdots, \tilde{H}_{M-1}, \tilde{H}_M, 0, \cdots, 0) \subset \tilde{H}_{tb}^N.
\]

We see that the row is identical to the \( i \)-th row of

\[
(H_{\nu^+}, \cdots, H_1, H_0, 0, \cdots, 0, H_M, \cdots, H_{\nu^+ + 1}) \subset H_{tb}^N.
\]

Similarly, the \( i \)-th \((1 \leq i \leq n_0 - k_0)\) row of

\[
(0, \tilde{H}_0, \tilde{H}_1, \cdots, \tilde{H}_{M-1}, \tilde{H}_M, 0, \cdots, 0) \subset \tilde{H}_{tb}^N
\]

is identical to the \( i \)-th row of

\[
(H_{\nu^+ + 1}, \cdots, H_1, H_0, 0, \cdots, 0, H_M, \cdots, H_{\nu^+ + 2}) \subset H_{tb}^N.
\]

Due to the cyclic structures of \( H_{tb}^N \) and \( \tilde{H}_{tb}^N \), similar correspondences hold successively. Hence, \( \tilde{H}_{tb}^N \) is given as a row permutation of \( H_{tb}^N \).

A procedure for computing \( \tilde{H}'(D) \) is obtained based on the above proposition.

**Procedure for computing \( \tilde{H}'(D) \):**
i) Compute a characteristic matrix $Y$ for the dual code $C_{\perp}$ based on $\tilde{H}_{N}^{tb}$, where $Y$ consists of $n_0$ rows and their right cyclic shifts by integer multiple of $n_0$.

ii) Choosing $(n-k)$ rows from $Y$, form $\tilde{H}'$, where $\tilde{H}'$ has the properties:

1) The rows of $\tilde{H}'$ are linearly independent and thus generate $C_{\perp}$.

2) $\tilde{H}'$ consists of $(n_0 - k_0)$ rows and their right cyclic shifts by integer multiple of $n_0$.

iii) Let $\tilde{H}'(D)$ be the polynomial matrix whose TBGM is $\tilde{H}'$.

iv) Note that $G'$ and $\tilde{H}'$ are equivalent to $G_{N}^{tb}$ and $\tilde{H}_{N}^{tb}$, respectively. Hence, $\tilde{H}'$ is a check matrix corresponding to $G'$. Then it follows from Proposition 5.3 that $\tilde{H}'(D)$ is a reciprocal dual encoder associated with $G'(D)$.

v) $Y$ is not necessarily unique. Hence, if necessary, try i) ~ iv) using another characteristic matrix $Y'$ for $C_{\perp}$.

The following is an example where trellis reduction is realized using a reciprocal dual encoder.

**Example 4:** Consider the rate $R = 2/3$ TB convolutional code $C$ of section length $N = 5$ with generator matrix

$$G(D) = \left(\begin{array}{cc}
1 + D & D \\
D^2 & 1 + D + D^2
\end{array}\right)$$

$$= \left(\begin{array}{cc}
1 & 0 \\
0 & 1
\end{array}\right) + \left(\begin{array}{cc}
1 & 0 \\
0 & 1
\end{array}\right) D + \left(\begin{array}{cc}
0 & 0 \\
1 & 0
\end{array}\right) D^2$$

$$\triangleq G_0 + G_1 D + G_2 D^2.$$ 

(72)

Based on the associated TBGM, i.e.,

$$G_{N}^{tb} = \left(\begin{array}{cccc}
G_0 & G_1 & G_2 & 0 \\
0 & G_0 & G_1 & G_2 \\
G_2 & 0 & G_0 & G_1 \\
G_1 & G_2 & 0 & 0
\end{array}\right),$$

(73)

a characteristic matrix $X$ for $C$ is computed as follows:

$$X = \left(\begin{array}{cccc}
1 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{array}\right).$$

(74)

