2012

MINIMALITY AND DUALITY OF TAIL-BITING TRELLISES FOR LINEAR CODES

Elizabeth A. Weaver
University of Kentucky, eaweaver1s@gmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Elizabeth A. Weaver, Student

Dr. Heide Gluesing-Luerssen, Major Professor

Dr. Peter Perry, Director of Graduate Studies

MINIMALITY AND DUALITY OF TAIL-BITING TRELLISES FOR LINEAR CODES

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Elizabeth A. Weaver
Lexington, Kentucky

Director: Dr. Heide Gluesing-Luerssen, Professor of Mathematics
Lexington, Kentucky 2012

ABSTRACT OF DISSERTATION

MINIMALITY AND DUALITY OF TAIL-BITING TRELLISES FOR LINEAR CODES

Codes can be represented by edge-labeled directed graphs called trellises, which are used in decoding with the Viterbi algorithm. We will first examine the well-known product construction for trellises and present an algorithm for recovering the factors of a given trellis. To maximize efficiency, trellises that are minimal in a certain sense are desired. It was shown by Koetter and Vardy that one can produce all minimal tail-biting trellises for a code by looking at a special set of generators for a code. These generators along with a set of spans comprise what is called a characteristic pair, and we will discuss how to determine the number of these pairs for a given code. Finally, we will look at trellis dualization, in which a trellis for a code is used to produce a trellis representing the dual code. The first method we discuss comes naturally with the known BCJR construction. The second, introduced by Forney, is a very general procedure that works for many different types of graphs and is based on dualizing the edge set in a natural way. We call this construction the local dual, and we show the necessary conditions needed for these two different procedures to result in the same dual trellis.

KEYWORDS: linear block codes, tail-biting trellises, characteristic generators, dualization, BCJR-construction.

Author's signature: Elizabeth A. Weaver

Date: April 19, 2012

MINIMALITY AND DUALITY OF TAIL-BITING TRELLISES FOR LINEAR CODES

By
Elizabeth A. Weaver

Director of Dissertation: Heide Gluesing-Luerssen

Director of Graduate Studies: Peter Perry

Date: April 19, 2012

This is dedicated to the RandomMatrix() command in Maple, a source of countless useful examples.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Heide Gluesing-Luerssen for her patience with me over the years. Without her guidance, I would not have made it this far. I would also like to thank my committee members Drs. Nagel, Corso, Klapper, and Marek for their help with my dissertation and defense.

Additionally, I cannot overlook all of the teachers who have challenged me and motivated me to pursue a career in mathematics. Specifically, I'm so grateful that Dr. Lee provided me with invaluable advice regarding preliminary exams, Dr. Ray took me on as a crazy advisee during my undergraduate work, and Dr. Sheets showed me the beauty of algebra.

I am also deeply indebted to my supportive department. I would not have had the time to find a job and finish this dissertation without the assistance of Sheri Rhine. I wouldn't have made it through Real Analysis or my first year of graduate school without Justin Taylor. I would not have my sanity if not for the levity provided by my office mates and Friday night dinner companions Nick, Beth, Chris, Megan, Richie, and Laura. Finally, I am blessed to have Dennis and his love and companionship.

Last, but not least, I must thank my family. I am so lucky to have parents who have supported me unconditionally, grandparents that have provided such encouragement, and a crazy Aunt Debbi to keep things groovy.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1 Introduction

An important aspect of coding theory lies in finding efficient decoding algorithms. While there are many algebraic approaches to this matter, much interest has arisen in the area of search decoding algorithms based on graphical representations for codes, called trellises. These are graphs where the vertices are sorted along a time axis and the codewords appear as paths through the graph. Trellis representations and trellis decoding algorithms were widely used for convolutional codes, but it wasn't until a paper by Bahl, Cocke, Jelinek, and Raviv in 1974, see [1], that their use spread to linear block codes. Because the efficiency of such a search algorithm depends on the complexity of the trellis, constructions that produce trellises that are minimal in a certain sense are desired. In the case of conventional trellises, it has been shown that there exists a unique conventional trellis, up to isomorphism, that simultaneously minimizes every conceivable complexity measure, see [4], [14], and [13]. Many constructions for this trellis exist, and Vardy provides a thorough overview of conventional trellises and their minimal constructions in [16].

Recently, tail-biting trellises, which are trellises that have a circular structure, have gained interest due to the fact that the complexity of a tail-biting trellis for a code may be lower than that of the minimal conventional trellis; however, the situation is much more complicated than that of conventional trellises. Most notably, for a given complexity measure, there may be many non-isomorphic minimal trellises for a code.

In this thesis, we will study various constructions of tail-biting trellises. Throughout we will mainly consider simultaneously minimizing the number of vertices at every time in our trellis. Additionally, we will discuss the relationship between trellises and the dual of a code.

We begin by considering a method for constructing tail-biting trellises called the product construction. This construction was introduced in the context of conventional trellises by Kschischang and Sorokine in [12] and extended to tail-biting trellises by Koetter and Vardy in [11]. To create a trellis for a given code, we begin with a set of generators for the code, and then choose a support-containing interval, called a span, for each of these generators. An elementary trellis is then constructed for each generator, and the structure of this trellis is completely dependent on the chosen span. These elementary trellises are then combined using a direct product-like construction to obtain a trellis for the given code.

The product construction results in trellises that are linear in structure, and in [10], Koetter and Vardy present the very important result that every trellis with such a linear structure is in fact a product trellis. This result prompts the question, if we have a linear trellis, how can we recover the elementary trellis factors used in its construction? In chapter 4, we show that the spans of the elementary trellis factors completely determine the structural isomorphism class of a trellis. A similar result has been derived independently and at the same time by Conti and Boston in [3]. Additionally, the spans of the elementary trellis factors determine the structure of the cycles of all of the codewords represented by the trellis. Using this fact, we present an algorithm that returns the spans of the elementary factors for a given linear trellis and discuss how one can recover a set of generating codewords for the trellis as well.

Because these generating spans play such a large role in the complexity of a product trellis, it is natural to turn attention to them in the search for minimal tail-biting trellises. In [11], Koetter and Vardy show that every code has a set of shortest spans, which along with corresponding codewords, form what they call a characteristic pair for a code. They go on to prove that every minimal tail-biting trellis for a code is a product of elementary trellises based on these generators and characteristic spans. It is important to note, however, that not every product trellis based on these generators and spans is minimal. Despite this fact, this class of trellises, which we refer to as KV-trellises, has many nice properties that we will investigate in the remainder of the thesis.

While Koetter and Vardy show that the set of characteristic spans for a code is unique, there may be many different sets of corresponding codewords that we can take to form a characteristic pair for a code. In Chapter 6, we show that the total number of characteristic pairs for a code is dependent on the number of containments between the characteristic spans themselves. This leads to many useful results. For example, using a lemma by Koetter and Vardy, we prove that a code and its dual have the same number of characteristic pairs. We also discuss under what condition a code has a unique characteristic matrix, and show that the class of MDS codes has this property. Finally, we illustrate that a single characteristic pair may not generate every minimal tail-biting trellis for a code. We conclude this chapter by showing that a single characteristic pair will produce every minimal trellis, up to structural isomorphism. All of this clarifies some subtle inconsistencies contained in [11].

Another construction of tail-biting trellises is provided in [15]. It is an extension of the well-known BCJR-construction for conventional trellises. We introduce both general and span-based versions of the tail-biting BCJR construction, as well as discuss several important properties of these trellises. For instance, a KV-trellis is isomorphic to its corresponding span-based BCJR-trellis, see also [8]. This construction is also a very important tool for trellis dualization in the following chapter.

Next we examine the process of trellis dualization, that is a process by which a trellis for a code is transformed into a trellis that represents the dual code. We first present a dualization that comes naturally with the BCJR construction. We also introduce a procedure we call local dualization, which is a specialization of the local dualization introduced by Forney in [5] in the case of general normal graphs. Often these dual trellises possess undesirable properties, and many times these two dualization procedures result in non-isomorphic trellises. For a reasonably nice class of trellises, we provide a list of necessary and sufficient conditions for the two dualizations being isomorphic. One of these conditions is an easy-to-check criterion in terms of the primal trellis; the other conditions show in essence that the local dual behaves well if and only if it is isomorphic to the BCJR-dual. Our examples of local duals lacking certain basic properties, which first appeared in [7], sparked the interest of Forney. In a joint collaboration, he and Gluesing-Luerssen studied how intrinsic graph realization properties behave under local dualization, see [6].

We close the thesis by showing that in the case of KV-trellises, not only do these dualizations lead to isomorphic trellises, but the dual trellis is a KV-trellis for the dual code. While this is also a consequence of the main result in [7, Chapter IV], we arrive at this conclusion in a much simpler fashion.

Chapter 2 Basic Notions

In this section we will introduce the basic definitions and properties for tail-biting trellises. Throughout, let $\mathbb{F} = \mathbb{F}_q$ be a finite field.

A **tail-biting trellis** $T = (V, E)$ of depth n over $\mathbb{F}$ is an edge-labeled directed graph where E is the edge set, and the vertex set V can be decomposed into a union of n disjoint subsets $V = V_0 \cup \ldots \cup V_{n-1}$ such that every edge that begins at a vertex in V_i ends at a vertex in $V_{(i+1) \bmod n}$ for $i = 0, 1, \ldots, n-1$. The edges are labeled with elements from $\mathbb{F}$, and the edge set E can also be partitioned into $E = E_0 \cup \ldots \cup E_{n-1}$ where E_i contains the edges between vertices in V_i and vertices in V_{i+1}. We will denote edges in E_i as triples consisting of the starting vertex, the edge label, and the ending vertex. Thus, for $i = 0, 1, \ldots, n-1$ we have $E_i := \{(v, a, \hat{v}) :$ there exists an edge in T from $v \in V_i$ to $\hat{v} \in V_{i+1}$ with label $a\} \subseteq V_i \times \mathbb{F} \times V_{i+1}$. We will identify the time axis $\mathcal{I} := \{0, 1, \ldots, n-1\}$ with $\mathbb{Z}_n$. We will often refer to V_i as the state space of T at time i, and E_i as the transition space at time i.

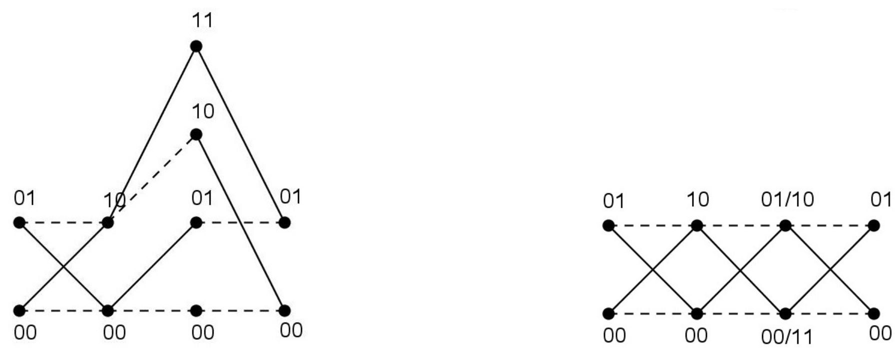
A **cycle** in T is a closed path of length n, that is, it begins and ends at the same vertex in $V_0 = V_n$. When $|V_0| = 1$, we call the trellis T **conventional**. (We will not deal with conventional trellises where $|V_0| > 1$ in this thesis.) We say that T is **trim** if every vertex in T lies on a cycle, T is **edge-reduced** if every edge in T is part of a cycle, and T is **reduced** if it is both trim and edge-reduced. If any two edges starting or ending at the same vertex have distinct labels, we say that T is **biproper**.

We say that T is **linear** if each state space V_i is a vector space and each transition space E_i is a subspace of $V_i \times \mathbb{F} \times V_{i+1}$. The set $\mathcal{S}(T) := \{(v_0, c_0, v_1, \ldots, v_{n-1}, c_{n-1}, v_0) : (v_i, c_i, v_{i+1}) \in E_i \text{ for } i \in \mathcal{I}\}$ is called the **label code** of T. If T is linear and reduced, then $\mathcal{S}(T)$ is a subspace of $V_0 \times \mathbb{F} \times V_1 \times \ldots \times V_{n-1} \times \mathbb{F} \times V_0$. The set $\mathcal{C}(T) := \{(c_0, c_1, \ldots, c_{n-1}) : \exists (v_0, c_0, v_1, \ldots, v_{n-1}, c_{n-1}, v_0) \in \mathcal{S}(T)\}$ is called the **edge-label code** of T. We say that T represents the code $\mathcal{C}$ if $\mathcal{C}(T) = \mathcal{C}$. In the case that there is a bijection between the cycles of T and the codewords in $\mathcal{C}(T)$, we say that T is **one-to-one**. The **state complexity profile**, or SCP, of a linear trellis T is defined as $\mathrm{SCP}(T) = (\xi_0, ..., \xi_{n-1})$ where $\xi_i = \dim V_i$ for $i = 0, \ldots, n-1$. Similarly, the **edge complexity profile**, or ECP, of a linear trellis T is $\mathrm{ECP}(T) = (\epsilon_0, \ldots, \epsilon_{n-1})$ where $\epsilon_i = \dim E_i$ for $i = 0, \ldots, n-1$.

Two trellises $T = (V, E)$ and $T' = (V', E')$ are **isomorphic** if there exists a bijection $\phi : V \to V'$ such that $\phi(V_i) = V_i'$ for $i \in \mathcal{I}$ and if $(v, a, u) \in E_i$ then $(\phi(v), a, \phi(u)) \in E_i'$ for $i \in \mathcal{I}$. If T and T' are linear, we also require $\phi|_{V_i}$ to be isomorphisms. We say that two trellises T and T' are **structurally isomorphic**, denoted by $\cong_S$, if there exists a bijection $\phi : V \to V'$ such that $\phi(V_i) = V_i'$ for $i \in \mathcal{I}$ and if the number of edges from $v \in V_i$ to $u \in V_{i+1}$ equals the number of edges from $\phi(v) \in V_i'$ to $\phi(u) \in V_{i+1}'$.

A trellis $T = (V, E)$ is said to be **mergeable** if there exist $u, v \in V_i$ for some $i \in \mathcal{I}$ that can be merged, that is replaced by a single vertex that inherits the edges incident to both u and v, without changing $\mathcal{C}(T)$. If no vertices in T can be merged, we say that T is **nonmergeable**.

Example 2.1. *Throughout the thesis, we will use the convention that when trellises represent binary codes, we use dashed lines to represent edges with label 0 and solid edges to represent edges with label 1. Trellis T, shown in Figure 2.1, is reduced, biproper, and linear. It has an SCP of (1,1,2) and an ECP of (2,2,2) and represents the code $\mathcal{C} = \{000, 011, 110, 101\}$. This trellis is mergeable at time $i = 2$ since merging the vertices 01 and 10 does not change the trellis code of T. The vertices 00 and 11 in V_2 may also be merged. The trellis T' shows the original trellis after these merges. Also note that while the trellis T is one-to-one, the trellis T' is not since there are two cycles that represent the codeword 000.*



[T] [T']

Figure 2.1: A mergeable trellis T over $\mathbb{F}_2$, and the trellis T' obtained after merging.

A trellis T is **weakly connected** if there exists a path (possibly undirected) between any two vertices in T. In other words, T is weakly connected if the underlying undirected graph of T is connected. For tail-biting trellises, the characterization of weak connectedness in the following proposition is very useful.

Proposition 2.2. *For a linear and trim trellis T, the following are equivalent.*

(i) T is weakly connected.

(ii) For all $i \in \mathcal{I}$, the trellis T has the property that for all $v \in V_i$ there exists a path of length n starting at v and ending at $0 \in V_i$.

(iii) For some $i \in \mathcal{I}$, the trellis T has the property that for all $v \in V_i$ there exists a path of length n starting at v and ending at $0 \in V_i$.

Proof. "(i) $\Rightarrow$ (ii)" Suppose that T is weakly connected. Because a tail-biting trellis is a circular object, we can cyclically shift the trellis without changing the structual properties. Thus, without loss of generality, we will consider $i = 0$. Let $v \in V_0$. By assumption, there exists an undirected path from v to $0 \in V_0$. Since T is trim, every vertex in T lies on a cycle. Therefore, whenever this undirected path wraps around the trellis at time 0, we may insert a suitable cycle. In this way the undirected path from v to 0 can be made to zigzag through the trellis without ever wrapping around the circular structure of the tail-biting trellis. Then by [10, Lemma 6.8] there exists a directed path of length n from v to $0 \in V_0 = V_n$. "(ii) $\Rightarrow$ (iii)" is clear. "(iii) $\Rightarrow$ (i)" As above, without loss of generality, we can assume that T has the property given in (iii) for $i = 0$. Thus V_0 is a subset of one of the connected components of T. Since T is trim, every vertex in T lies on a cycle. Since every cycle passes through V_0, the trellis T has one connected component. □

While there are many notions of trellis minimality in use, in this thesis we will focus on the following definition.