The span list for $X$ is given by

$$T = \{(0, 4), (1, 8), (2, 6), (3, 7), (4, 11), (5, 9), (6, 10), (7, 14), (8, 12), (9, 13), (10, 2), (11, 0), (12, 1), (13, 5), (14, 3)\}.$$ 

(75)

Choosing 10 rows from $X$, let

$$G' = \left(\begin{array}{cccc}
1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right).$$

(76)
The span list for $G'$ is given by

$$S = \{(0, 4), (2, 6), (3, 7), (5, 9), (6, 10), (8, 12), (9, 13), (11, 0), (12, 1), (14, 3)\}. \quad (77)$$

We see that the rows of $G'$ are linearly independent and thus generate $C$, i.e., $G'$ is equivalent to $G'^{tb}_0$. Also, note that $G'$ is the TBGM associated with

$$G'(D) = \left( \begin{array}{ccc} 1 + D & D & 1 \\ D^2 & 0 & 1 + D \end{array} \right) \quad (78)$$

$$= \left( \begin{array}{ccc} 1 & 0 & 1 \\ 0 & 0 & 1 \end{array} \right) + \left( \begin{array}{ccc} 1 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) D + \left( \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 0 & 0 \end{array} \right) D^2$$

$$\triangleq G'_0 + G'_1 D + G'_2 D^2. \quad (79)$$

Hence, the original TB convolutional code is equally represented as a TB convolutional code defined by $G'(D)$.

Observe that the constraint length of $G'(D)$ is $\nu' = 3$ and is equal to that of $G(D)$. Also, notice that the second column of $G'(D)$ has a factor $D$. However, $\nu'$ is not reduced by dividing the column by $D$. In general, it is difficult to tell a possibility of reduction of $G'(D)$ just by looking at its entries. So, we will compute a reciprocal dual encoder $\tilde{H}'(D)$ associated with $G'(D)$.

We begin with a reciprocal dual encoder $\tilde{H}(D)$ associated with $G(D)$. $\tilde{H}(D)$ is given by

$$\tilde{H}(D) = (1 + D + D^2 + D^3, 1 + D + D^3, 1 + D^2 + D^3). \quad (80)$$

Based on $\tilde{H}^{tb}_5$, a characteristic matrix $Y$ for $C^\perp$ is computed as follows:

$$Y = \left( \begin{array}{ccccccccc} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right). \quad (81)$$

The span list for $Y$ is given by

$$\tilde{T} = \{(0, 11), (1, 12), (2, 10), (3, 14), (4, 0), (5, 13), (6, 2), (7, 3), (8, 1), (9, 5), (10, 6), (11, 4), (12, 8), (13, 9), (14, 7)\}. \quad (82)$$

Note that if the span list for $X$ is $T = \{(a_l, b_l), \ l = 1, \cdots, 15\}$, then the span list for $Y$ is given by $\tilde{T} = \{(b_l, a_l), \ l = 1, \cdots, 15\}$. \textbf{10, 13.}

Next, choosing 5 rows from $Y$, let

$$\tilde{H}' = \left( \begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 \end{array} \right) \quad (83)$$

The span list for $\tilde{H}'$ is given by

$$\tilde{S} = \{(2, 10), (5, 13), (8, 1), (11, 4), (14, 7)\}. \quad (84)$$

We see that $\tilde{H}'$ is equivalent to $\tilde{H}^{tb}_5$. Thus $\tilde{H}'$ is a scalar check matrix corresponding to $G'$. Also, note that $\tilde{H}'$ is the TBGM associated with $\tilde{H}'(D) = (D + D^2, D^3, 1)$. We already know that $G'$ is the TBGM associated with $G'(D)$. Hence, by Proposition 5.3, a reciprocal dual encoder associated with $G'(D)$ is given by

$$\tilde{H}'(D) = (D + D^2, D^3, 1). \quad (85)$$
Observe that $\tilde{H}(D) = (D + D^2, D^3, 1)$ has a factor $D$ in the first column and a factor $D^2$ in the second column. Then sweeping these factors out of the corresponding columns, the constraint length of $\tilde{H}(D)$ is reduced to one. This fact implies that the constraint length of $G(\Delta)$ can also be reduced.

In the following, we will show that reduction of $G(\Delta)$ is actually realized. For the purpose, a check matrix corresponding to $G(\Delta)$, i.e.,

$$H'(D) = (D^{-1} + D^{-2}, D^{-3}, 1) \times D^3 = (D + D^2, 1, D^3)$$

(86)

is used.

Let $G(\Delta)$ and $H(\Delta)$ be a generator matrix and a corresponding check matrix for a convolutional code, respectively. In the following, this relation is denoted by $G(\Delta) \iff H(\Delta)$. It is shown [27] that $G(\Delta)$ and $H(\Delta)$ can be reduced simultaneously, if reduction is possible, where the relation $\iff$ is retained in the whole reduction process. We apply the method to our case under consideration.

Step 1: For $G(\Delta)$, add the first row multiplied by $D$ to the second row. By Proposition 4.3, this is a TB-equivalent transformation. As a result, we have

$$G''(D) = \begin{pmatrix} 1 + D & D & 1 \\ D & D^2 & 1 \end{pmatrix} \iff H'(D) = (D + D^2, 1, D^3).$$

Step 2: Divide the second column of $G''(D)$ by $D$, while divide the first and third columns of $H'(D)$ by $D$. Then we have

$$G'''(D) = \begin{pmatrix} 1 + D & 1 & 1 \\ D & D & 1 \end{pmatrix} \iff H''(D) = (1 + D, 1, D^2).$$

Step 3: Multiply the third column of $G'''(D)$ by $D$, while divide the third column of $H''(D)$ by $D$. Then we have

$$G^{(4)}(D) = \begin{pmatrix} 1 + D & 1 & D \\ D & D & D \end{pmatrix} \iff H^{(4)}(D) = (1 + D, 1, D).$$

Step 4: Note that $G^{(4)}(D)$ is not basic [5]. Using an invariant-factor decomposition [5] of $G^{(4)}(D)$, an equivalent basic matrix

$$G^{(5)}(D) = \begin{pmatrix} 1 + D & 1 & D \\ 1 & 1 & 1 \end{pmatrix}$$

(87)

is obtained. Note that $G^{(4)}(D)$ and $G^{(5)}(D)$ are TB-equivalent (cf. Proposition 4.4).

In the above reduction process for $G(\Delta)$, except for TB-equivalent transformations, the second column is divided by $D$, whereas the third column is multiplied by $D$. Accordingly, for each TB path, let us cyclically shift the second component of each branch to the left by one branch and cyclically shift the third component of each branch to the right by one branch. Then the modified TB paths are represented as a TB convolutional code defined by $G^{(5)}(D)$ (see Propositions 4.1 and 4.2). The trellis for $G^{(5)}(D)$ is shown in Fig.3. For example, take an information sequence

$$u = 01 \ 00 \ 01 \ 00 \ 00$$

and the corresponding TB path

$$w = uG_{5}^{tb} = 011 \ 001 \ 110 \ 001 \ 101.$$  

By cyclically shifting the second component of each branch to the left by one branch, and by cyclically shifting the third component of each branch to the right by one branch, we have

$$w_m = 001 \ 011 \ 101 \ 000 \ 111.$$  

We see that $w_m$ is a TB path which starts and ends in state (0) in Fig.3.

Fig. 3. The tail-biting convolutional code defined by $G^{(5)}(D)$ ($N = 5$).