Definition 2.3. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear block code and $T = (V, E)$ be a linear trellis for $\mathcal{C}$. Then T is called minimal if there exists no linear trellis $T' = (V', E')$ for $\mathcal{C}$ such that $|V_i'| \le |V_i|$ for all $i \in \mathcal{I}$ and $|V_j'| < |V_j|$ for some $j \in \mathcal{I}$.*

It is well established that for conventional trellises, there is a unique minimal conventional trellis, up to isomorphism. More specifically a linear and reduced conventional trellis is biproper if and only if it is nonmergeable if and only if it is minimal (see [16]); however, the situation for tail-biting trellises is more complicated. In [11], Koetter and Vardy provide the following chain of inclusions

$$\{\text{minimal trellises}\} \subsetneq \{\text{nonmergeable trellises}\} \subsetneq \{\text{biproper trellises}\}. \tag{2.1}$$

Finally, we fix the following notation pertaining to the code under consideration and its representation. Throughout, let

$$\mathcal{C} = \operatorname{im} G = \ker H^\top \subseteq \mathbb{F}^n \text{ be a } k\text{-dimensional code with support } \mathcal{I} = \{0, \ldots, n-1\}, \tag{2.2}$$

where the latter means that for each $j \in \mathcal{I}$ there exists a codeword $(c_0, \ldots, c_{n-1}) \in \mathcal{C}$ such that $c_j \neq 0$. Here, $\operatorname{im} M := \{\alpha M \mid \alpha \in \mathbb{F}^m\}$ and $\ker M := \{\alpha \in \mathbb{F}^m \mid \alpha M = 0\}$ denote the row space and left kernel of the matrix $M \in \mathbb{F}^{m \times n}$, respectively. We assume $G \in \mathbb{F}^{r \times n}$, hence $\operatorname{rk} G = k \le r$, and will explicitly state when $r = k$ and thus G is a full row rank encoder matrix. Throughout, $H \in \mathbb{F}^{(n-k) \times n}$ is a full row rank parity check matrix. Furthermore, we fix the notation

$$G = (g_{lj})_{\substack{l=1,\ldots,r \\ j=0,\ldots,n-1}} = \begin{pmatrix} G_0^\top & \ldots & G_{n-1}^\top \end{pmatrix} \in \mathbb{F}^{r \times n} \quad \text{and} \quad H^\top = \begin{pmatrix} H_0 \\ \vdots \\ H_{n-1} \end{pmatrix} \in \mathbb{F}^{n \times (n-k)}, \tag{2.3}$$

where $G_j^\top \in \mathbb{F}^r$ and $H_j^\top \in \mathbb{F}^{n-k}$ are the columns of G and H, respectively. As for the matrices G and H above, we will use the notation $M_j^\top$ for the j-th column of the matrix M and we will employ the (Maple) notation $\operatorname{row}(M, l)$ for the l-th row of M. It will also be convenient to have a notion for the indicator function of a subset $\mathcal{A} \subseteq \mathcal{I}$. Thus, we define $I^{\mathcal{A}} \in \mathbb{F}^n$ as the vector with entries $I_j^{\mathcal{A}} = 1$ if $j \in \mathcal{A}$ and $I_j^{\mathcal{A}} = 0$ otherwise.

Chapter 3 The Product Construction and Product Trellises

In this chapter we will discuss a method of creating trellises known as the product construction. This was introduced in the case of conventional trellises in [12] and extended to tail-biting trellises in [2] and [11]. We will begin by defining elementary trellises, which represent one-dimensional codes and have a structure based on a given parameter called a span. We then combine these trellises through a type of direct product type construction to obtain a trellis representing a larger code. After investigating properties of these product trellises, we conclude with the important result by Koetter and Vardy, introduced in [10], that every linear and reduced trellis is a product trellis.

Throughout we will make use of the following interval notation. For $a,\, b \in \mathcal{I}$ we define

$$[a,\, b] := \begin{cases} \{a, a+1, \ldots, b\}, & \text{if } a \leq b, \\ \{a, a+1, \ldots, n-1, 0, 1, \ldots, b\}, & \text{if } a > b, \end{cases}$$

and $(a,\, b] := [a, b]\backslash\{a\}$. In the case where $a \leq b$, we say that the interval is conventional, and if $a > b$, the interval is called circular.

Definition 3.1. *For a vector $c = (c_0, \ldots, c_{n-1}) \in \mathbb{F}^n\backslash\{0\}$ we call the interval $(a,\, b]$ a span of c if $c_j = 0$ for $j \notin [a,\, b]$ (that, is $[a,\, b]$ contains the support of c). We call $(a,\, b]$ a proper span of c if $(a,\, b]$ is a span of c and $c_a \neq 0 \neq c_b$. Although the entire time axis cannot be expressed as a half-open interval, we also allow $\mathcal{I}$ to be a span (but not a proper span) of c. We will take the closure of a span $(a, b]$ to be $\overline{(a, b]} = [a, b]$. Note that $\overline{\mathcal{I}} = \mathcal{I}$.*

While a vector can have many different spans, it will always have a unique conventional proper span.

As we will see in the next definition, a span for a given vector can be used to construct a trellis for the code represented by the image of the vector.

Definition 3.2. *[11, p. 2089] Let s be a span of the nonzero vector $c = (c_0, \ldots, c_{n-1}) \in \mathbb{F}^n$. The elementary trellis for the pair $\big(c, s\big)$ is defined as $T_{c,s} := (V, E)$, where the state spaces and transition spaces are given by $V_i = \operatorname{im}(I_i^s) \subseteq \mathbb{F}$ and $E_i = \operatorname{im}(I_i^s, c_i, I_{i+1}^s) \subseteq V_i \times \mathbb{F} \times V_{i+1}$. Recall that I^s is the indicator function of the set s, and thus $I_j^s = 1$ if $j \in s$ and $I_j^s = 0$ otherwise. Precisely, if $s = (a, b]$ and $a = b$, then*

$$V_i := \{0\} \text{ for all } i \in \mathcal{I},$$
$$E_i := \begin{cases} \{(0, 0, 0)\}, & \text{if } i \neq a, \\ \{(0, \alpha c_i, 0) \mid \alpha \in \mathbb{F}\}, & \text{if } i = a; \end{cases}$$

If $s = (a, b]$ *and* $a \neq b$*, then*

$$V_i := \begin{cases} 0, & \text{if } i \notin (a,\, b], \\ \mathbb{F}, & \text{if } i \in (a,\, b], \end{cases}$$

$$E_i := \begin{cases} \{(0,0,0)\}, & \text{if } i \notin [a,\, b], \\ \{(0, \alpha c_i, \alpha) \mid \alpha \in \mathbb{F}\}, & \text{if } i = a, \\ \{(\alpha, \alpha c_i, 0) \mid \alpha \in \mathbb{F}\}, & \text{if } i = b, \\ \{(\alpha, \alpha c_i, \alpha) \mid \alpha \in \mathbb{F}\}, & \text{if } i \in (a, b-1]. \end{cases}$$

Finally, if $s = \mathcal{I}$*, then*

$$V_i := \mathbb{F} \text{ for all } i \in \mathcal{I},$$
$$E_i := \{(\alpha, \alpha c_i, \alpha) \mid \alpha \in \mathbb{F}\} \text{ for all } i \in \mathcal{I}.$$

Remark 3.3. *Let* s *be a span of* $c \in \mathbb{F}^n \backslash \{0\}$*. Then* $T_{c,s}$ *is a one-to-one, linear, and reduced trellis with* $SCP(T_{c,s}) = I^s$ *and* $ECP(T_{c,s}) = I^{\overline{s}}$*. The trellis is conventional if and only if* s *is a conventional span, and the trellis is nonmergeable (and hence biproper) if and only if* s *is a proper span of* c*.*
Moreover, the trellis is weakly connected if and only if the chosen span for c *is not* $\mathcal{I}$*. This follows immediately from the definition of the transition spaces* E_i *and the fact that when* $s \neq \mathcal{I}$ *we know that* $V_i = \{0\}$ *for at least one time index* i*.*

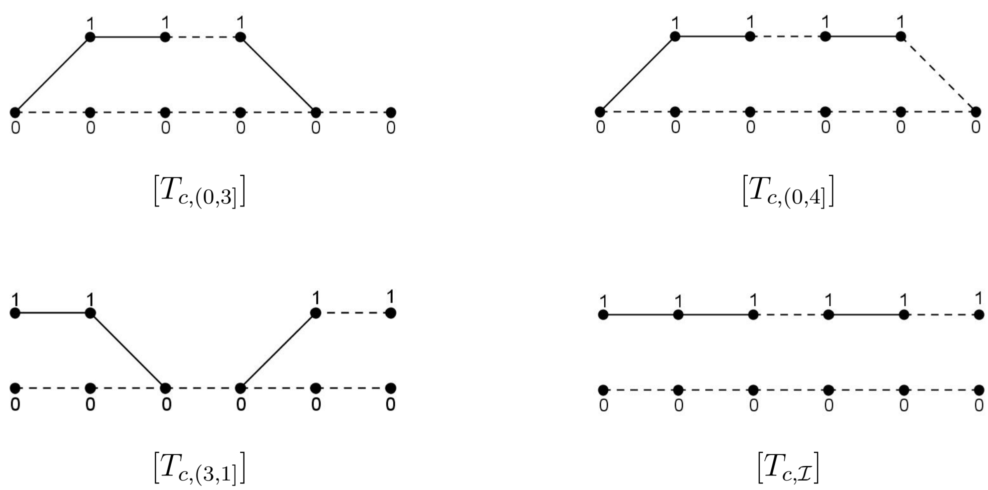


Figure 3.1: Elementary trellises for the vector $c = (1, 1, 0, 1, 0)$.

Example 3.4. *Consider the vector* $c = (1, 1, 0, 1, 0) \subseteq \mathbb{F}_2^5$*. It has two possible conventional spans,* $(0, 3]$ *and* $(0, 4]$*, and of these only* $(0, 3]$ *is a proper span. This vector also has many possible circular spans, including* $(1, 0]$*,* $(3, 1]$*, and* $(3, 2]$*. Of these spans, the first two are proper. Any of these spans, as well as* $\mathcal{I}$*, can be used to construct an elementary trellis for the code generated by* c*, and several of these elementary trellises are pictured in Fig.3.1. From Fig. 3.1, one can see that* $SCP(T_{c,(0,4]}) = (0, 1, 1, 1, 1)$

which agrees with $I^{(0,4]}$*, and ECP*$(T_{c,(0,4]}) = (1,1,1,1,1) = \overline{I^{(0,4]}}$ *as well. It is also worth noting that since* $(0,4]$ *is a conventional interval,* $T_{c,(0,4]}$ *is a conventional trellis; however, since* $(0,4]$ *is not a proper span for* c*, this trellis is mergeable at time* $i=4$*, and one can see that it is not biproper at time* $i=4$*.*

An important trellis construction is the product construction, described in the following proposition. We can use this procedure to take smaller trellises and use them to construct trellises for larger codes. Most often, we will construct elementary trellises for the generators of a code and take their product to get a trellis that represents the entire code.

Proposition 3.5. *Let* $T=(V,E)$ *and* $T'=(V',E')$ *be linear and reduced trellises of depth* n *over* $\mathbb{F}$*, and let* $\mathcal{C}=\mathcal{C}(T)$ *and* $\mathcal{C}'=\mathcal{C}(T')$*. Then the product* $\hat{T} := T \times T'$ *is defined as the trellis* $\hat{T} = (\hat{V}, \hat{E})$*, where* $\hat{V}_i := V_i \times V_i'$ *and* $\hat{E}_i = \big\{\big((v,v'), a + a', (w,w')\big) \,\big|\, (v,a,w) \in E_i, (v',a',w') \in E_i'\big\}$ *for all* $i \in \mathcal{I}$*. The product trellis* $\hat{T}$ *is linear and reduced and satisfies* $\mathcal{C}(\hat{T}) = \mathcal{C} + \mathcal{C}'$*. Moreover,*

(a) $\dim \hat{V}_i = \dim V_i + \dim V_i'$ *for all* $i \in \mathcal{I}$*.*

(b) If $\mathcal{C} \cap \mathcal{C}'$ *does not contain a codeword of weight* 1*, then* $\dim \hat{E}_i = \dim E_i + \dim E_i'$ *for all* $i \in \mathcal{I}$*.*

(c) If $\hat{T}$ *is one-to-one, then* T *and* T' *are one-to-one. The converse is true when* $\mathcal{C} \cap \mathcal{C}' = \{0\}$*.*

(d) $\hat{T}$ *is weakly connected if and only if* T *and* T' *are weakly connected.*

(e) If $\hat{T}$ *is nonmergeable, then* T *and* T' *are nonmergeable.*

(f) If $\hat{T}$ *is biproper, then* T *and* T' *are biproper.*

Proof. (a) This follows directly from the product construction.

(b) The statement $\dim \hat{E}_i = \dim E_i + \dim E_i'$ follows easily from the construction as long as we verify that no two products of edges from E_i and E_i' result in the same edge in $\hat{E}_i$. Suppose that $((v_1,v_1'), a+a', (v_2,v_2')) = ((w_1,w_1'), b+b', (w_2,w_2')) \in \hat{E}_i$. Then we must have that $v_i = w_i$ and $v_i' = w_i'$ for $i=1,2$. Thus there exist edges $(0, a-b, 0) \in E_i$ and $(0, a'-b', 0) \in E_i'$. Therefore, both T and T' have cycles that represent multiples of the i-th standard basis vector. By linearity, this implies that $\mathcal{C} \cap \mathcal{C}'$ contains a codeword of weight 1.

(c) We will prove the first statement by showing that the contrapositive is true. Suppose that T is not one-to-one. Since T is linear, we may assume without loss of generality that there are two distinct cycles in T that correspond to the zero codeword, say $(v_0, 0, v_1, \ldots, v_{n-1}, 0, v_0)$ and the all-zero cycle. Note that the all-zero cycle is also a cycle in T'. Thus the product construction yields the all-zero cycle in $\hat{T}$ as well as the cycle $((v_0,0), 0, (v_1,0), \ldots, (v_{n-1},0), 0, (v_0,0))$. This implies that $\hat{T}$ is not one-to-one. For the second statement we will also employ the contrapositive. Suppose that T and T' are both one-to-one, but $\hat{T}$ is not one-to-one.

Then, without loss of generality, there are two distinct cycles in $\hat{T}$ that correspond to the zero codeword, say $((v_0, v_0'), 0, (v_1, v_1'), \ldots, (v_{n-1}, v_{n-1}'), 0, (v_0, v_0'))$ and the all-zero cycle. Thus, there must exist cycles $(v_0, a_0, v_1, ..., v_{n-1}, a_{n-1}, v_0)$ in T and $(v_0', a_0', v_1', ..., v_{n-1}', a_{n-1}', v_0')$ in T' such that $a_i + a_i' = 0$ and v_i and v_i' are not both zero for all $i \in \mathcal{I}$. Hence, we have $(a_0, ..., a_{n-1}) = (-a_0', ..., -a_{n-1}') \in \mathcal{C} \cap \mathcal{C}'$. Since T is one-to-one, and at least one v_i is non-zero, we know that $(a_0, ..., a_{n-1}) \neq 0$ and therefore $\mathcal{C} \cap \mathcal{C}' \neq \{0\}$.

(d) This follows easily from Proposition 2.2 and the fact that every path in $\hat{T}$ is the product of a path in T and a path in T'.

(e) We will use the contrapositive to prove this statement. Suppose that T is mergeable, that is there exist $u, v \in V_j$ for some $j \in \mathcal{I}$ such that merging u and v does not create any codewords that are not in $\mathcal{C}(T)$. Since every cycle in $\hat{T}$ is the product of a cycle in T and a cycle in T', it follows that the vertices $(u, w'), (v, w') \in \hat{V}_j$, for all $w' \in V_j'$, can be merged without creating codewords that are not in $\mathcal{C}(\hat{T})$.

(f) We will prove this statement by showing that the contrapositive is true. Suppose that T is not biproper. Then there exist edges $(v, a, w), (v, a, w') \in E_j$ for some $j \in \mathcal{I}$. Taking the product of these edges with $(0, 0, 0) \in E_j'$ results in the edges $((v, 0), a, (w, 0)), ((v, 0), a, (w', 0)) \in \hat{E}_j$. Hence $\hat{T}$ is not biproper.

□

Example 3.6. *Let* $\mathcal{C} = \operatorname{im} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \operatorname{im} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \subseteq \mathbb{F}_2^3$. *We can construct a trellis that represents* $\mathcal{C}$ *by taking the product of elementary trellises representing* c_1 *and* c_2*. Specifically, if we take* $T_{c_1,(0,2]}$ *and* $T_{c_2,(1,0]}$ *as in Figure 3.2, we can use the product construction to build the trellis* T *for* $\mathcal{C}$*. As in Proposition 3.5, the figure shows that* $SCP(T) = SCP(T_{c_1,(0,2]}) + SCP(T_{c_2,(1,0]}) = (1, 1, 2)$*. Also, since the codes generated by* c_1 *and* c_2 *have a trivial intersection, we can obtain the* $ECP(T) = (2, 2, 2)$ *by adding up the ECPs of* $T_{c_1,(0,2]}$ *and* $T_{c_2,(1,0]}$*. It is also important to note that while the product of two one-to-one trellises is again one-to-one when the intersection of the represented codes is trivial, the same cannot be said for nonmergeability. In this case, both of the elementary trellises are nonmergeable since they are based on proper spans for the given codewords; however, the resulting product trellis is mergeable at time* $i = 2$*, and this merging results in a trellis that is not one-to-one.*

For convenience, we will use the following notation.