Remark: We remark that in the above argument, it is assumed that $G'$ is equivalent to $G_{5}^{tb}$ (i.e., the equivalence has been checked beforehand). In general, however, $k = k_0N$ is relatively large for high rate codes. Hence, it is preferable that the equivalence of $G'$ and $G_{5}^{tb}$ is derived without checking it beforehand. Actually, the equivalence of $G'$ and $G_{5}^{tb}$ is derived from the equivalence of $H'$ and $H_{5}^{tb}$ using the result of Gluesing-Luerssen and Weaver [10] Theorem IV.3 (see Appendix A).
D. Relation Between Trellis Reduction and Section Length

In the proposed trellis reduction method, the section length $N$ is an important parameter. Actually, the method is effective for TB convolutional codes of short to moderate section length. This is because the span lengths of characteristic generators increase as $N$ grows (see Section III-D). We have already shown that a TB trellis with generator matrix $G(D) = (1 + D + D^2, 1 + D^2)$ can be reduced for the case of $N = 5$. Consider the same trellis. This time, however, $N$ is set to 6. Then $G_6^{tb}$ is given by

$$
G_6^{tb} = 
\begin{pmatrix}
G_0 & G_1 & G_2 & 0 & 0 & 0 \\
0 & G_0 & G_1 & G_2 & 0 & 0 \\
0 & 0 & G_0 & G_1 & G_2 & 0 \\
G_2 & 0 & 0 & 0 & G_0 & G_1 \\
G_1 & G_2 & 0 & 0 & 0 & G_0
\end{pmatrix}
$$

(88)

$$
G_6^{tb} = 
\begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{pmatrix}
$$

(89)

Note that to each generator in $G_6^{tb}$, its span is assigned in the natural manner. Observe that the span lengths of these spans are the same, i.e., 6. A characteristic matrix is computed as follows:

$$
X = 
\begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1
\end{pmatrix}
$$

(90)

Thus the set of basic spans is given by

$$
T_0 = \{ (0, 5), (1, 8) \}.
$$

(91)

With respect to $G'$, there are two cases. When the first row of $X$ is used as a basic generator of $G'$, $G'$ is identical to $G_6^{tb}$. When the second row of $X$ is used as a basic generator of $G'$, the span lengths of rows of $G'$ are 8 and are greater than 6. These facts mean that in either case, trellis reduction is not realized using the proposed method. On the other hand, this example implies that the upper bound for $N$ can be estimated by comparing the span lengths of generators in $G'$ with those of generators in $G_N^{tb}$.

Let $X$ be a characteristic matrix for a TB convolutional code of section length $N$. We already know that the associated span list $T$ consists of the set of basic spans

$$
T_0 = \{ (0, b_0), (1, b_1), \ldots, (n_0 - 1, b_{n_0 - 1}) \}
$$

and $\rho_{in_0}(T_0)$ ($i = 1, 2, \cdots, N - 1$). Also, the sum of span lengths of spans in $T_0$ is given by

$$
\ell = n_0((n_0 - k_0)N + 1).
$$

In the proposed method, $G'$ consists of $k$ generators in $X$. From a span viewpoint, this corresponds to choosing $k_0$ spans from $T_0$. Accordingly, the sum of span lengths of these $k_0$ spans, denoted by $\ell'$, is approximated by

$$
\ell' = (k_0 / n_0) \times n_0((n_0 - k_0)N + 1) = k_0((n_0 - k_0)N + 1).
$$

(92)
On the other hand, consider $G^{tb}_N$, where to each generator, its span is assigned in the natural manner. Then the span list consists of the set of basic spans $\tilde{T}_0$ and $\rho_n(\tilde{T}_0)$ ($i = 1, 2, \ldots, N - 1$). We evaluate the sum of span lengths of spans in $\tilde{T}_0$, denoted by $\hat{\ell}$. Let $\nu_i$ be the degree of the $i$th row of $G(D)$. Here take notice of the first block of $k_0$ rows in $G^{tb}_N$, i.e.,

$$(G_0, G_1, \ldots, G_{L-1}, G_L, 0, \ldots, 0).$$

The span length of the $i$th ($1 \leq i \leq k_0$) row is approximated by $n_0(\nu_i + 1)$. Hence, we have