Definition 3.7. *Let* $\mathcal{C} = \operatorname{im} G$*, where* $G \in \mathbb{F}^{r \times n}$ *has no zero rows, and let* $\mathcal{S} := \left[s_l,\, l = 1, \ldots, r\right]$ *be a span list for* G*, that is,* s_l *is a span for the row* g_l*,* $l = 1, \ldots, r$*. Then the product trellis* $T_{g_1,s_1} \times \ldots \times T_{g_r,s_r}$ *is denoted by* $T_{G,\mathcal{S}}$*.*

Remark 3.8. *(a) Suppose that a product trellis* T *has factors* T_{g_1,s_1} *and* T_{g_2,s_2} *where* s_1 *and* s_2 *are the same empty span, that is a span of the form* $(a, a]$*. It is easy to show that the removal of either one (but not both) of these factors leaves* T *unchanged. Thus throughout the thesis, we will always assume that no two*

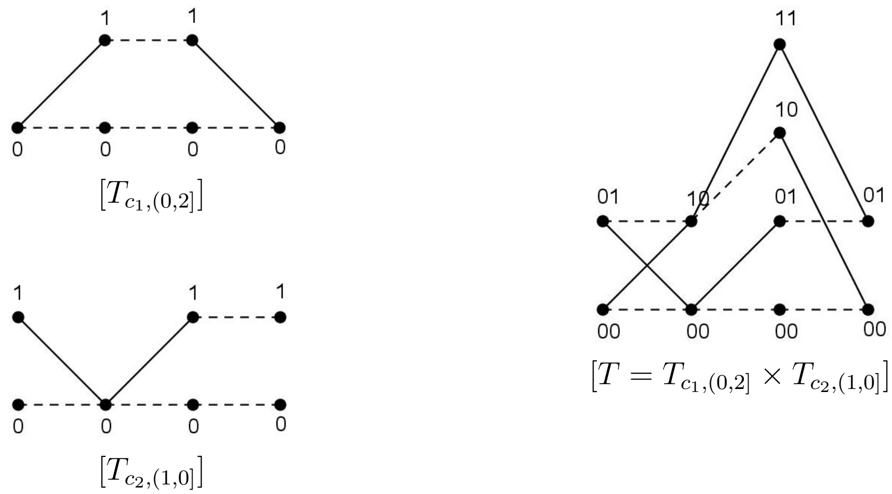


Figure 3.2: An example of the product construction.

elementary trellises in a given product trellis will be based on the same empty span.

(b) *Since the elementary trellis $T_{g,s}$ is biproper if and only if s is a proper span of g, it follows from Proposition 3.5, that if $T_{G,\mathcal{S}}$ is biproper, then $\mathcal{S}$ consists of proper spans for the rows of G. Similarly, $T_{G,\mathcal{S}}$ is weakly connected if and only if none of the spans in $\mathcal{S}$ is the time axis $\mathcal{I}$. It is also easy to check that if $T_{G,\mathcal{S}}$ is biproper, then the starting points of the spans not equal to $\mathcal{I}$ in $\mathcal{S}$ are distinct and the same is true for the ending points.*

Theorem 3.9. *[8, Thm. III.6], [7, Prop. 11.2] Let $T := T_{G,\mathcal{S}}$ be as in Definition 3.7. Then T is a linear and reduced trellis representing the code $\mathcal{C}$, that is, $\mathcal{C}(T) = \mathcal{C}$. Moreover, we have the following.*

1. *The state spaces of T are given by $V_i = \operatorname{im} M_i$, where*

$$\left.\begin{aligned} M_i &= \begin{pmatrix} \mu_i^1 & & \\ & \ddots & \\ & & \mu_i^r \end{pmatrix} \in \mathbb{F}^{r\times r}, \\ \mu_i^l &= \begin{cases} 1, & \text{if } i \in s_l, \\ 0, & \text{if } i \notin s_l. \end{cases} \end{aligned}\right\} \tag{3.1}$$

   The transition spaces are given by $E_i = \operatorname{im}(M_i, G_i^\top, M_{i+1})$ for $i \in \mathcal{I}$.

2. *T is one-to-one if and only if $\operatorname{rk} G = r$.*

3. *SCP(T)=$\sum_{l=1}^r I^{s_l}$ and ECP(T)=$\sum_{l=1}^r I^{\overline{s_l}}$.*

4. *Suppose $s_l = (a_l, b_l]$ for $l = 1, ..., r$, that is, no span in $\mathcal{S}$ is $\mathcal{I}$, and define $\mathcal{A} = \{a_1, \dots, a_r\}$ and $\mathcal{B} = \{b_1, \dots, b_r\}$.*

(a) If $a_1, \ldots, a_r$ are distinct, then ECP(T) = SCP(T) + $I^{\mathcal{A}}$.

(b) If $b_1, \ldots, b_r$ are distinct, then ECP(T) = σ(SCP(T)) + $I^{\mathcal{B}}$, where $\sigma(s_0, \ldots, s_{n-1}) := (s_1, \ldots, s_{n-1}, s_0)$.

(c) Let $a_1, \ldots, a_r$ be distinct and $b_1, \ldots, b_r$ be distinct. Then

$$s_{i+1} = \begin{cases} s_i, & \text{if } i \in \mathcal{A} \cap \mathcal{B} \text{ or } i \notin \mathcal{A} \cup \mathcal{B}, \\ s_i + 1, & \text{if } i \in \mathcal{A} \setminus \mathcal{B}, \\ s_i - 1, & \text{if } i \in \mathcal{B} \setminus \mathcal{A}, \end{cases}$$

and if each s_l is a proper span, the trellis T is biproper.

Proof. Everything, with the exception of the biproperness, follows directly from the properties of elementary trellises and product trellises found in Remark 3.3 and Proposition 3.5.
For the biproperness, first observe that $\mathrm{rk}(M_i, G_i^\top) \leq \mathrm{rk}(M_i) + 1$. Now, by (3.1) and the fact that s_l is proper, we know that $g_{l,a_l} \neq 0$ and $\mu^l_{a_l} = 0$. We also have that $g_{l,j} = \mu^l_j g_{l,j}$ for $j \neq a_l$. Thus, we obtain that $G_i^\top$ is not in the column space of M_i exactly when $i \in \mathcal{A}$, and we have the equivalence $\big[i \in \mathcal{A} \Longleftrightarrow \mathrm{rk}(M_i, G_i^\top) = \mathrm{rk}\, M_i + 1\big]$. Similarly, $\big[i \in \mathcal{B} \Longleftrightarrow \mathrm{rk}(G_i^\top, M_{i+1}) = \mathrm{rk}\, M_{i+1} + 1\big]$. Combining this with the formulas for the ECP, we obtain $\mathrm{rk}(M_i, G_i^\top) = \mathrm{rk}(G_i^\top, M_{i+1}) = \mathrm{rk}(M_i, G_i^\top, M_{i+1})$ for all $i \in \mathcal{I}$. This in turn implies the identities $\ker(M_i, G_i^\top) = \ker(G_i^\top, M_{i+1}) = \ker(M_i, G_i^\top, M_{i+1})$, from which the biproperness follows. □

Throughout the following shift property will be useful.

Remark 3.10. *[8, Rem. III.7] Let $\sigma : \mathbb{F}^n \longrightarrow \mathbb{F}^n$, $(c_0, \ldots, c_{n-1}) \longmapsto (c_1, \ldots, c_{n-1}, c_0)$ be the left cyclic shift on $\mathbb{F}^n$, and let $G^* \in \mathbb{F}^{r \times n}$ be the matrix consisting of the shifted rows $\sigma(g_l)$, $l = 1, \ldots, r$. If the list $\mathcal{S} = \big[s_l,\, l = 1, \ldots, r\big]$ is a span list for G, then $\mathcal{S}^* = \big[s_l^*, l = 1, \ldots, r\big]$ where $s_l^* = \mathcal{I}$ if $s_l = \mathcal{I}$ and $s_l^* = (a_l - 1,\, b_l - 1]$ if $s_l = (a_l, b_l]$, forms a span list for G^*. Furthermore, with the notation as in (3.1), the state spaces of the product trellis $T_{G^*,\mathcal{S}^*}$ are given by $V_i^* = \mathrm{im}\, M_{i+1}$, and the transition spaces are $E_i^* = \mathrm{im}(M_{i+1}, G_{i+1}^\top, M_{i+2})$.*

We close this section with the following factorization theorem by Koetter and Vardy.

Theorem 3.11. *[10, Thm. 6.2] Every linear and reduced tail-biting trellis is a product trellis of the form $T_{G,\mathcal{S}}$.*

While we may not know how a given trellis was constructed, the above theorem allows us to easily determine whether the trellis follows from the product construction. In the next section, we will determine how to recover the factors used in the construction of a given trellis.

In this section we will look at how the factors used to construct a product trellis affect the cycles of the codewords in the trellis as well as the structure of the entire trellis. In Proposition 4.7, we will show that the underlying structure of a linear trellis is solely dependent on the spans used in the elementary factors. Finally, we will present an algorithm to recover the elementary factors used to construct a given one-to-one product trellis.

Definition 4.1. *Let T be a linear and reduced trellis, and let $c = (v_0, c_0, \ldots, c_{n-1}, v_0) \in \mathcal{S}(T)\backslash\{0\}$.*

(a) The trellis span of c in T is the set $\operatorname{tspan}(c) := \{i : v_i \neq 0\}$.

(b) We say that c diverges at time i if $i \notin \operatorname{tspan}(c)$ and $i+1 \in \operatorname{tspan}(c)$.

(c) We say that c converges at time i if $i \in \operatorname{tspan}(c)$ and $i+1 \notin \operatorname{tspan}(c)$.

Note that the concept of divergence and convergence has been introduced for conventional trellises by Kschischang and Sorokine in [12].

Now consider the product trellis T in Figure 3.2. The cycle $(00, 1, 10, 0, 10, 1, 00)$ in T has a trellis span of $(0, 2]$, diverges at time 0, and converges at time 2. Note that this is the unique cycle representing $c_1 = (1, 0, 1)$, and c_1 along with the span $(0, 2]$ is used in the construction of T. It is not a coincidence that this chosen span for c_1 is equal to the trellis span of the cycle representing c_1. This will always be the case for codewords that are generators of the trellis as we will see in the next proposition.

Note that in the case that the generating codewords are linearly independent, one has a one-to-one trellis. Therefore every codeword will be represented by a unique cycle, and thus every codeword will be associated with exactly one trellis span. In this case, we will refer to the trellis span of the cycle representing a codeword as the trellis span of the codeword itself.

Now, consider the following proposition.

Proposition 4.2. *Let $T = T_{g_1,s_1} \times \ldots \times T_{g_r,s_r}$ where $g_1, \ldots, g_r$ are possibly linearly dependent, no g_l is the zero codeword, and no two spans are the same empty span. Define the matrix $L = (M_0|G_0^\top|M_1|\ldots|M_{n-1}|G_{n-1}^\top|M_0)$, where each M_i is as in Theorem 3.9 and each $G_i^\top$ represents the i-th column of the matrix with rows $g_1, ..., g_r$. Then the rows of L form a basis for $\mathcal{S}(T)$.*

Proof. We will first show that the matrix L has full row rank. To begin, we will define the following useful set, $\mathcal{L}_i = \{l \in \{1, \ldots, r\} : i \in s_l\}$. Now assume that $\alpha L = 0$ for some $\alpha \in \mathbb{F}^r$. Then $\alpha \in \ker M_i$ for all i. Because the nonzero rows of M_i are linearly independent, we have that $\alpha_l = 0$ for $l \in \cup_{i=0}^{n-1}\mathcal{L}_i$. Now, suppose $l \notin \cup_{i=0}^{n-1}\mathcal{L}_i$. In this case, $l \notin \mathcal{L}_i$ for all i, and thus s_l must be the empty set. Since no two empty spans in our span set are equal, each g_l with an empty span is a multiple of a different

standard basis vector. Therefore, since $\alpha G = 0$, we have that $\alpha_l = 0$ in this case as well. Thus $\alpha = 0$ and $\operatorname{rk} L = r$.
As a consequence of the product construction, we have that $\operatorname{im} L \subseteq \mathcal{S}(T)$, so it remains to show that every cycle in $\mathcal{S}(T)$ is in the image of L.
Let $c = (v_0, c_0, \dots, c_{n-1}, v_0)$
$\in \mathcal{S}(T)$. By Theorem 3.9, we know that $(v_i, c_i, v_{i+1}) = \alpha^{(i)}(M_i, G_i^\top, M_{i+1})$ for all i, where $\alpha^{(i)} \in \mathbb{F}^r$. Since c is a cycle, we also have that $\alpha^{(i+1)} - \alpha^{(i)} \in \ker M_{i+1}$ for all i. Thus $\alpha_l^{(i+1)} = \alpha_l^{(i)}$ when $l \in \mathcal{L}_{i+1}$. If s_l is of the form $(a_l, b_l]$, this implies $\alpha_l^{(a_l)} = ... = \alpha_l^{(b_l)}$. Similarly, for a span of the form $\mathcal{I}$ we have that $\alpha_l^{(0)} = ... = \alpha_l^{(n-1)}$. We will now construct the vector $\beta = (\beta_1, ..., \beta_r) \in \mathbb{F}^r$ in the following way: if $s_l = (a_l, b_l]$, set $\beta_l = \alpha_l^{(b_l)}$, and if $s_l = \mathcal{I}$, set $\beta_l = \alpha_l^{(0)}$. We claim that $c = \beta L$. Now, when $i \notin [a_l, b_l]$, we know that $g_{l,i} = 0$, and hence $\alpha_l^{(i)} g_{l,i} = \beta_l g_{l,i}$. On the other hand, when $i \in [a_l, b_l]$ we have $\alpha_l^{(i)} = \beta_l$ and hence $\alpha_l^{(i)} g_{l,i} = \beta_l g_{l,i}$. Thus, we get that $\beta G_i^\top = \sum_{l=1}^r \beta_l g_{l,i} = \sum_{l=1}^r \alpha_l^{(i)} g_{l,i} = \alpha^{(i)} G_i^\top = c_i$. By a similar argument, one can also show that $\beta M_i = \alpha^{(i)} M_i = v_i$. Therefore, $c = \beta L \in \operatorname{im} L$. □

Recall from Remark 3.8(a) that $T_{g,(a,a]} \cong T_{g,(a,a]} \times T_{g',(a,a]}$, and thus the restriction on the spans of the generators in the previous proposition does not limit our choice of linear trellises. Also, notice that the l-th row of L is a cycle in T representing g_l with trellis span s_l. Thus we can think of these cycles as generating cycles of our trellis. We will use this idea in the following proposition.

Proposition 4.3. *Let $T = T_{g_1,s_1} \times \dots \times T_{g_r,s_r}$, where $g_1, \dots, g_r$ are possibly linearly dependent, no g_l is the zero codeword, and no two spans are the same empty span. Let $c \in \mathcal{S}(T) \backslash \{0\}$. By Proposition 4.2, we can express c as*

$$c = \sum_{l=1}^{r} \alpha_l (v_0^{(l)}, g_{l,0}, v_1^{(l)}, \dots, v_{n-1}^{(l)}, g_{l,n-1}, v_0^{(l)})$$

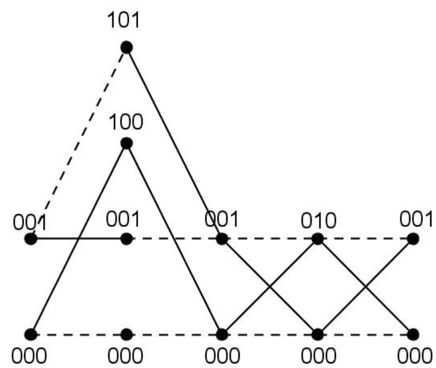
where $\alpha_l \in \mathbb{F}$ for $l = 1, \dots, r$ and $(v_0^{(l)}, g_{l,0}, v_1^{(l)}, \dots, v_{n-1}^{(l)}, g_{l,n-1}, v_0^{(l)})$ is the l-th row of the matrix L. Then $\operatorname{tspan}(c) = \bigcup_{\substack{l \in \{1,\dots,r\} \\ \alpha_l \neq 0}} s_l$.

Proof. Note that $c = (\sum_{l=1}^r \alpha_l v_0^{(l)}, c_0, \dots, \sum_{l=1}^r \alpha_l v_{n-1}^{(l)}, c_{n-1}, \sum_{l=1}^r \alpha_l v_0^{(l)})$, where $(c_0, \dots, c_n) = (\alpha_0, \dots, \alpha_r)G$. It is also important to notice that by Theorem 3.9 (1) that at every time i, the nonzero v_i^l's are standard basis vectors and thus are linearly independent. Therefore we have the following equivalences.

$$i \in \operatorname{tspan}(c) \iff \sum_{l=1}^{r} \alpha_l v_i^{(l)} \neq 0 \iff \exists l \text{ s.t. } \alpha_l \neq 0 \text{ and } i \in s_l \iff i \in \bigcup_{\substack{l \in \{1,\dots,r\} \\ \alpha_l \neq 0}} s_l.$$

□

Remark 4.4. *As a consequence of Proposition 4.3, cycles in a trellis T can only diverge at the starting points of the generating cycles' spans and can only converge*



$[T_{1100,(0,1]} \times T_{0011,(2,3]} \times T_{1011,(3,2]}]$

Figure 4.1: A product trellis with three generators.

at the cycles' ending points. If we also consider part (3) of 3.9, we can see that ECP(T)-SCP(T)=$(\delta_0, ..., \delta_{n-1})$ *where* δ_j *is the number of generating spans that diverge at* $j \in \mathcal{I}$*. This enables one to read these starting and ending points of the generating spans directly from the trellis. It is also worth noting that this formula is true even when the same span appears several times in the list of generating spans.*

Example 4.5. *Let* $\mathcal{C} = \operatorname{im}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \subseteq \mathbb{F}_2^4$*, and consider the trellis* T *for* $\mathcal{C}$ *shown in Figure 4.1. Note that* T *is constructed by taking the product of the elementary trellises of the rows of the above generator matrix based on the spans (0,1], (2,3], and (3,2] respectively. Now, the cycle representing the codeword 0111 results from adding together the cycles representing the generators 1100 and 1011. Thus by Proposition 4.3 we should obtain the trellis span associated with 0111 by taking the union of the intervals (0,1] and (3,2]. One can see that the cycle in* T *corresponding to the vector 0111 does indeed have a trellis span of* $(3, 2]$*. While trellis spans can be expressed as unions of intervals, not every trellis span can be expressed as a single interval. For instance, the cycle corresponding to the codeword 1111 has a trellis span of* $(0, 1] \cup (2, 3]$*. Some cycles, such as the one representing 1000, never pass through the zero state and have a trellis span of* $\mathcal{I}$*.*

In a particular sense, the spans of the generators used to construct a product trellis are minimal among all trellis spans. A specific result illustrating this is presented for one-to-one trellises in the following proposition.

Proposition 4.6. *Let* $T = T_{g_1,s_1} \times \ldots \times T_{g_k,s_k}$ *where* $g_1, ..., g_k$ *are linearly independent. Then* $\sum_{l=1}^{k} |s_l| = min\{\sum_{l=1}^{k} |tspan(c_l)|\}$ *where the minimum is taken over all sets of* k *linearly independent codewords. Furthermore, only* $s_1, ..., s_k$ *give this minimum.*

Proof. Suppose $c_1, .., c_k \in \mathcal{C}(T)$ are linearly independent. Since $g_1, ..., g_k$ also form a basis for $\mathcal{C}(T)$, there exists a bijection ψ from $\{c_1, ..., c_k\}$ to $\{g_1, ..., g_k\}$ where each

$c_j = \Sigma_{l=1}^{k} \alpha_l g_l$ is mapped to a g_l where $\alpha_l \neq 0$. Without loss of generality, let $\psi(c_l) = g_l$ for $l = 1, ..., k$. Thus by Proposition 4.3, we have $|\operatorname{tspan}(c_l)| \geq |s_l|$ for $l = 1, ..., k$, and this gives $\sum_{l=1}^{k} |s_l| \leq \sum_{l=1}^{k} |\operatorname{tspan}(c_l)|$. Therefore, if we have equality in the previous statement, that is $\sum_{l=1}^{k} |s_l| = \sum_{l=1}^{k} |\operatorname{tspan}(c_l)|$, we must have $|\operatorname{tspan}(c_l)| = |s_l|$ for all l. Since $s_l \subseteq \operatorname{tspan}(c_l)$ for $l = 1, ..., k$, we get that $s_l = \operatorname{tspan}(c_l)$ for $l = 1, ..., k$. $\square$

At this point in the section, we will take a quick diversion from trellis spans to introduce the following result about structurally isomorphic trellises. The notation developed in the proof of this result will be used in Algorithm 4.8 to recover the spans of the elementary factors from a product trellis.

This result was also proved independently by Conti and Boston (see [3]). Originally our result was for one-to-one trellises, and upon discussion with Conti, we realized that this was an unnecessary restriction.

Proposition 4.7. *Suppose $T = T_{g_1,s_1} \times \ldots \times T_{g_r,s_r}$ and $T' = T_{g'_1,s'_1} \times \ldots \times T_{g'_{r'},s'_{r'}}$ are two trellises for $\mathcal{C}$. Let $\mathcal{S} = [s_l : l = 1, ..., r]$ and $\mathcal{S}' = [s'_l : l = 1, ..., r']$. Then T and T' are structurally isomorphic if and only if $\mathcal{S} = \mathcal{S}'$ (up to ordering).*

Proof. " $\Leftarrow$" If $\mathcal{S} = \mathcal{S}'$ (up to ordering), it is clear from the product construction that $T \cong_S T'$.
" $\Rightarrow$" Suppose $T \cong_S T'$. Since $T = (V, E)$ and $T' = (V', E')$ are linear trellises, the structural isomorphism maps $0 \in V_i$ to $0 \in V'_i$ for $i = 0, ..., n-1$. Thus we have $S = [\operatorname{tspan}(c) : c \in \mathcal{S}(T)] = [\operatorname{tspan}(c) : c \in \mathcal{S}(T')] = S'$. Since T and T' are structurally isomorphic, they have the same SCP and the same ECP. Thus Remark 4.4 implies that $r = r'$. Now suppose that $\mathcal{S} \neq \mathcal{S}'$. Then there exists a span $s \in \mathcal{S} \backslash \mathcal{S}'$. Without loss of generality, let s be a span of shortest length in $\mathcal{S} \backslash \mathcal{S}'$. Now, let N equal the number of times that s appears in S, and let α equal the number of times that s appears in $\mathcal{S}$. By Proposition 4.3, a cycle in T can have the trellis span s in one of three ways. It can be a linear combination of generators with spans equal to s, which can occur in $A = q^{\alpha} - 1$ ways. It can be a linear combination of generators with spans strictly contained in s. Say this occurs in B ways. Finally it could be a linear combination of both generators with spans strictly contained in s and generators with spans equal to s. If we let C equal the number of unions of spans in $\mathcal{S}$ such that the union is strictly contained in s, this occurs in $AB + AC$ ways. Thus we get that $N = A + B + AB + AC$. We can similarly define the quantities N', A', B', and C' for trellis T'. Because s is the span of shortest length in $\mathcal{S} \backslash \mathcal{S}'$, we know that $B = B'$, $C = C'$, and $A > A'$. This implies that $N > N'$, which contradicts the fact that $S = S'$. Thus we must have that $\mathcal{S} = \mathcal{S}'$. $\square$

Now that we have introduced the notation in the previous proof, we will proceed with the main result of this section. Let T be a linear, reduced, and connected trellis. From Theorem 3.11 and Proposition 3.5, we know that T is of the form $T_{g_1,(a_1,b_1]} \times \ldots \times T_{g_r,(a_r,b_r]}$. Using the following algorithm we can recover the spans of the elementary factors that can be used in the construction of T.

Algorithm 4.8. *Given a linear, reduced, and connected trellis T of depth n, set $S := [\text{tspan}(c) : c \in \mathcal{S}(T)]$ and $\mathcal{A} := [a_1, ..., a_r]$, the list of indices of the divergences.*
(Note that $\mathcal{A}$ and S can be read directly from the trellis T.)
Set $\hat{\mathcal{S}} = \emptyset$.
For $i = 0$ to $n-1$
For $j = 1$ to r
If $(a_j, a_j + i] \in S$, then calculate
$N =$ the number of times $(a_j, a_j + i]$ appears in S
$\hat{B} =$ the number of ways $(a_j, a_j + i]$ occurs as a union of strictly smaller spans in $\hat{\mathcal{S}}$
$\hat{C} =$ the number of unions of spans in $\hat{\mathcal{S}}$ such that the union is strictly contained in $(a_j, a_j + i]$
$\hat{\alpha} = \log_q \left(\frac{N+1+\hat{C}}{1+\hat{B}+\hat{C}}\right)$
and add $\hat{\alpha}$ copies of $(a_j, a_j + i]$ to $\hat{\mathcal{S}}$.
When $|\hat{\mathcal{S}}| = r$, stop. Output $\hat{\mathcal{S}}$.

Proposition 4.9. *For a linear, reduced, and connected trellis $T_{g_1,(a_1,b_1]} \times \ldots \times T_{g_r,(a_r,b_r]}$, the output $\hat{\mathcal{S}}$ of Algorithm 4.8 equals $\mathcal{S} = [(a_l, b_l] : l = 1, \ldots, r]$.*

Proof. For a given trellis span $(a, b]$, we have $N = A + B + AB + AC$, where N, A, B, and C are as in the proof of Proposition 4.7. Recall that α equals the number of times that $(a, b]$ appears in $\mathcal{S}$. Since $A = q^\alpha - 1$, we get that $\alpha = \log_q \left(\frac{N+1+C}{1+B+C}\right)$. Thus we must show that at every step in the algorithm, $B = \hat{B}$ and $C = \hat{C}$ so that $\alpha = \hat{\alpha}$. When $i = 0$, the algorithm is only adding spans of length zero to $\hat{\mathcal{S}}$. Since no span of length zero can be written as a union of strictly shorter spans or contain shorter spans, $B = C = 0$ for all trellis spans of this type. Similarly, $\hat{B} = \hat{C} = 0$.
Suppose the algorithm has run correctly for $i = 0$ to $i = m - 1$. This means that the list $\hat{\mathcal{S}}$ consists of all spans in $\mathcal{S}$ up to length $m - 1$. Now if a span of length m contains or can be written as a union of strictly shorter spans in $\mathcal{S}$, these spans will be in $\hat{\mathcal{S}}$ as well. Thus $B = \hat{B}$ and $C = \hat{C}$. $\square$

Note that with an additional step, this algorithm can also be used when l of the generating spans are $\mathcal{I}$. In this case, the trellis is disconnected and has q^l components with identical structure. The remaining spans can be determined by running the algorithm on the trellis component containing the zero cycle, which is a linear trellis itself.
We will now limit our focus to one-to-one trellises as we explore how to recover the generating codewords for a trellis. While the spans used in the elementary factors of a given product trellis determine the structure, the codewords used (up to scalar multiples) determine the isomorphism class of the trellis. One would like to simply be able to use the trellis to read the label-sequences of cycles with the desired trellis spans in order to determine the generating codewords. However, there is no guarantee that a generating span corresponds to the trellis span of only one codeword. We will show that in a particular case switching one generating codeword of a product trellis

for another codeword with the same trellis span does lead to an isomorphic trellis. In order to do so, we will need the following result.

Theorem 4.10. *Let G and G' be matrices in $\mathbb{F}^{k\times n}$ of rank k with rows $g_1, g_2, ..., g_k$ and $g'_1, g_2, ..., g_k$, respectively, where s_j is a span for g_j, $j = 1, ..., k$, and s_1 is a span for both g_1 and g'_1. If $\operatorname{im} G = \operatorname{im} G'$ and $g'_1 \in span_{\mathbb{F}}\{g_l : s_l \subseteq s_1\}$, then $T_{G,\mathcal{S}} \cong T_{G',\mathcal{S}}$ where $\mathcal{S} := [s_j : j = 1, ..., k]$.*

Proof. We will begin by setting up the following notation. Since $\operatorname{im} G = \operatorname{im} G'$, there exists an invertible matrix $T \in \mathbb{F}^{k\times k}$ such that $G' = TG$. It is easy to see that T is of the form $\left(\begin{smallmatrix} a & b \\ 0 & I \end{smallmatrix}\right)$ where $a \in \mathbb{F}\backslash\{0\}$, $b = (b_2, ..., b_k) \in \mathbb{F}^{k-1}$, and $I \in \mathbb{F}^{(k-1)\times(k-1)}$. It is easy to see that T^{-1} is of the form $\left(\begin{smallmatrix} a^{-1} & -a^{-1}b \\ 0 & I \end{smallmatrix}\right)$.
Now, let $T_{G,\mathcal{S}} = (V, E)$ and $T_{G',\mathcal{S}} = (V', E')$ where $V_i = V'_i = \operatorname{im} M_i$ and M_i is defined as in Theorem 3.9. We claim that $\varphi : E_i \to E'_i$ where $\alpha(M_i, G_i^\top, M_{i+1}) \mapsto \alpha T^{-1}(M_i, {G'_i}^\top, M_{i+1})$ for all $\alpha \in \mathbb{F}^k$ is the desired trellis isomorphism. To show that φ is in fact well-defined, as well as one-to-one, we will prove that $\ker M_i = \ker T^{-1}M_i$ for all i. In this proof it will be helpful to view M_i as the following block matrix, $M_i = \left(\begin{smallmatrix} \mu_i^1 & 0 \\ 0 & \hat{M}_i \end{smallmatrix}\right)$ where $\hat{M}_i = \begin{pmatrix} \mu_i^2 & & \\ & \ddots & \\ & & \mu_i^k \end{pmatrix}$. Thus, $T^{-1}M_i = \left(\begin{smallmatrix} a^{-1}\mu_i^1 & -a^{-1}b\hat{M}_i \\ 0 & \hat{M}_i \end{smallmatrix}\right)$.
We will now consider two cases. First, if $\mu_i^1 \neq 0$, then we have the following equivalences.

$$\begin{aligned} \alpha = (\alpha_1, ..., \alpha_k) \in \ker T^{-1}M_i &\iff [\alpha_1 = 0 \text{ and } (\alpha_2, ..., \alpha_k) \in \ker \hat{M}_i] \\ &\iff \alpha \in \ker M_i \end{aligned}$$

Thus $\ker M_i = \ker T^{-1}M_i$ as desired.
Second, suppose that $\mu_i^1 = 0$. Since $G' = TG$, we have $g'_1 = ag_1 + \sum_{i=2}^k b_i g_i$. Now consider the set $\mathcal{L} := \{l : b_l \neq 0\}$. We then have that $b_l = 0$ for $l \notin \mathcal{L}$. Because $\mu_i^1 = 0$, we know that $i \notin s_1$. Since the spans s_l where $l \in \mathcal{L}$ are contained in s_1, we also know that $i \notin s_l$ for $l \in \mathcal{L}$, and therefore $\mu_i^l = 0$ for $l \in \mathcal{L}$. Thus we get that $b\hat{M}_i = 0$, and it follows that $M_i = T^{-1}M_i$. Again, we have $\ker M_i = \ker T^{-1}M_i$.
Thus φ is indeed well-defined, and it easily follows that φ is an isomorphism. □

It is worth noting that Theorem 4.10 extends Proposition III.14 in [8] to a much larger class of trellises.

This now brings us to another main result of this section.

Theorem 4.11. *Let $T_{G,\mathcal{S}}$ be a product trellis where $\operatorname{rk} G = k$, the rows of G are given by $g_1, ..., g_k$, and $\mathcal{S} := [s_1, ..., s_k]$. Suppose that $g'_1, ..., g'_k$ is a linearly independent set of codewords in $\mathcal{C}(T)$ whose trellis spans are $s_1, ..., s_k$ respectively, and where $g'_j \in span_{\mathbb{F}}\{g_l : s_l \subseteq s_j\}$ for all j. Then $T_{G,\mathcal{S}} \cong T_{G',\mathcal{S}}$, where the rows of G' are $g'_1, ..., g'_k$.*

Proof. Note that if g has trellis span s, then s is also a span of g. Therefore, since $g'_1 \in \operatorname{span}_{\mathbb{F}}\{g_l : s_l \subseteq s_1\}$, Theorem 4.10 gives that $T_{G,\mathcal{S}} \cong T_{g'_1,s_1} \times T_{g_2,s_2} \times ... \times T_{g_k,s_k}$. By repeating the above argument we obtain the desired result. □

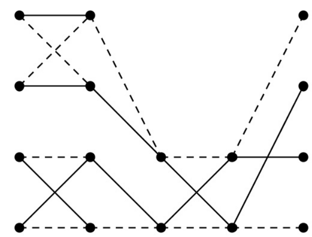


Figure 4.2: A connected one-to-one product trellis representing a code over $\mathbb{F}_2$.

We can now revisit the problem of identifying the generating codewords of a one-to-one trellis by following cycles in the trellis. Let $T_{G,\mathcal{S}}$ be a one-to-one product trellis where the rows of G are given by $g_1, ..., g_k$, and $\mathcal{S} := [s_1, ..., s_k]$. Suppose that a codeword $c = \sum_{l=1}^{k} \alpha_l g_l \in \mathcal{C}(T)$ where $\alpha_l \in \mathbb{F}$ has trellis span s_1. Since

$$\operatorname{tspan}(c) = \bigcup_{\substack{l \in \{1,...,k\} \\ \alpha_l \neq 0}} s_l$$

by Proposition 4.3, we know that c must be a linear combination of codewords in $\mathcal{C}(T)$ with trellis spans contained in s_1. Thus by the above theorem, $T_{c,s_1} \times T_{g_2,s_2} \times \ldots T_{g_k,s_k}$ is isomorphic to T. Therefore, once the generating spans of a trellis have been recovered, one can simply use any set of linearly independent codewords whose trellis spans are the generating spans in order to find a set of elementary factors for the trellis.

Example 4.12. *We will now use Algorithm 4.8 and Theorem 4.11 to find the elementary factors of the trellis shown in Figure 4.2. From the trellis we can see that* $\mathcal{A} = [0, 2, 3]$, $\mathcal{C}(T) = \{0000, 1000, 0100, 1100, 0011, 1011, 0111, 1111\}$, *and* $S = [\mathcal{I}, \mathcal{I}, (0, 1], (3, 2], (2, 0], (2, 1], (3, 2]]$.
To begin the algorithm, we will set $\hat{\mathcal{S}} = \emptyset$, *let* $i = 0$, *and first consider the spans* $(0, 0], (2, 2]$, *and* $(3, 3]$. *Since none of these spans are in* S, *we will move on to* $i = 1$. *While* $(2, 3]$ *and* $(3, 0]$ *are not in* S, *the span* $(0, 1]$ *is, so we must carry out the following computations. We can observe that for the span* $(0, 1]$ *we have* $N = 1$ *and* $\hat{B} = \hat{C} = 0$. *Thus we calculate that* $\alpha = 1$, *and we now set* $\hat{\mathcal{S}} = [(0, 1]]$. *Moving on to the considered spans of length 2, we see that only* $(2, 0]$ *is in* S. *The calculation is identical to the one in the previous step, so we again obtain* $\alpha = 1$ *and set* $\hat{\mathcal{S}} = [(0, 1], (2, 0]]$. *When* $i = 3$, *we see that* $(0, 3] \notin S$, *but* $(2, 1]$ *and* $(3, 2]$ *are trellis spans of the given trellis. For* $(2, 1]$ *we get that* $N = \hat{B} = 1$ *and* $\hat{C} = 2$ *which gives* $\alpha = 0$, *and thus* $\hat{\mathcal{S}}$ *remains unchanged. However, for the span* $(3, 2]$, *we obtain* $N = 2$, $\hat{B} = 0$, *and* $\hat{C} = 1$, *which gives* $\alpha = 1$. *Thus we now have* $\hat{\mathcal{S}} = [(0, 1], (2, 0], (3, 2]]$, *and since* $|\hat{\mathcal{S}}| = k = 3$, *we can end the algorithm because we now have the full set of generating spans. For this particular trellis, we see that this set of generating spans leads to two sets of codewords,* $\{1100, 1011, 1111\}$ *and* $\{1100, 1011, 0011\}$, *both of which are linearly independent. Note that 0011=1111+1100, and the trellis span of 1100 is (3,2], which is contained in the trellis span of 1111. By Theorem 4.11, these*

two sets of codewords will result in isomorphic trellises, and either set along with the generating spans can be used to construct the given trellis.