$$\hat{\ell} \equiv n_0(\nu_1 + 1) + n_0(\nu_2 + 1) + \cdots + n_0(\nu_{k_0} + 1)$$

$$= n_0(\nu + \nu_2 + \cdots + \nu_{k_0}) + n_0k_0$$

$$= n_0(\nu + k_0),$$

(93)

where $\nu \triangleq \nu_1 + \nu_2 + \cdots + \nu_{k_0}$ is the constraint length of $G(D)$. Since trellis reduction is realized in the case where $G'$ consists of generators with short span length, we can take the inequality

$$\ell' \leq \hat{\ell}$$

as a criterion for trellis reduction. That is, we can estimate the upper bound for $N$ using the inequality

$$(\sharp) \quad k_0((n_0 - k_0)N + 1) \leq n_0(\nu + k_0).$$

(94)

For several concrete cases, we will show the condition $(\sharp)$.

1. $R = 1/2$:

$$\ell' \equiv N + 1$$

$$\ell \equiv 2\nu + 2 \quad (\sharp) \quad N \leq 2\nu + 1.$$

2. $R = 1/3$:

$$\ell' \equiv 2N + 1$$

$$\ell \equiv 3\nu + 3 \quad (\sharp) \quad N \leq \lfloor (3/2)\nu \rfloor + 1.$$

3. $R = 2/3$:

$$\ell' \equiv 2N + 2$$

$$\ell \equiv 3\nu + 6 \quad (\sharp) \quad N \leq \lfloor (3/2)\nu \rfloor + 2.$$

We observe that the TB convolutional codes presented in Section V-B all satisfy the condition $(\sharp)$. Also, the rate $R = 2/3$ TB convolutional code discussed in the previous section satisfies the condition $(\sharp)$.

VI. CONCLUSION

In this paper, we have derived several basic properties of a characteristic matrix for a TB convolutional code. We have shown that the characteristic span list consists of some basic spans and their right cyclic shifts. Using the derived results, we have shown that a trellis associated with a given TB convolutional code can be reduced in some cases. As candidates for trellis reduction, we have taken the generator matrices from the tables in [12] Chapter 8 in principle. For example, the rate $R = 1/2$ encoders in Section V-B were chosen from [12] TABLE 8.1. On the other hand, good TB convolutional encoders have been obtained [12, 25]. Here, for a given rate $R = k_0/n_0$, the optimal encoder of memory length $L$ produces the largest minimum distance $d$ for each section length $N$. We have applied the proposed reduction method to some of such encoders (see [12] TABLE 8.19]). As a result, for example, we have obtained $G = (6, 7)$ ($\nu = 2$, $N = 5$) from $G = (50, 64)$ ($\nu = 3$, $N = 5$), where the octal notation for generator matrices is used. Similarly, we have obtained $G = (54, 60)$ ($\nu = 3$, $N = 6$) from $G = (46, 60)$ ($\nu = 4$, $N = 6$). Note that both $G = (6, 7)$ ($\nu = 2$, $N = 5$) and $G = (54, 60)$ ($\nu = 3$, $N = 6$) are listed in the same table.

Finally, we remark that the proposed trellis reduction method depends on the choice of a characteristic matrix for a given convolutional code. Though the number of characteristic matrices to be examined is rather restricted (cf. Proposition 5.1), the method is not fully constructive. Also, a detailed condition that trellis reduction is realized has to be clarified.
We first prove the following.