Chapter 5 Characteristic Generators and KV-trellises

For conventional trellises, it is well-known that for a given code there is a unique minimal trellis, up to isomorphism, in the sense of Definition 2.3. In [12], it is shown that the minimal conventional trellis can be constructed by taking the product of elementary trellises of generators with conventional spans that are minimal in a certain sense. Although there is not a unique minimal trellis in the tail-biting case, this idea can be naturally extended to produce a set of generators from which all minimal tail-biting trellises can be constructed. This extension was introduced by Koetter and Vardy in [11]; however, they introduced this set of generators as the product of a procedure. In this chapter, we will present the characteristic generators as a more natural extension of the situation for conventional trellises and then follow the approach taken in Section III of [7]. Throughout this chapter, we will assume that $\mathcal{C}$ has support $\mathcal{I}$.

In [4], Forney remarks on the concept of trellis oriented generator matrices for a code, and McEliece provides an excellent discussion on this topic in [13, Chapter IV]. Because of the importance of these matrices in the case of the minimal conventional trellis and the role they play in the results of Koetter and Vardy, we will begin this chapter by introducing some relevant properties of these matrices. The following results can be found in [12] and [13].

Remark 5.1. *A* ***minimal-span generator matrix*** *(or MSGM) for a code $\mathcal{C}$ is a generator matrix G such that G and its conventional proper span list $\mathcal{S} = \big[(a_l, b_l],\ l = 1, \ldots, k\big]$ have the following properties.*

(i) Among all generator matrices for $\mathcal{C}$ where the rows are taken to have proper conventional spans, the sum $\sum_{l=1}^{k} |(a_l, b_l]|$ is minimal.

(ii) If $c \in \mathcal{C}$ has a proper conventional span $(a_l, \hat{b}]$ where $l \in \{1, \ldots, k\}$, then $(a_l, b_l] \subseteq (a_l, \hat{b}]$.

(iii) The spans in $\mathcal{S}$ have distinct starting points and distinct ending points.

(iv) The spans in $\mathcal{S}$ have the predictable span property. That is, a codeword $c = \alpha_1 g_1 + \ldots \alpha_k g_k \in \mathcal{C}$ has the proper conventional span given by $(\min a_l, \max b_l]$ where the minimum and the maximum are taken over all $l \in \{1, \ldots, k\}$ such that $\alpha_l \neq 0$.

(v) The product trellis $T = T_{g_1,(a_1,b_1]} \times \ldots \times T_{g_1,(a_k,b_k]}$ is the minimal conventional trellis for $\mathcal{C}$.

Every code $\mathcal{C}$ has a minimal-span generator matrix, and such a matrix has full rank.

In the following definition, we will extend property (ii) to include circular intervals.

Definition 5.2. *A codeword $x \in \mathcal{C}$ with proper span $(a, b]$ is defined to be a characteristic generator of $\mathcal{C}$ if for any $c \in \mathcal{C}$ with proper span $(a, \hat{b}]$ we have $(a, b] \subseteq (a, \hat{b}]$. We say that $(a, b]$ is a characteristic span of $\mathcal{C}$.*

In this definition, one could easily use the ending points of the spans instead of the starting points as the two definitions are equivalent. This follows from an argument similar to that found in the proof of Proposition 5.3 (a).

Just like the spans of the MSGM, the characteristic spans of a code possess several nice properties outlined below.

Proposition 5.3. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a k-dimensional code with support $\mathcal{I}$.*

(a) No two distinct characteristic spans of $\mathcal{C}$ start or end at the same position.

(b) $\mathcal{C}$ has exactly n distinct characteristic spans.

(c) For all $j \in \mathcal{I}$, there exist exactly $n - k$ characteristic spans of $\mathcal{C}$ that contain j.

Proof. (a) From Definition 5.2 it is clear that no two distinct characteristic spans start at the same position. Suppose there exist distinct characteristic spans of $\mathcal{C}$ with the same endpoint, say $(a, b]$ and $(\hat{a}, b]$. Without loss of generality, assume that $(\hat{a}, b] \subsetneq (a, b]$. By definition, we know that there exist codewords in $\mathcal{C}$ with proper spans $(a, b]$ and $(\hat{a}, b]$, say c and $\hat{c}$ respectively. Now, $c - c_b(\hat{c}_b)^{-1}\hat{c} \in \mathcal{C}$ has a proper span that starts at a and is properly contained in $(a, b]$. Since $(a, b]$ is a characteristic span of $\mathcal{C}$, this is a contradiction.

(b) Since $\mathcal{C}$ has full support, for every $i \in \mathcal{I}$ there exists a codeword with a proper span starting at i. Thus $\mathcal{C}$ has at least n characteristic spans, and by part (a), $\mathcal{C}$ must have exactly n characteristic spans.

(c) The basic arguments found in the proof of Theorem 5.10 in [11] still apply in this setting. Let $T = T_{G,\mathcal{S}}$ be the minimal conventional trellis of $\mathcal{C}$. Thus by Remark 5.1, the set $\mathcal{S}$ must consist of k proper conventional characteristic spans, and these k spans in $\mathcal{S}$ do not contain 0. On the other hand, we will now show that the remaining $n - k$ characteristic spans of $\mathcal{C}$ do contain 0. Suppose there exists a characteristic generator x of $\mathcal{C}$ with a characteristic span $(a, b]$ that is not in $\mathcal{S}$ and does not contain 0. Then $(a, b]$ must be a conventional interval, and thus $a \leq b$. Since G generates $\mathcal{C}$ we can write x as a linear combination of the rows of G. Hence, Remark 5.1 (v) implies that $(a, b]$ must have the same starting point as a span in $\mathcal{S}$, but this contradicts part (a). Thus 0 is contained in exactly $n - k$ characteristic spans. By a similar argument applied to the cyclic shifts of $\mathcal{C}$, part (c) follows.

□

Following the approach of [7], we will now present the definition for a characteristic matrix, first introduced by Koetter and Vardy, in terms of the important properties that it possesses.

Definition 5.4. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a k-dimensional code with support $\mathcal{I}$. A characteristic pair of $\mathcal{C}$ is defined to be a pair $(X, \mathcal{T})$, where*

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{F}^{n\times n} \text{ and } \mathcal{T} = \big[(a_l, b_l],\, l = 1, \ldots, n\big] \tag{5.1}$$

have the following properties.

(i) $x_l \in \mathcal{C}$ *for* $l = 1, \ldots, n$

(ii) $\mathcal{T}$ *is the list of characteristic spans of* $\mathcal{C}$.

(iii) $(a_l,\, b_l]$ *is a span of* x_l *for* $l = 1, \ldots, n$.

We call X a characteristic matrix of $\mathcal{C}$ and $\mathcal{T}$ the characteristic span list.

Remark 5.5. *(a) Let σ be the left cyclic shift on $\mathbb{F}^n$ as in Remark 3.10. If $(X, \mathcal{T})$ as in (5.1) is a characteristic pair of $\mathcal{C}$, then $\big(X^*, \mathcal{T}^*\big)$ is a characteristic pair of $\sigma(\mathcal{C})$, where*

$$X^* := \begin{pmatrix} \sigma(x_1) \\ \vdots \\ \sigma(x_n) \end{pmatrix}, \quad \mathcal{T}^* := \big[(a_l - 1,\, b_l - 1],\, l = 1, \ldots, n\big].$$

(b) Note that every characteristic matrix of $\mathcal{C}$ contains an MSGM of $\mathcal{C}$, and thus every characteristic matrix of $\mathcal{C}$ contains a generating set of $\mathcal{C}$. Every characteristic matrix also contains a shifted MSGM for the code $\sigma^j(\mathcal{C})$ for each $j = 0, \ldots, n-1$. This is how Koetter and Vardy originally defined the characteristic pair in [11].

We will now focus on a particular class of trellises built by taking products of elementary trellises based on characteristic generators. These trellises are relevant because they possess many useful properties which we will explore in the remainder of the thesis.

Definition 5.6 ([8, Def.III.1])**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a k-dimensional code with support $\mathcal{I}$, and let $(X, \mathcal{T})$ be a characteristic pair of $\mathcal{C}$ as in (5.1). Any trellis of the form $T_{x_{l_1},(a_{l_1},b_{l_1}]} \times \ldots \times T_{x_{l_k},(a_{l_k},b_{l_k}]}$, where $x_{l_1}, \ldots, x_{l_k}$ are linearly independent rows of X, is called a $\mathrm{KV}_{(X,\mathcal{T})}$-trellis of $\mathcal{C}$. Every trellis that is a $\mathrm{KV}_{(X,\mathcal{T})}$-trellis for some characteristic pair $(X, \mathcal{T})$ of $\mathcal{C}$ is called a* KV*-trellis.*

Just as the minimal conventional trellis is built from the product of the rows of the MSGM, Koetter and Vardy showed that minimal tail-biting trellises are all, in fact, KV-trellises.

Theorem 5.7. *[11, Thm. 5.5] Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a k-dimensional code with support $\mathcal{I}$, and let T be a minimal trellis of $\mathcal{C}$ in the sense of Definition 2.3. Then T is one-to-one, and there exists a characteristic pair $(X, \mathcal{T})$ such that T is a $\mathrm{KV}_{(X,\mathcal{T})}$-trellis.*

It is important to note that while all minimal trellises are KV-trellises, not every KV-trellis is minimal.

Example 5.8. *Let* $\mathcal{C} = \operatorname{im}\begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \subseteq \mathbb{F}_2^{3\times 6}$*. It is easy to verify that*

$$X = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \mathcal{T} = [(0,2], (1,4], (2,5], (3,0], (4,1], (5,3]]$$

is a characteristic pair for $\mathcal{C}$*. The trellis generated by the first three rows of* X *leads to the minimal conventional trellis for* $\mathcal{C}$*, which has an SCP of (0,1,2,2,2,1). Consider the trellis* T *generated by the rows of* X *with spans* $(1,4]$,$(3,0]$*, and* $(5,3]$*. While this trellis is a KV-trellis for* $\mathcal{C}$*, it has an SCP of (2,1,2,2,2,1) and comparison with the minimal conventional trellis shows that* T *is not minimal.*

It is also important to note that we do not get the same set of KV-trellises from each characteristic pair $(X, \mathcal{T})$ of $\mathcal{C}$, hence the distinction in Definition 5.6. In fact, the set of minimal trellises generated by a particular characteristic pair of $\mathcal{C}$ may not contain all of the minimal trellises for $\mathcal{C}$, as we will see in the next example.

Example 5.9. *Let* $\mathbb{F} = \mathbb{F}_2$ *and consider the code* $\mathcal{C} = \operatorname{im}\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \subseteq \mathbb{F}_2^6$*. The two pairs* $(X, \mathcal{T})$ *and* $(X', \mathcal{T})$*, where*

$$X = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, X' = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}, \tag{5.2}$$

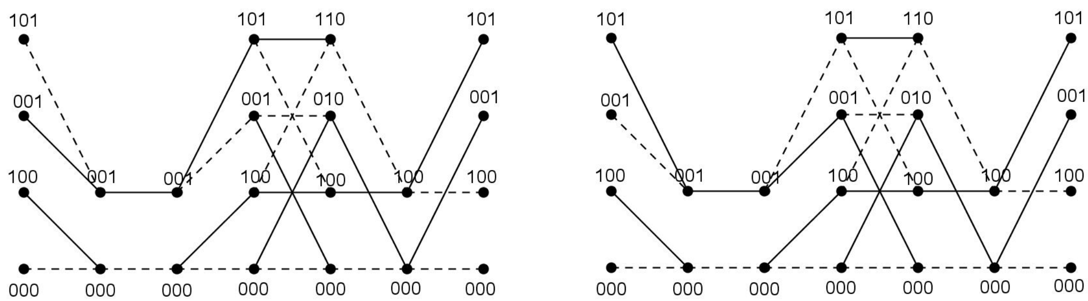
and $\mathcal{T} = \big[(0,2],\ (1,5],\ (2,0],\ (3,4],\ (4,1],\ (5,3]\big]$*, are both characteristic pairs of* $\mathcal{C}$*.*

Notice that while the last three rows of X *are linearly dependent and thus do not lead to a KV-trellis for* $\mathcal{C}$*, the last three rows of* X' *do in fact generate a KV-trellis for* $\mathcal{C}$*. By Proposition 4.7, this trellis cannot occur as the product of generators with spans other than (3,4], (4,1], and (5,3], and thus this trellis is not a* $\mathrm{KV}_{(X,\mathcal{T})}$*-trellis. Now, consider the following matrices and span list taken from the characteristic pairs above.*

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, G' = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \tag{5.3}$$

and $\mathcal{S} = \big[(2,0], (3,4], (5,3]\big]$. The product trellises $T_{G,\mathcal{S}}$ and $T_{G',\mathcal{S}}$, shown in Figure 5.1 are both minimal KV-trellises for $\mathcal{C}$ based on the same choice of spans. However, these two trellises, while structurally isomorphic, are not isomorphic. One can check that the cycle representing 110101 in $T_{G,\mathcal{S}}$ passes through the zero state at times $i = 4$ and $i = 5$, while the cycle representing 110101 in $T_{G',\mathcal{S}}$ never passes through the zero state.

Note that Theorem 4.10 is not applicable to these trellises. The codeword 011101 with span (5,3] in G' is equal to the sum of all three of the rows of G; however, the spans (2,0] and (3,4] are not contained in (5,3].



$[T_{G,\mathcal{S}}]$ $[T_{G',\mathcal{S}}]$

Figure 5.1: Nonisomorphic product trellises $T_{G,\mathcal{S}}$ and $T_{G',\mathcal{S}}$

This example became a major turning point in our investigation of characteristic matrices and KV-trellises because it detected a subtle oversight in [11].

We will see later that while we do not get every minimal trellis from a characteristic pair, we will get every minimal SCP from any characteristic pair. Furthermore, although we have seen in Example 5.9 that two KV-trellises based on the same span set are not necessarily isomorphic, they are structurally isomorphic by Proposition 4.7.

Chapter 6 Counting Characteristic Matrices

In this section, we will look at how the relationships between the characteristic spans for a given code affect the number of characteristic matrices for that code. Using a result of Koetter and Vardy, we can easily count the number of characteristic matrices for the dual code as well. In the case where a code has a unique characteristic matrix, we are able to determine the characteristic span list and show that all KV-trellises for the code are minimal. We conclude by addressing an issue from the previous section. Although, in general, we cannot obtain all KV-trellises for a code from one characteristic matrix, we can achieve all of the minimal SCPs for a code with any characteristic pair. Throughout this chapter, let $\mathcal{C} \subseteq \mathbb{F}^n$ be a k-dimensional code with support $\mathcal{I}$.

Because counting containments between characteristic spans plays a pivotal role in the upcoming results, we will introduce the following notation.

Definition 6.1. *Let $\mathcal{T} = [(a_1, b_1],\ldots,(a_n, b_n]]$ be the characteristic span list of $\mathcal{C}$. Then, for $l = 1, \ldots, n$, the set S_l is defined as follows*

$$S_l := \{r : (a_r, b_r] \subsetneq (a_l, b_l]\}.$$

Since generators that differ by only a scalar multiple lead to isomorphic trellises, it will simplify the counting process if we restrict ourselves to normalized generators, defined below.

Definition 6.2. *A characteristic generator x of $\mathcal{C}$ with characteristic span $(a, b]$ is normalized if $x_a = 1$. A characteristic matrix X of $\mathcal{C}$ is normalized if X consists of normalized characteristic generators.*

An important aspect in counting the number of characteristic matrices for a given code is first determining how many codewords in the code fit each characteristic span, and we carry out this calculation in the next few lemmas. We can then use the multiplication principle to find the total number of characteristic matrices for a given code, and using a result by Koetter and Vardy, we can show that a code and its dual have the same number of characteristic matrices.

Lemma 6.3. *Let $\mathcal{T} = [(a_1, b_1],\ldots,(a_n, b_n]]$ be the characteristic span list of $\mathcal{C}$, and for $l \in \{1, \ldots, n\}$. Then there exist linearly independent codewords $v_1, \ldots, v_{|S_l|}$ in $\mathcal{C}$ such that $(a_{r_j}, b_{r_j}]$ is a span of v_j for $j = 1, \ldots, |S_l|$.*

Proof. Let $(a_l, b_l] \in \mathcal{T}$. Since no characteristic span of $\mathcal{C}$ is the entire time axis $\mathcal{I}$, there exists $i \in \{0, \ldots, n-1\}$ such that $i \notin (a_l, b_l]$. Then, $(a_l - i, b_l - i]$ is a conventional span and each $(a_{r_j} - i, b_{r_j} - i]$ is a conventional span for $r_j \in S_l$. Thus, without loss of generality, we will assume that $(a_l, b_l]$ is a conventional span. Since each $(a_{r_j}, b_{r_j}]$ where $r_j \in S_l$ is a characteristic span of $\mathcal{C}$, we can find a codeword $v_j \in \mathcal{C}$ with span $(a_{r_j}, b_{r_j}]$ for each $r_j \in S_l$. Since the spans $(a_{r_j}, b_{r_j}]$ where $r_j \in S_l$ are distinct conventional characteristic spans, we get that the codewords v_j where $r_j \in S_l$ are linearly independent. $\square$

Lemma 6.4. *Given $(a_l, b_l]$, a characteristic span of $\mathcal{C}$, there exist $q^{|S_l|}$ normalized characteristic generators of $\mathcal{C}$ having this span.*

Proof. As in the previous proof, without loss of generality, we may assume that $(a_l, b_l]$ is a conventional span. Since $(a_l, b_l]$ is a characteristic span there exists a characteristic generator of $\mathcal{C}$, say x_l, with this span. Let $v_1, \dots, v_{|S_l|}$ be as in Lemma 6.3. Let c be any codeword. We claim that c is a normalized characteristic generator of $\mathcal{C}$ with span $(a_l, b_l]$ if and only if $c \in x_l + \mathrm{span}_{\mathbb{F}}\{v_r | r = 1, \dots, |S_l|\}$.