**Proposition A.1:** Let $G(D)$ be as in Section III. Consider a TB convolutional code $C$ generated by $G_N^A$ and the corresponding dual code $C^\perp$. Let $X$ be a characteristic matrix for $C$ with span list $T$. Also, let $Y$ be a characteristic matrix for $C^\perp$ with span list $\hat{T}$. Let $\tilde{X}$ and $\tilde{Y}$ be submatrices of $X$ and $Y$, respectively, where $\tilde{X}$ consists of $k$ rows in $X$, whereas $\tilde{Y}$ consists of $(n - k)$ rows in $Y$. Denote by $S$ and $\hat{S}$ the span lists for $\tilde{X}$ and $\tilde{Y}$, respectively. Here assume the following:

- Each span in $S$ does not include any spans in $\hat{T}$ except itself.
- $\hat{S}$ satisfies $(b, a) \leftrightarrow (a, b) \notin S$ (that is, $\hat{S}$ consists of the spans in $\hat{T}$ whose reverse is not in $S$).

Then the rows of $\tilde{X}$ and $\tilde{Y}$ are linearly independent and thus generate $C^\perp$. Moreover, $Y$ consists of $(n_0 - k_0)$ rows and their right cyclic shifts by integer multiple of $n_0$.

**Remark:** When $\tilde{X}$ consists of generators in $X$ with short span length, it is probable that the condition ii) holds. Similarly, when $\tilde{Y}$ consists of generators in $Y$ with short span length, it is probable that the condition v) holds.

**Proof:** From ii), it follows that $\tilde{X}$ is common to all the characteristic matrices for $C$. Similarly, from v), it follows that $\tilde{Y}$ is common to all the characteristic matrices for $C^\perp$. Also, by vi), $(\tilde{X}, \tilde{Y})$ is a dual selection of $(X, Y)$ [10] Definition IV.2].

As a result [10] Theorem IV.3, we have

1. rank $\tilde{X} = k \leftrightarrow$ rank $\hat{Y} = n - k$
2. Let rank $\tilde{X} = k$. Then the KV trellises [10] based on $(\tilde{X}, S)$ and $(\hat{Y}, \hat{S})$ are dual to each other.

Let us go back to Example 4. In this example, the code $C$ generated by $G_N^A$ and the dual code $C^\perp$ generated by $H_N^A$, are considered. $X$ is a characteristic matrix for $C$, whereas $Y$ is a characteristic matrix for $C^\perp$. Note that neither $X$ nor $Y$ are unique. These are observed from the relation of inclusion in the associated span lists $T$ and $\hat{T}$, where

\[
T = \{(0, 4], (1, 8], (2, 6], (3, 7], (4, 11], (5, 9], (6, 10], (7, 14], \\
(8, 12], (9, 13], (10, 2], (11, 0], (12, 1], (13, 5], (14, 3)\}
\]

\[
\hat{T} = \{(0, 11], (1, 12], (2, 10], (3, 14], (4, 0], (5, 13], (6, 2], (7, 3], \\
(8, 1], (9, 5], (10, 6], (11, 4], (12, 8], (13, 9], (14, 7)\}.
\]

Next, take notice of the matrices $G'$ and $\hat{H}'$, which are submatrices of $X$ and $Y$, respectively. The corresponding span lists are given by

\[
S = \{(0, 4], (2, 6], (3, 7], (5, 9], (6, 10], \\
(8, 12], (9, 13], (11, 0], (12, 1], (14, 3)\}
\]

\[
\hat{S} = \{(2, 10], (5, 13], (8, 1], (11, 4], (14, 7)\}.
\]

Here note the following:

- Each span in $S$ does not include any spans in $T$ except itself.
- Each span in $\hat{S}$ does not include any spans in $\hat{T}$ except itself.

Moreover, $\hat{S}$ consists of the spans in $\hat{T}$ whose reverse is not in $S$. Actually, by reversing the spans in $\hat{S}$, we have

\[
(10, 2], (13, 5], (1, 8], (4, 11], (7, 14).
\]

We see that these spans are not in $S$.

All these facts show that the conditions in Proposition A.1 are satisfied, when $\tilde{X}$ and $\tilde{Y}$ are replaced by $G'$ and $\hat{H}'$, respectively. Hence, the equivalence of $G'$ and $G_N^A$ is derived.

**Remark:** [10] Theorem IV.3 holds only for a pair $(X, Y)$, where $X$ is a characteristic matrix for $C$ and $Y$ is the corresponding dual one for $C^\perp$ (see [10]). On the other hand, the pair $(X, Y)$ computed above may not be in the duality relation. However, $G'$ is common to all the characteristic matrices for $C$, and $\hat{H}'$ is common to all the characteristic matrices for $C^\perp$ as well. Hence, the theorem can be applied to our case.

**Acknowledgment**

The author would like to thank Prof. Heide Gluesing-Luerssen for valuable comments on the duality of Koetter-Vardy (KV) trellises.
REFERENCES

[1] J. B. Anderson and S. M. Hladik, “An optimal circular Viterbi decoder for the bounded distance criterion,” IEEE Trans. Commun., vol. 50, no. 11, pp. 1736–1742, Nov. 2002.
[2] 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, no. 2, pp. 284–287, March 1974.
[3] A. R. Calderbank, G. D. Forney, Jr., and A. Vardy, “Minimal tail-bitting trellises: The Golay code and more,” IEEE Trans. Inform. Theory, vol. 45, no. 5, pp. 1435–1455, July 1999.
[4] D. Conti and N. Boston, “On the algebraic structure of linear tail-biting trellises,” IEEE Trans. Inform. Theory, vol. 61, no. 5, pp. 2283–2299, May 2015.
[5] G. D. Forney, Jr., “Convolutional codes I: Algebraic structure,” IEEE Trans. Inform. Theory, vol. IT-16, no. 6, pp. 720–738, Nov. 1970.
[6] , “Structural analysis of convolutional codes via dual codes,” IEEE Trans. Inform. Theory, vol. IT-19, no. 4, pp. 512–518, July 1973.
[7] , “Coset codes—Part II: Binary lattices and related codes,” IEEE Trans. Inform. Theory, vol. 34, no. 5, pp. 1152–1187 (Appendix A), Sept. 1988.
[8] , “Dimension/length profiles and trellis complexity of linear block codes,” IEEE Trans. Inform. Theory, vol. 40, no. 6, pp. 1741–1752, Nov. 1994.
[9] H. Gluesing-Luerssen and E. A. Weaver, “Linear tail-biting trellises: Characteristic generators and the BCJR-construction,” IEEE Trans. Inform. Theory, vol. 57, no. 11, pp. 7418–7430, Nov. 2011.
[10] , “Characteristic generators and dualization for tail-biting trellises,” IEEE Trans. Inform. Theory, vol. 59, no. 10, pp. 6597–6610, Oct. 2013.
[11] R. Johannesson and K. S. Zigangirov, Fundamentals of Convolutional Coding, New York: IEEE Press, 1999.
[12] R. Koetter and A. Vardy, “The structure of tail-biting trellises: Minimality and basic principles,” IEEE Trans. Inform. Theory, vol. 49, no. 9, pp. 2081–2105, Sep. 2003.
[13] F. R. Kschischang and V. Zorin, “On the trellis structure of block codes,” IEEE Trans. Inform. Theory, vol. 41, no. 6, pp. 1924–1937, Nov. 1995.
[14] S. Lin and R. Y. Shao, “General structure and construction of tail-biting trellises for linear block codes,” in Proc. ISIT2000, p. 117, June 2000.
[15] H. H. Ma and J. K. Wolf, “On tail-biting convolutional codes,” IEEE Trans. Commun., vol. COM-34, no. 2, pp. 104–111, Feb. 1986.
[16] R. J. McEliece, “On the BCJR trellis for linear block codes,” IEEE Trans. Inform. Theory, vol. 42, no. 4, pp. 1072–1092, July 1996.
[17] R. J. McEliece and W. Lin, “The trellis complexity of convolutional codes,” IEEE Trans. Inform. Theory, vol. 42, no. 6, pp. 1855–1864, Nov. 1996.
[18] D. J. Muder, “Minimal trellises for block codes,” IEEE Trans. Inform. Theory, vol. 34, no. 5, pp. 1049–1053, Sept. 1988.