Now, if $c \in x_l + \mathrm{span}_{\mathbb{F}}\{v_r | r = 1, \dots, |S_l|\}$, then c has the conventional span $(a_l, b_l]$ since the span of each v_r is properly contained in $(a_l, b_l]$ and no two characteristic spans have the same endpoints. Since x_l is normalized and $v_{r,a_l} = 0$ for all $r = 1, \dots, |S_l|$, we get that c is normalized.

On the other hand, suppose that c is a normalized characteristic generator of $\mathcal{C}$ with span $(a_l, b_l]$. Consider the codeword $c - x_l$. Since $(c - x_l)_{a_l} = 0$, it follows that $c - x_l$ has a conventional span that is properly contained in $(a_l, b_l]$. Since $v_1, \dots, v_{|S_l|}$ are linearly independent characteristic generators of $\mathcal{C}$ with conventional spans, they can be expanded to an MSGM for $\mathcal{C}$, say $v_1, \dots, v_k$. Then, $c - x_l = \alpha_1 v_1 + \ldots + \alpha_k v_k$ where $\alpha_i \in \mathbb{F}$ for $i = 1, .., k$. By Remark 5.1, we get $\alpha_i = 0$ for v_i's with conventional spans that are not properly contained in $(a_l, b_l]$. Thus $c - x_l \in \mathrm{span}_{\mathbb{F}}\{v_r | r = 1, \dots, |S_l|\}$ and the claim follows. $\square$

Once we know how many codewords in $\mathcal{C}$ fit a particular characteristic span, it is simple to calculate the total number of different normalized characteristic matrices for the code. We simply use the multiplication principle of counting to obtain that if there exist $q^{|S_l|}$ normalized characteristic generators with span $(a_l, b_l]$, then $\mathcal{C}$ has $\prod_{l=1}^{n} q^{|S_l|} = q^{\sum_{l=1}^{n} |S_l|}$ normalized characteristic matrices. A Maple procedure is included in the appendix that will return the set of normalized codewords that fit each characteristic span and the number of normalized characteristic matrices for a given code.

Additionally, the following result by Koetter and Vardy makes it very simple to determine the set of characteristic spans for the dual of a given code, that is, the set $\mathcal{C}^\perp = \{w \in \mathbb{F}^n : wv^\top = 0 \text{ for all } v \in \mathcal{C}\}$.

Lemma 6.5. *[11, Lem. 5.11] Let $\mathcal{C} \subseteq \mathbb{F}^n$ and $\mathcal{C}^\perp$ both be codes with support $\mathcal{I}$, and let the characteristic span list of $\mathcal{C}$ be given by $\mathcal{T} = \big[(a_l, b_l],\ l = 1, \dots, n\big]$. Then the characteristic span list of $\mathcal{C}^\perp$ is given by $\big[(b_l, a_l],\ l = 1, \dots, n\big]$.*

Theorem 6.6. *If $\mathcal{C}$ and $\mathcal{C}^\perp$ both have full support, then $\mathcal{C}$ and $\mathcal{C}^\perp$ have the same number of normalized characteristic matrices.*

Proof. Let $\mathcal{T} = [(a_1, b_1], \dots, (a_n, b_n]]$ be the characteristic span list of $\mathcal{C}$. Thus, by Lemma 6.5, $[(b_1, a_1], \dots, (b_n, a_n]]$ is the characteristic span list of $\mathcal{C}^\perp$. Define the set $\hat{S}_l := \{r | (b_r, a_r] \subsetneq (b_l, a_l]\}$. Now, by Lemma 6.4, for each characteristic span $(a_l, b_l]$ in $\mathcal{C}$ there are $q^{|S_l|}$ codewords that fit the span as a normalized characteristic generator.

Thus, $\mathcal{C}$ has $q^{\sum_{l=1}^{n}|S_l|}$ normalized characteristic matrices. Because $\mathcal{C}$ and $\mathcal{C}^{\perp}$ have full support, no characteristic span of $\mathcal{C}$ is empty. Thus, we have $(a_r, b_r] \subsetneq (a_l, b_l]$ if and only if $(b_l, a_l] \subsetneq (b_r, a_r]$. Hence, $r \in S_l$ if and only if $l \in \hat{S}_r$. Therefore, $\sum_{l=1}^{n}|S_l| = \sum_{l=1}^{n}|\hat{S}_l|$, and $\mathcal{C}$ and $\mathcal{C}^{\perp}$ have the same number of normalized characteristic matrices. □

Since the number of characteristic matrices for a given code is solely dependent on the number of containments between the characteristic spans, there is only one set of characteristic spans that will result in a unique characteristic matrix.

Corollary 6.7. *The code $\mathcal{C}$ has a unique normalized characteristic matrix if and only if all of the characteristic spans for $\mathcal{C}$ are of length $n-k$. In this case, the characteristic spans of $\mathcal{C}$ are $(0, n-k]$,$\ldots$,$(k-1, n-1]$,$(k, 0]$,$\ldots$,$(n-1, n-k-1]$.*

Proof. Since $\mathcal{C}$ has $q^{\sum_{l=1}^{n}|S_l|}$ normalized characteristic matrices, it is clear that $\mathcal{C}$ has a unique normalized characteristic matrix if and only if no characteristic span of $\mathcal{C}$ contains any other characteristic span of $\mathcal{C}$. Thus in the case that all of the characteristic spans of $\mathcal{C}$ are of length $n-k$, it is easy to see that they must be of the form stated above, and thus no characteristic span of $\mathcal{C}$ contains any other.

Assume that $\mathcal{C}$ has a unique normalized characteristic matrix. We will first show that $|(a_j, b_j]| \leq n-k$ for all j. By way of contradiction, suppose there is a characteristic span $(a_j, b_j]$ of $\mathcal{C}$, where $|(a_j, b_j]| > n-k$. Then, we can shift all of the spans so that $a_j = 0$. Without loss of generality, let $(a_1, b_1]$,$\ldots$,$(a_k, b_k]$ be conventional spans where $a_1 < a_2 < \ldots < a_k$, and let $(a_{k+1}, b_{k+1}]$,$\ldots$,$(a_n, b_n]$ be circular spans. Thus, $j = 1$ and $(0, b_1]$ is a characteristic span where $b_1 > n-k$. Since no characteristic span of $\mathcal{C}$ contains any other characteristic span of $\mathcal{C}$, we know that $b_1 < b_2 < \ldots < b_k$. So, k endpoints of characteristic spans must lie in the interval $[n-k+1, n-1]$, which only contains $k-1$ integers. Thus, we must have $b_1 \leq n-k$ and $|(a_l, b_l]| \leq n-k$ for all $l = 1, \ldots, n$. Now, suppose there is a characteristic span $(a_j, b_j]$, where $|(a_j, b_j]| < n-k$. Since no characteristic span of $\mathcal{C}$ contains any other characteristic span of $\mathcal{C}$, no characteristic span of $\mathcal{C}$ is empty. Thus, $|\chi(\mathcal{C}^{\perp})| = n$, and $(b_j, a_j]$ is a characteristic span of $\mathcal{C}^{\perp}$ such that $|(b_j, a_j]| > k$. Hence, we get a contradiction by the above argument since $\mathcal{C}^{\perp}$ also has a unique characteristic matrix by Theorem 6.6. Therefore, $|(a_l, b_l]| = n-k$ for all $l = 1, \ldots, n$. □

In [9, Cor. 3], Kan and Shen show that cyclic codes have characteristic spans of the above type and thus have unique normalized characteristic matrices. This is true for the class of MDS codes as well. Recall that an $[n, k, d]$ code is MDS (or maximum distance separable) if $d = n-k+1$. This is equivalent to the statement for any set of d indices $D = \{i_1, ..., i_d\}$ there exists a codeword in $\mathcal{C}$ with support D. Moreover, the dual of an MDS code is again MDS.

Corollary 6.8. *If $\mathcal{C}$ is an MDS code, then $\mathcal{C}$ has a unique normalized characteristic matrix.*

Proof. Since $\mathcal{C}$ is an MDS code, we know that the distance of $\mathcal{C}$ is $d = n-k+1$, and for all $0 \leq i_1 < i_2 < \ldots < i_d \leq n-1$ there exists a codeword $c \in \mathcal{C}$ such that

$c_i \neq 0$ if and only if $i \in \{i_1, \ldots, i_d\}$. Now, consider the set $\{0, \ldots, n-k\}$. This set has cardinality d, so by the above, there exists $c \in \mathcal{C}$ with this support and thus proper span $(0, n-k]$. Because $d = n-k+1$, no codeword in $\mathcal{C}$ can have a proper span with length less than $n-k$. Thus the characteristic spans of $\mathcal{C}$ must have a length of at least $n-k$. Since $\mathcal{C}^\perp$ is also an MDS code, Lemma 6.5 gives that no characteristic span can have length greater than $n-k$. Thus all characteristic spans of $\mathcal{C}$ have length $n-k$, and by Corollary 6.7, $\mathcal{C}$ has a unique normalized characteristic matrix. □

Additionally, Corollary 6.7 leads to an interesting result regarding minimality. In [11], Koetter and Vardy also address other notions of minimality with regard to KV-trellises in particular. In their Theorem 5.6, they prove that trellises that are minimal under several other minimality orders are also KV-trellises. One such order is the product order.

Definition 6.9. *Let $\mathcal{C} \subset \mathbb{F}^n$ be a linear block code and $T = (V, E)$ and $T' = (V', E')$ be two trellises for $\mathcal{C}$. We define the product order, denoted $\leq_{\prod}$, where $T \leq_{\prod} T'$ if $\prod_{i=0}^{n-1} |V_i| \leq \prod_{i=0}^{n-1} |V_i'|$.*

Note that for a linear trellis T, we have that $\prod_{i=0}^{n-1} |V_i'| = \prod_{i=0}^{n-1} q^{\xi_i}$ where SCP(T)= $(\xi_0, ..., \xi_{n-1})$. Additionally, if a linear trellis is minimal under the product order, it is also minimal in the sense of Definition 2.3. With this new notion of minimality, we can now prove the following result.

Theorem 6.10. *Let $\mathcal{T} = [(a_1, b_1], \ldots, (a_n, b_n]]$ be the characteristic span list of the code $\mathcal{C}$. If $\mathcal{C}$ has a unique normalized characteristic matrix, then every KV-trellis of $\mathcal{C}$ is product-minimal and thus minimal.*

Proof. By Corollary 6.7 we have $|(a_l, b_l]| = n-k$ for all $l = 1, \ldots, n$. Let $T = (V, E)$ be a KV-trellis for $\mathcal{C}$ based on characteristic generators with spans $(a_{l_1}, b_{l_1}], \ldots, (a_{l_k}, b_{l_k}]$ and $SCP(T) = (s_0, \ldots, s_{n-1})$. Then $\prod_{i=0}^{n-1} |V_i| = \prod_{i=0}^{n-1} q^{s_i} = q^{\sum_{i=0}^{n-1} s_i} = q^{\sum_{i=1}^{k} |(a_{l_i}, b_{l_i}]|}$. Therefore, since $\mathcal{C}$ has a unique characteristic matrix Corollary 6.7 implies that $\sum_{i=1}^{k} |(a_{l_i}, b_{l_i}]| = k(n-k)$, and we get that for every KV-trellis of $\mathcal{C}$ we have $\prod_{i=0}^{n-1} |V_i| = q^{k(n-k)}$. Thus, every KV-trellis for $\mathcal{C}$ is product-minimal. Since product-minimality implies minimality in the sense of Definition 2.3, every KV-trellis for $\mathcal{C}$ is minimal. □

In Section 5 we addressed the fact that not all KV-trellises for $\mathcal{C}$ come from a single characteristic pair of $\mathcal{C}$. Since we are often looking for the minimal trellises for a given code, we might wonder if one has to look at every KV characteristic pair for a code to find all of the minimal tail-biting trellises. If we are only looking at the size and structure of the minimal trellises, the answer is no. The following result shows that if certain characteristic spans generate a minimal trellis for $\mathcal{C}$ using one characteristic matrix, then each set of generators corresponding to these spans will be linearly independent. While these minimal trellises may not be isomorphic, they will be structurally isomorphic by Proposition 4.7.

Theorem 6.11. *If the characteristic generators $x_1,\ldots,x_k$ for the code $\mathcal{C}$ with the spans $(a_1,b_1],\ldots,(a_k,b_k]$, respectively, create a minimal trellis, T for $\mathcal{C}$ then any set of characteristic generators for $\mathcal{C}$ with spans $(a_1,b_1],\ldots(a_k,b_k]$ is linearly independent.*

Proof. If $x_1,\ldots,x_k$ are linearly independent, then $x_{1,a_1}^{-1}x_1,\ldots,x_{1,a_1}^{-1}x_k$ are also linearly independent. Thus, without loss of generality, suppose that x_1 is normalized. Let x_1' be a characteristic generator for $\mathcal{C}$ with span $(a_1,b_1]$, where $x_1' \neq x_1$. Suppose $x_1',x_2,\ldots,x_k$ are linearly dependent. Then, $\alpha_1x_1'+\alpha_2x_2+\ldots+\alpha_kx_k = 0$ for some $\alpha_1,\ldots,\alpha_k \in \mathbb{F}$. Since $x_2,\ldots,x_k$ are linearly independent, $\alpha_1 \neq 0$. Then, $x_1' = \beta_2x_2+\ldots\beta_kx_k$ for some $\beta_2,\ldots,\beta_k \in \mathbb{F}$. Without loss of generality, assume that x_1' is normalized. Since x_1 and x_1' have the same characteristic span and are both normalized with respect to that span, the proof of Lemma 6.4 gives that $x_1' = x_1+c$, where $c \in \text{span}_{\mathbb{F}}\{x_r|r \in S_1\}$. Thus, it follows that $x_1+c = \beta_2x_2+\ldots+\beta_kx_k$. Since $x_1+c \in \text{span}_{\mathbb{F}}\{x_2,\ldots,x_k\}$ and $x_1 \notin \text{span}_{\mathbb{F}}\{x_2,\ldots,x_k\}$, we also have that $c \notin \text{span}_{\mathbb{F}}\{x_2,\ldots,x_k\}$. Since $c \in \text{span}_{\mathbb{F}}\{x_r|r \in S_1\}$, there is a span of c, say $(a',b']$ that is properly contained in $(a_1,b_1]$, the characteristic span of x_1. Thus, the product trellis generated by $c,x_2,\ldots,x_k$ with the spans $(a',b'],(a_2,b_2],\ldots,(a_k,b_k]$, respectively, represents $\mathcal{C}$ and is strictly smaller than T. This contradicts the fact that T is minimal, and thus $x_1',x_2,\ldots,x_k$ are linearly independent. Hence, if $x_1',\ldots,x_k'$ are characteristic generators for $\mathcal{C}$ with spans $(a_1,b_1],\ldots,(a_k,b_k]$, respectively, the generators $x_1,\ldots,x_k$ can be switched to $x_1',\ldots,x_k'$ one at a time as above to maintain a set of linearly independent characteristic generators. □

In addition to the product construction, we will make use of the tail-biting BCJR-construction introduced by Nori and Shankar in [15]. It is an extension of the construction for conventional trellises introduced by Bahl, Cocke, Jelinek, and Raviv in [1], hence the name BCJR. We will begin with a general definition, and later introduce a specific type of BCJR-trellis.

Definition/Theorem 7.1. *[15, Lem. 2, Lem. 3] Let the code $\mathcal{C}$ and the matrices $G \in \mathbb{F}^{r\times n}$, $H \in \mathbb{F}^{(n-k)\times n}$ be as in (2.2) and (2.3). Let $D \in \mathbb{F}^{r\times(n-k)}$ be any matrix. For $i \in \mathcal{I}$ define the matrices*

$$N_0 = D \text{ and } N_i = N_{i-1} + G_{i-1}^\top H_{i-1} \text{ for } i > 0. \tag{7.1}$$

Then $N_n = N_0$. We define $T_{(G,H,D)}$ to be the trellis with state spaces $V_i := \operatorname{im} N_i \subseteq \mathbb{F}^{n-k}$ and transition spaces $E_i = \operatorname{im}(N_i, G_i^\top, N_{i+1})$. Then, $T_{(G,H,D)}$ is a linear, reduced, and biproper trellis representing the code $\mathcal{C}$. We call D the displacement matrix for the trellis $T_{(G,H,D)}$, and the trellis $T_{(G,H,D)}$ is called a general (tail-biting) BCJR-trellis of $\mathcal{C}$.

All statements in this result can easily be proven. Notice, in particular, that $N_n = N_0$ simply follows from $N_i = N_0 + \sum_{j=0}^{i-1} G_j^\top H_j$, $i = 1,\ldots,n$, along with $0 = GH^\top = \sum_{j=0}^{n-1} G_j^\top H_j$. The identity $\mathcal{C}(T_{(G,H,D)}) = \mathcal{C}$ will also be a consequence of Proposition 7.2 below. In the above definition if one chooses the displacement matrix D to be the zero matrix, then the resulting general BCJR-trellis is conventional and is in fact the conventional BCJR-trellis introduced in [1].

A useful property of general BCJR-trellises is stated in the proposition below. It is important to note that this result does not hold for all classes of trellises. Product trellises do not always possess this attribute. For example, the product trellis shown in Figure 4.2 has a path whose edge-label sequence is the zero codeword.

Proposition 7.2. *[8, Prop. IV.2] Let $T = T_{(G,H,D)}$ be as in Definition/Theorem 7.1. Then the edge-label sequence of a path of length n starting at time 0 is a codeword if and only if the path is a cycle.*

Proof. Let $(\alpha^{(i)}N_i, \alpha^{(i)}G_i^\top, \alpha^{(i)}N_{i+1})$, $i \in \mathcal{I}$ where $\alpha^{(i)} \in \mathbb{F}^r$ be a path through T. Thus $\alpha^{(i)}N_{i+1} = \alpha^{(i+1)}N_{i+1}$ for $i = 0,\ldots,n-2$. By induction and the recursive definition of the N_i matrices, we obtain $\alpha^{(i+1)}N_{i+1} = \alpha^{(0)}N_0 + \sum_{j=0}^{i} \alpha^{(j)}G_j^\top H_j$ for $i = 0,\ldots,n-2$. This implies

$$\alpha^{(n-1)}N_0 = \alpha^{(n-1)}N_n = \alpha^{(n-1)}N_{n-1} + \alpha^{(n-1)}G_{n-1}^\top H_{n-1} = \alpha^{(0)}N_0 + \sum_{j=0}^{n-1} \alpha^{(j)}G_j^\top H_j.$$

Hence we have the equivalences

$$\begin{aligned}\alpha^{(n-1)}N_0 = \alpha^{(0)}N_0 \quad &\Longleftrightarrow \quad \sum_{i=0}^{n-1} \alpha^{(i)}G_i^\top H_i = 0 \\ &\Longleftrightarrow \quad (\alpha^{(0)}G_0^\top, \ldots, \alpha^{(n-1)}G_{n-1}^\top) \in \ker H^\top = \mathcal{C}.\end{aligned}$$

This shows that the path is a cycle if and only if its edge-label sequence is a codeword. □

This proposition leads to a useful result regarding nonmergeability. Recall that a trellis $T = (V, E)$ is said to be mergeable if there exist $u, v \in V_i$ for some $i \in \mathcal{I}$ that can be merged, that is replaced by a single vertex that inherits the edges incident to both u and v, without changing $\mathcal{C}(T)$. It is often difficult to check a trellis for mergeability; however, in the case of general BCJR-trellises connectedness and mergeability are equivalent.

Proposition 7.3. *Let $T = T_{(G,H,D)}$ be a general BCJR-trellis. Then T is weakly connected if and only T is nonmergeable.*

Proof. "⇒" Assume that T is connected. Suppose that $u, v \in V_i$ can be merged. Without loss of generality, we may also assume that $i = 0$. Since T is linear, we can assume that one of the vertices, say u, is 0. Thus, every path from v to 0 represents a codeword. Such a path exists by Proposition 2.2 since T is connected. By Proposition 7.2, since this path represents a codeword, this path must be a cycle, and thus $v = 0$. Hence, T is nonmergeable.
"⇐" Suppose that T is not connected. Then T has at least two connected components, say T_1 and T_2. Let $v_1, v_2 \in V_i$, where $v_i \in T_i$, $i = 1, 2$. Suppose we merge v_1 and v_2 into a single vertex v. This will not create any new cycles in T. For example, if a path starts in T_1 and crosses over to T_2 at v, there is no way for the path to return to T_1 to complete the cycle. Therefore T is mergeable. □

While some general BCJR-trellises have nice features, as we will see in the next example, they are not always well-behaved.

Example 7.4. *(a) Consider the code $\mathcal{C} = \operatorname{im} G = \ker H^\top \subseteq \mathbb{F}_2^3$ where*

$$G = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}.$$

Let $D = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The general BCJR-trellis $T_{(G,H,D)}$ representing $\mathcal{C}$ is shown in Figure 7.1. It has state and transition spaces given by $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im}(N_j, G_j^\top, N_{j+1})$, where

$$\begin{aligned}&(N_0|G_0^\top|N_1|G_1^\top|N_2|G_2^\top|N_0) \\ &= \left(\begin{array}{c|c|c|c|c|c|c} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right).\end{aligned}$$

We have staggered the N_j matrices with the columns of G in order to easily show all of the transition spaces E_j.

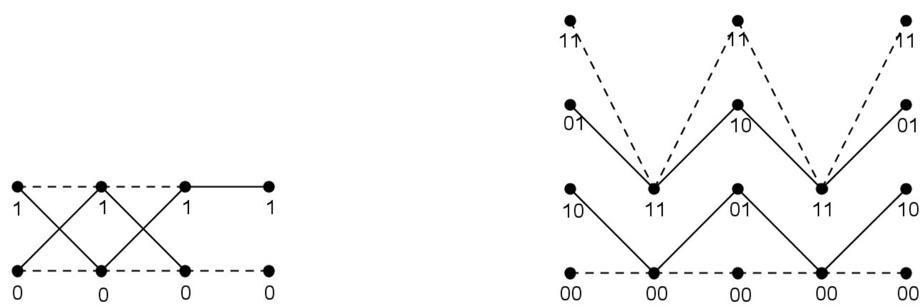
(b) *Let $\mathcal{C}' = \operatorname{im} G' = \ker H'^{\top} \subseteq \mathbb{F}_2^4$ where*

$$G' = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = H'.$$

Let $D' = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. The trellis $T_{(G',H',D')}$ has state and transition spaces given by $V'_j = \operatorname{im} N'_j$ and $E'_j = \operatorname{im}(N'_j, {G'}_j^{\top}, N'_{j+1})$, where

$$(N'_0|{G'_0}^{\top}|N'_1|{G'_1}^{\top}|N'_2|{G'_2}^{\top}|N'_3|{G'_3}^{\top}|N'_0) =$$
$$\left(\begin{array}{cc|c|cc|c|cc|c|cc|c|cc} 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \end{array}\right).$$

As one can see in Figure 7.1, the trellis $T_{(G',H',D')}$ is not connected and hence is also mergeable. It is also not one-to-one.



$[T_{(G,H,D)}]$ $[T_{(G',H',D')}]$

Figure 7.1: Two general BCJR-trellises

We will now move from the general BCJR-construction from Definition/Theorem 7.1 to one based on spans for the generators. This type of trellis possesses more useful properties than its general counterpart, and later in this section we will discuss the relationship between these trellises and product trellises.

Definition 7.5. *Let the code $\mathcal{C}$ and the matrices $G \in \mathbb{F}^{r\times n}$, $H \in \mathbb{F}^{(n-k)\times n}$ be as in (2.2) and (2.3). Let $\mathcal{S} := [(a_l, b_l], l = 1, \ldots, r]$ be a span list of G. Then the trellis $T_{(G,H,\mathcal{S})}$ is defined as $T_{(G,H,D)}$, where*

$$D = \begin{pmatrix} d_1 \\ \vdots \\ d_r \end{pmatrix} \in \mathbb{F}^{r\times(n-k)} \text{ is such that } d_l = \operatorname{row}(D, l) = \sum_{j=a_l}^{n-1} g_{lj} H_j \text{ for } j = 1, \ldots, r. \tag{7.2}$$

The trellis $T_{(G,H,\mathcal{S})}$ is called a span-based (tail-biting) BCJR-trellis of $\mathcal{C}$.

Note that in [8] and [7], only proper spans are considered; however, in the case of span-based BCJR-trellises the use of non-proper spans does not create a problem. If $(a, b]$ is a non-proper span for g_l, then it is easy to check that $\sum_{j=a}^{n-1} g_{lj} H_j = \sum_{j=a'}^{n-1} g_{lj} H_j$ where a' is the first index past a where $g_{l,a}$ is nonzero. We will, however, not include any spans of the form $\mathcal{I}$.

Example 7.6. *(a) By inspection, one can see that for both parts (a) and (b) of Example 7.4, there are no spans that will generate the given displacement matrix as per Definition 7.5. These trellises are general BCJR-trellises, but they are not span-based BCJR-trellises.*

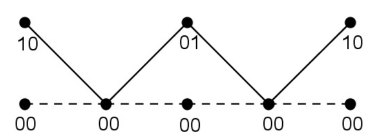
(b) Let $\mathcal{C}' = \operatorname{im} G' = \ker H'^{\top} \subseteq \mathbb{F}_2^4$ where

$$G' = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = H',$$

as in 7.4(b). Let $\mathcal{S} := [(3, 0], (1, 2]]$ be a span list of G'. Thus the trellis $T_{(G',H',\mathcal{S})}$ has state and transition spaces given by $V_j' = \operatorname{im} N_j'$ and $E_j' = \operatorname{im}(N_j', {G'_j}_j^{\top}, N_{j+1}')$, where

$$\begin{aligned} &(N_0' | {G_0'}^{\top} | N_1' | {G_1'}^{\top} | N_2' | {G_2'}^{\top} | N_3' | {G_3'}^{\top} | N_0') \\ &= \left(\begin{array}{cc|c|cc|c|cc|c|cc|c|cc} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right). \end{aligned}$$

As one can see in Figure 7.2, the trellis $T_{(G',H',\mathcal{S})}$ consists of the connected component of $T_{(G',H',D')}$ from Figure 7.1 that contains the zero cycle.



$[T_{(G',H',\mathcal{S})}]$

Figure 7.2: A span-based BCJR-trellis

Remark 7.7. *[8, Rem. IV.4] Span-based BCJR-trellises do not depend on the choice of the parity check matrix. Precisely, for any $U \in GL_{n-k}(\mathbb{F})$ the trellises $T_{(G,H,\mathcal{S})}$ and $T_{(G,UH,\mathcal{S})}$ are isomorphic. Indeed, if D is as in (7.2) then DU^{T} is the corresponding displacement matrix for $T_{(G,UH,\mathcal{S})}$. As a consequence, U^{T} restricted to all state spaces furnishes an isomorphism between the two trellises. However, if G and $UG \in \mathbb{F}^{r \times n}$, $U \in GL_r(\mathbb{F})$, are two generator matrices with the same span list $\mathcal{S}$, then the BCJR-trellises $T_{(G,H,\mathcal{S})}$ and $T_{(UG,H,\mathcal{S})}$ need not be isomorphic.*

We saw in Remark 3.10 that product trellises behave well with cyclic shifting. We now present a similar remark regarding BCJR-trellises.

Remark 7.8. *(a) Let the code $\mathcal{C}$ and the matrices $G \in \mathbb{F}^{r\times n}$, $H \in \mathbb{F}^{(n-k)\times n}$ be as in (2.2) and (2.3). Let $D \in \mathbb{F}^{r\times(n-k)}$ be any matrix. As in Remark 3.10, let $G^* \in \mathbb{F}^{r\times n}$ be the matrix consisting of the shifted rows $\sigma(g_l)$, $l = 1,\dots,r$, and define H^* accordingly. Let N_i, $i \in \mathcal{I}$, be the state space matrices as in (7.1) for the BCJR-trellis $T_{(G,H,D)}$. Then the state space matrices for the BCJR-trellis $T_{(G^*,H^*,\mathcal{S}^*)}$ are given by $N_i^* = N_{i+1}$ for $i \in \mathcal{I}$. This follows inductively due to (7.1).*

(b) Let $\mathcal{S}$ and D be as in Definition 7.5. Then $\mathcal{S}^* := \big[(a_l-1,\, b_l-1],\, l = 1,\dots,r\big]$ is a span list for the rows of G^*. Also, by (7.2) the l-th row of N_0^* is defined as $\sum_{j=a_l-1}^{n-1} g_{lj}^* H_j^* = \sum_{j=a_l-1}^{n-1} g_{l,j+1} H_{j+1} = \sum_{j=a_l}^{n-1} g_{lj} H_j + g_{l0} H_0$, and the latter is the l-th row of N_1. This way we obtain again that $N_i^* = N_{i+1}$ for all $i \in \mathcal{I}$.

We will close this section with some important results regarding span-based BCJR-trellises. The first states a nice property of the state space matrices of span-based BCJR-trellises that allows for easy comparison to the state space matrices of a product trellis. The second gives the property that all span-based BCJR-trellises are nonmergeable, and presents a relationship between the BCJR and KV-trellises for a given G with span set $\mathcal{S}$.

Proposition 7.9. *Let $T_{(G,H,\mathcal{S})}$ be a span-based BCJR-trellis with state space matrices N_i, $i \in \mathcal{I}$ and where $\mathcal{S} = [(a_l, b_l] : l = 1,\dots,r]$. Then $\mathrm{row}(N_j, l) = 0$ for all l such that $j \notin (a_l, b_l]$.*

Proof. If $(a_l, b_l]$ is conventional, then all nonzero entries of g_l are in the interval $[a_l, n-1]$. Therefore $d_l = \sum_{j=a_l}^{n-1} g_{lj} H_j = \sum_{j=0}^{n-1} g_{lj} H_j = 0$, due to $GH^\top = 0$. This proves the statement for $j = 0$. Suppose that $j \notin (a_l, b_l]$, and apply the left cyclic shift σ, as in Remark 7.8, to our data j times to obtain the trellis T' with state space matrices N_i'. Then, by the above argument we have that $\mathrm{row}(N_0', l) = 0$ for all l such that $0 \notin (a_l - j, b_l - j]$. Shifting back gives that $\mathrm{row}(N_j, l) = 0$ for all l such that $j \notin (a_l, b_l]$. □

Theorem 7.10. *[8, Cor. IV.7, Thm. IV.9, Rem. IV.13] Let G, H, and the span list $\mathcal{S}$ be as in Definition 7.5.*

(a) The BCJR-trellis $T_{(G,H,\mathcal{S})}$ is non-mergeable.

(b) If the product trellis $T_{G,\mathcal{S}}$ is non-mergeable, then $T_{G,\mathcal{S}}$ is isomorphic to $T_{(G,H,\mathcal{S})}$.

(c) If the product trellis $T_{G,\mathcal{S}}$ has the same SCP as the BCJR-trellis $T_{(G,H,\mathcal{S})}$, then these two trellises are isomorphic.

Proof. Let V_i be the i-th state space of $T_{(G,H,\mathcal{S})}$. Due to the shift property of $T_{(G,H,\mathcal{S})}$ as described in Remark 7.8, it suffices to show that $T_{(G,H,\mathcal{S})}$ is non-mergeable at $i = 0$. Furthermore, by linearity of the trellis, it is enough to show that no state $v \in V_0\backslash\{0\}$ can be merged with $0 \in V_0$. Thus, let $v \in V_0\backslash\{0\}$. If we can show that there exists a path of length n from v to $0 \in V_0$, then Proposition 7.2 implies that the corresponding edge-label sequence is not a codeword. Hence v and $0 \in V_0$ cannot be merged (because after merging this path would be a cycle).

Thus, it remains to establish the existence of a path of length n from v to 0. In order to do so, consider the product trellis $T_{G,\mathcal{S}}$. Let N_i and M_i as in (7.2) and (3.1) be the state space matrices for $T_{(G,H,\mathcal{S})}$ and $T_{G,\mathcal{S}}$, respectively. Moreover, let $v = \alpha^{(0)} N_0$, where $\alpha^{(0)} \in \mathbb{F}^k$. Then, $w := \alpha^{(0)} M_0$ is a state at time zero in the product trellis $T_{G,\mathcal{S}}$. Hence Remark 3.3(b) and Proposition 3.5 imply that $T_{G,\mathcal{S}}$ is weakly connected. Thus by Proposition 2.2 there exists a path in $T_{G,\mathcal{S}}$ from w to $0 \in V_0$. Precisely, there exist vectors $\alpha^{(1)}, \ldots, \alpha^{(n-1)} \in \mathbb{F}^k$ such that

$$\left.\begin{aligned} &\alpha^{(i)} - \alpha^{(i+1)} \in \ker M_{i+1},\ i = 0, \ldots, n-2, \\ &\alpha^{(n-1)} M_0 = 0. \end{aligned}\right\} \tag{7.3}$$

This also gives rise to a path in the BCJR-trellis $T_{(G,H,\mathcal{S})}$. Indeed, Proposition 7.9 and the definition of M_i in (3.1) show that $\ker M_i \subseteq \ker N_i$. Then (7.3) implies $\alpha^{(i)} - \alpha^{(i+1)} \in \ker N_{i+1}$ for $i = 0, \ldots, n-2$, and $\alpha^{(n-1)} N_0 = 0$. In other words, we obtain a path from $v = \alpha^{(0)} N_0$ to 0 in the trellis $T_{(G,H,\mathcal{S})}$. This concludes the proof of part (a). Proofs of the statements (b) and (c) may be found in [8], see results Corollary IV.7, Theorem IV.9, and Remark IV.13. □

As a consequence of Theorem 7.10, the nonmergeability of a product trellis $T_{G,\mathcal{S}}$ can easily be tested by comparing the data sets $\mathrm{SCP}(T_{G,\mathcal{S}})$ and $\mathrm{SCP}(T_{(G,H,\mathcal{S})})$. In the case where $T_{G,\mathcal{S}}$ is not isomorphic to the corresponding BCJR-trellis, and thus is mergeable, Theorem IV.9 in [8] shows that the product trellis can be merged to $T_{(G,H,\mathcal{S})}$.