[19] A. V. Nori and P. Shankar, “Unifying views of tail-biting trellis constructions for linear block codes,” IEEE Trans. Inform. Theory, vol. 52, no. 10, pp. 4431–4443, Oct. 2006.
[20] S. Riedel, “Symbol-by-symbol MAP decoding algorithm for high-rate convolutional codes that use reciprocal dual codes,” IEEE J. Select. Areas Commun., vol. 16, no. 2, pp. 175–185, Feb. 1998.
[21] Y. Shany and Y. Be’ery, “Linear tail-biting trellises, the square-root bound, and applications for Reed-Muller codes,” IEEE Trans. Inform. Theory, vol. 46, no. 4, pp. 1514–1523, July 2000.
[22] R. Y. Shao, S. Lin, and M. P. C. Fossorier, “Two decoding algorithms for tail-biting codes,” IEEE Trans. Commun., vol. 51, no. 10, pp. 1658–1665, Oct. 2003.
[23] G. Solomon and H. C. van Tilborg, “A connection between block and convolutional codes,” SIAM J. Appl. Math., vol. 37, no. 2, pp. 358–369, 1979.
[24] P. Stålhi, J. B. Anderson, and R. Johannesson, “Optimal and near-optimal encoders for short and moderate-length tail-biting trellises,” IEEE Trans. Inform. Theory, vol. 45, no. 7, pp. 2562–2571, Nov. 1999.
[25] M. Tajima, K. Okino, and T. Miyagoshi, “Trellis complexity analysis of convolutional codes using reciprocal dual codes,” in Proc. SITA2004 (Japanese Ed.), pp. 91–94, Dec. 2004.
[26] , “Simultaneous code/error-trellis reduction for convolutional codes using shifted code/error-subsequences,” in Proc. ISIT2011, pp. 2409–2413, Aug. 2011.
[27] M. Tajima and K. Okino, “Error-trellises for tailbiting convolutional codes,” in Proc. ISITA2012, pp. 648–652, Oct. 2012.
[28] M. Tajima, “Trellis reduction for tail-biting convolutional codes using characteristic matrices and cyclically shifted code-subsequences,” IEICE Technical Report (Japanese Ed.), IT2016-108, pp. 61–66, March 2017.
[29] H.-H. Tang and M.-C. Lin, “On (n, n–1) convolutional codes with low trellis complexity,” IEEE Trans. Commun., vol. 50, no. 1, pp. 37–47, Jan. 2002.
[30] H.-H. Tang, M.-C. Lin, and B. F. U. Filho, “Minimal trellis modules and equivalent convolutional codes,” IEEE Trans. Inform. Theory, vol. 52, no. 8, pp. 3738–3746, Aug. 2006.
[31] E. A. Weaver, “Minimality and duality of tail-biting trellises for linear codes,” Ph.D. dissertation, submitted to the University of Kentucky, Lexington, Kentucky, USA, April 2012.
[32] 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.

Masato Tajima (M’86–SM’13) was born in Toyama, Japan, on August 13, 1949. He received the B.E., M.E., and Dr. of Eng. degrees all in electrical engineering from Waseda University, Tokyo, Japan, in 1972, 1974, and 1979, respectively. He joined the Electronics Equipment Laboratory of Toshiba R&D Center in 1979, where he engaged in research and development of channel coding techniques with applications to satellite communication systems. From 1993 to 2006, he was with the Department of Intellectual Information Systems Engineering, Toyama University, first as an Associate Professor, next as a Professor. From 2006 to 2015, he was with the Graduate School of Science and Engineering, University of Toyama, as a Professor. He is currently a Professor Emeritus at University of Toyama. His research interests are in coding theory and its applications.