Example 7.11. *(a) Let $\mathcal{C} = \operatorname{im} G = \ker H^\top \subseteq \mathbb{F}_2^3$, where*

$$G = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},\ H = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

Consider the span list $\mathcal{S} = [(0,2],(1,0]]$ for G. Then the span-based BCJR-trellis $T = T_{(G,H,\mathcal{S})}$ has state and transition spaces given by $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im}(N_j, G_j^\top, N_{j+1})$, where

$$(N_0|G_0^\top|N_1|G_1^\top|N_2|G_2^\top|N_0) = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{array}\right).$$

The trellis T is shown in Figure 7.3. The product trellis $T_{G,\mathcal{S}}$ is shown in Figure 2.1. While this trellis is not isomorphic to its BCJR-counterpart, the merging in Figure 2.1, shows that the merging results in a trellis that is isomorphic to $T_{(G,H,\mathcal{S})}$.

(b) Consider the code $\mathcal{C}' = \operatorname{im} G' = \ker H'^\top \subseteq \mathbb{F}_2^5$, where

$$G' = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix},\quad H' = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

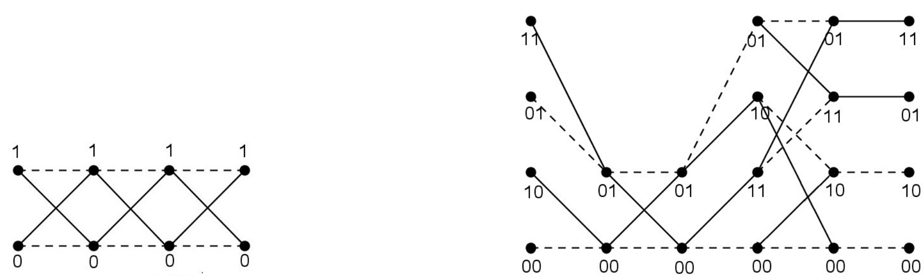
Then $\mathcal{S}' := \big[(1,3],\ (3,0],\ (2,1]\big]$ *is a span list for* G'*. Note that* $\mathcal{S}'$ *is not a set of characteristic spans since the codeword 00011 has a proper conventional span of* $(3,4]$*. Using the product construction, we obtain the following state space matrices for the trellis* $T' = T_{(G',\mathcal{S}')}$*.*

$$(M_0|M_1|M_2|M_3|M_4) = \left(\begin{array}{ccc|ccc|ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

Hence the SCP of T' *is given by* $(2,1,1,2,2)$*. The BCJR-construction for the trellis* $T_{(G,H,\mathcal{S})}$ *yields the following state space matrices.*

$$(N_0|N_1|N_2|N_3|N_4) = \left(\begin{array}{cc|cc|cc|cc|cc} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right)$$

Thus, the SCP of $T_{(G',H',\mathcal{S})}$ *is also given by* $(2,1,1,2,2)$*. Thus by Theorem 7.10, we have that* $T' = T_{G',\mathcal{S}'} \cong T_{(G',H',\mathcal{S}')}$*. Hence,* $T_{G',\mathcal{S}'}$ *is nonmergeable.*



$[T_{(G,H,\mathcal{S})}]$ $\qquad$ $[T_{G',\mathcal{S}'} \cong T_{(G',H',\mathcal{S}')}]$

Figure 7.3: Two span-based BCJR-trellises.

Finally we show that KV-trellises are always isomorphic to their corresponding BCJR-trellises and thus are also nonmergeable. This result is very useful as we will see in the following chapter.

Theorem 7.12. *[7, Thm. II.12] Assume that* $\mathcal{C}$ *and* $\mathcal{C}^\perp$ *have support* $\mathcal{I}$*, and let* X *be a characteristic matrix of* $\mathcal{C}$*. Let* $G \in \mathbb{F}^{r\times n}$ *be a selection of* r *rows of* X *with corresponding list of characteristic spans* $\mathcal{S} := [(a_l, b_l] : l = 1, \ldots, r]$*. Consider the BCJR-trellis* $T_{(G,H,\mathcal{S})}$ *with state space matrices* $N_j \in \mathbb{F}^{r\times(n-k)}$*. Then for all* $j \in \mathcal{I}$

1. $\mathrm{row}(N_j, l) = 0$ *for all* l *such that* $j \notin (a_l, b_l]$*,*

2. *the set* $\{\mathrm{row}(N_j, l) : l \text{ such that } j \in (a_l, b_l]\}$ *is linearly independent.*

3. $j \in \{b_1, \dots, b_r\} \Longleftrightarrow H_j \in \operatorname{im} N_j$.

As a consequence, the trellis $T_{(G,H,\mathcal{S})}$ is isomorphic to the corresponding product trellis $T_{G,\mathcal{S}}$, and if $r = k = \operatorname{rk} G$ we may call $T_{(G,H,\mathcal{S})}$ a KV-trellis.

Proof. (1) This statement has been proven in 7.9. Part (2) and the consequence have been shown in [8, Thm. IV.11] for the case where $G \in \mathbb{F}^{r\times n}$ has rank r. It can easily be checked that the same proof applies to this more general case.
(3) Using Remark 7.8 we may assume without loss of generality that $j = 0$.
"$\Rightarrow$" Let $0 = b_l$. Then (7.2) and the identity $GH^\top = 0$ along with the fact that $(a_l, b_l] = (a_l, 0]$ is the span of the l-th row of G imply $\operatorname{row}(N_0, l) = -g_{l0}H_0$, proving the desired result.
"$\Leftarrow$" Let $0 \notin \{b_1, \dots, b_k\}$ and assume $H_0 = \beta N_0$ for some $\beta \in \mathbb{F}^k$. Then $H_0 \neq 0$, due to our assumption that the dual code $\mathcal{C}^\perp$ has support $\mathcal{I}$. Define the set $\mathcal{L} := \{l : 0 \in (a_l, b_l], \beta_l \neq 0\}$. Using the definition of N_0 in (7.2) along with (1), we may write $H_0 = \sum_{l\in\mathcal{L}} \beta_l \sum_{j=a_l}^{n-1} g_{lj} H_j = \sum_{l\in\mathcal{L}} \beta_l \hat{g}_l H^\mathsf{T}$, where the vectors $\hat{g}_l = (\hat{g}_{l0}, \dots, \hat{g}_{l,n-1}) \in \mathbb{F}^n$ are defined via $\hat{g}_{lj} = g_{lj}$ if $j \geq a_l$ and $\hat{g}_{lj} = 0$ if $j < a_l$. As a consequence, $c := \sum_{l\in\mathcal{L}} \beta_l \hat{g}_l - e_0 \in \ker H^\mathsf{T} = \mathcal{C}$, where $e_0 \in \mathbb{F}^n$ is the first standard basis vector. The definition of $\hat{g}_l$ shows that the codeword c has span $(a_s, 0]$, where $a_s := \min\{a_l : l \in \mathcal{L}\}$. Notice that $a_s > 0$ due to the very definition of the set $\mathcal{L}$. Now we are in a position to invoke the characteristic spans. Indeed, Definition 5.2 implies $(a_s, b_s] \subseteq (a_s, 0]$, and since $(a_s, b_s]$ is circular, this in turn yields $b_s = 0$, contradicting our assumption that $0 \notin \{b_1, \dots, b_k\}$. This proves (3). □

Chapter 8 Dualizing Trellises

We will begin this chapter by discussing several methods for dualizing trellises to produce a trellis for the dual of a code. One follows easily from the BCJR-construction. Another, called the local dual, was introduced by Forney and is a very general procedure that works for many different types of graphs. It is based on dualizing the transition spaces in a natural way. Often the BCJR-dual and the local dual of a given trellis are non-isomorphic. In Theorem 8.10, we present conditions for a large class of well-behaved trellises under which the BCJR-dual and the local dual coincide. For KV-trellises, not only are these two dual trellises isomorphic, but we will show that the dual of a KV-trellis of $\mathcal{C}$ is a KV-trellis of $\mathcal{C}^\perp$. This proves and even extends a conjecture made by Koetter and Vardy, see [11, p. 2097]. We conclude the section by considering which types of minimality are preserved under the dualization process.

Throughout this section, we will only consider codes $\mathcal{C}$ where both $\mathcal{C}$ and $\mathcal{C}^\perp$ have full support.

We begin with the natural dualization that comes with the BCJR-construction. This was introduced by Nori and Shankar in [15].

Definition/Theorem 8.1. *Let $T = T_{(G,H,D)}$ be a general BCJR-trellis. Then the trellis $T^\perp = T_{(H,G,D^\top)}$ is called the BCJR-dual trellis of T and represents the dual code $\mathcal{C}^\perp$. Its state spaces are given by $\operatorname{im} {N_i}^\top$, $i \in \mathcal{I}$. Thus, $T_{(G,H,D)}$ and $T_{(H,G,D^\top)}$ have the same SCP, given by $(s_0, \ldots, s_{n-1})$, where $s_i = \operatorname{rk} N_i$. Therefore, if $T_{(G,H,D)}$ is a minimal trellis for $\mathcal{C}$, then $T^\perp$ is a minimal trellis for $\mathcal{C}^\perp$.*

The statements in this result can be easily verified, and a formal proof appears in [15]. Note that the BCJR-dual of a span-based BCJR-trellis is not guaranteed to be span-based. There may not be any suitable spans for H that will lead to the displacement matrix $N_0^\top$ as in Definition 7.5.

Example 8.2. *(a) Let $\mathcal{C} = \operatorname{im} G = \ker H^\top \subseteq \mathbb{F}_2^4$, where*

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \; H = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

Let $\mathcal{S}$ be the span list $\mathcal{S} = [(3,1],(1,2]]$ for G. Let $T = T_{(G,H,\mathcal{S})}$ with state space matrices N_i. Then the BCJR-dual trellis $T^\perp = T_{(H,G,{N_0}^\perp)}$ has state and transition spaces given by $V_j = \operatorname{im} {N_j}^\top$ and
$E_j = \operatorname{im}({N_j}^\top, H_j^\top, {N_{j+1}}^\top)$, where

$$({N_0}^\top | H_0^\top | {N_1}^\top | H_1^\top | {N_2}^\top | H_2^\top | {N_3}^\top | H_3^\top | {N_0}^\top) =$$
$$\left(\begin{array}{cc|c|cc|c|cc|c|cc|c|cc} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \end{array}\right).$$

The trellis is shown in Figure 8.1. It is also easy to see that the span list $\hat{\mathcal{S}} = (2,0],(3,2]]$ for H results in $T^\perp \cong T_{(H,G,\hat{\mathcal{S}})}$, and thus $T^\perp$ is a span-based BCJR-trellis.

(b) *Let* $\mathcal{C}' = \operatorname{im} G' = \ker H'^{\mathsf{T}} \subseteq \mathbb{F}_2^5$*, where*

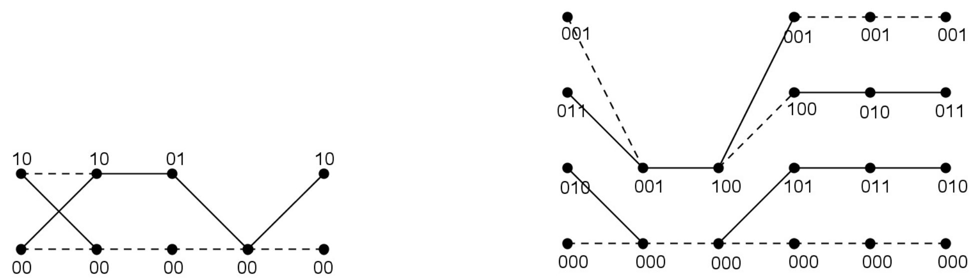
$$G' = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \quad H' = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix},$$

and let $\mathcal{S}' := \big[(1,3],\ (3,0],\ (2,1]\big]$ *be a span list for* G'*. Let* $T' = T_{(G',H',\mathcal{S}')}$ *with state space matrices* N_i'*. Then the trellis* $T'^{\perp}$ *has state and transition spaces given by* $V_j' = \operatorname{im} {N_0'}^{\top}$ *and* $E_j' = \operatorname{im}({N_j'}^{\top}, {H'}_j^{\top}, {N_{j+1}'}^{\top})$*, where*

$$({N_0'}^{\top} | {H_0'}^{\top} | {N_1'}^{\top} | {H_1'}^{\top} | {N_2'}^{\top} | {H_2'}^{\top} | {N_3'}^{\top} | {H_3'}^{\top} | {N_0'}^{\top}) =$$

$$\left(\begin{array}{ccc|c|ccc|c|ccc|c|ccc|c|ccc|c|ccc} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right).$$

As one can see in Figure 8.1, this trellis is not connected and hence is mergeable. Therefore, this is not a span-based BCJR-trellis by Theorem 7.10.



$[T^{\perp}]$ $[T'^{\perp}]$

Figure 8.1: Two BCJR-dual trellises.

We will now consider the process of local dualization which, in the particular case of tail-biting trellises, looks as follows.

Theorem 8.3. *Let* $T = (V, E)$ *be a linear trellis representing the code* $\mathcal{C} \subseteq \mathbb{F}^n$*. Let* $\hat{V}_j$*,* $j \in \mathcal{I}$*, be vector spaces such that* $\dim V_j = \dim \hat{V}_j$ *for all* $j \in \mathcal{I}$*, and fix non-degenerate bilinear forms* $\langle \cdot\,, \cdot \rangle$ *on* $V_j \times \hat{V}_j$*,* $j \in \mathcal{I}$*. For each transition space* $E_j \subseteq V_j \times \mathbb{F} \times V_{j+1}$*, define* $(E_j)^{\circ}$ *as the dual space with respect to the bilinear form*

$$(V_j \times \mathbb{F} \times V_{j+1}) \times (\hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1}) \longrightarrow \mathbb{F}, \quad \big((v,a,w),\ (\hat{v},b,\hat{w})\big) \longmapsto \langle v, \hat{v}\rangle + ab - \langle w, \hat{w}\rangle,$$

that is,

$$(E_j)^{\circ} := \big\{(\hat{v}, b, \hat{w}) \in \hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1} \,\big|\, \langle v, \hat{v}\rangle + ab - \langle w, \hat{w}\rangle = 0 \text{ for all } (v,a,w) \in E_j\big\}. \tag{8.1}$$

Then the trellis $T^\circ = (\hat{V}, E^\circ)$, where $\hat{V} = \bigcup_{j=0}^{n-1} \hat{V}_j$ and $E^\circ = \bigcup_{j=0}^{n-1} (E_j)^\circ$, is linear and represents $\mathcal{C}^\perp$. Furthermore, SCP(T°) = SCP$(T) := (s_0, \ldots, s_{n-1})$ and

$$\dim(E_j)^\circ = s_j + s_{j+1} + 1 - \dim E_j \text{ for } j \in \mathcal{I}. \tag{8.2}$$

We call T° the local dual of T.

It is clear that SCP(T)=SCP(T°), and the formula for ECP(T°) comes from the fact that $E_j \subseteq V_j \times \mathbb{F} \times V_{j+1}$. It is also straight-forward to check that $\mathcal{C}(T^\circ) \subseteq \mathcal{C}^\perp$. Let $(\hat{v}_0, \hat{a}_0, \ldots, \hat{a}_{n-1}, \hat{v}_0)$ be a cycle in T°. Then for any cycle $(v_0, a_0, \ldots, a_{n-1}, v_0)$ in T, we have $0 = \sum_{i=0}^{n-1}(\langle v_i, \hat{v}_i\rangle + a_i\hat{a}_i - \langle v_{i+1}, \hat{v}_{i+1}\rangle) = \sum_{i=0}^{n-1} a_i\hat{a}_i$. Thus $(\hat{a}_0, \ldots, \hat{a}_{n-1}) \in \mathcal{C}^\perp$. Forney's proof of this result can be found for a more general setting in [5, Sec. VII].

It should be noted that the isomorphism class of T° is not dependent on the choice of $\hat{V}_j$ or the non-degenerate bilinear forms used in the construction. If $\langle\,\cdot\,,\,\cdot\,\rangle_1$ and $\langle\,\cdot\,,\,\cdot\,\rangle_2$ are two such forms on $V_j \times \hat{V}_j$ and $V_j \times \tilde{V}_j$, then there exists an isomorphism $\phi_j : \hat{V}_j \to \tilde{V}_j$ such that $\langle v, w\rangle_1 = \langle v, \phi_j(w)\rangle_2$ for all $v \in V_j$, $w \in \hat{V}_j$. As a consequence, this isomorphism furnishes a trellis isomorphism between the two corresponding dual trellises.

Example 8.4. *(a) The BCJR-trellis $T_{(G,H,\mathcal{S})}$ used in Example 8.2 (a) has state and transition spaces $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im}(N_j, G_j^\top, N_{j+1})$ where*

$$(N_0|G_0^\top|N_1|G_1^\top|N_2|G_2^\top|N_3|G_3^\top|N_0) =$$

$$\left(\begin{array}{cc|c|cc|c|cc|c|cc|c|cc} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

For the local dual T°, we may use the standard inner product on $\mathbb{F}_2^2$ and $\hat{V}_j = V_j$ for $j = 1, 2, 3$; however, since the standard inner product on $V_0 \times V_0$ is degenerate, we take $\hat{V} = \operatorname{im}(1\ 0)$. Then

$$(E_0)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{array}\right),$$
$$(E_1)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 1 & 1 & 0 & 1 \end{array}\right),$$
$$(E_2)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 1 & 1 & 0 & 0 \end{array}\right),$$
$$(E_3)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 0 & 1 & 1 & 0 \end{array}\right).$$

Thus in this case the local dual trellis is isomorphic to the BCJR-trellis shown in Figure 8.1.

(b) Consider Example 8.2(b). The BCJR-trellis $T' := T_{(G',H',\mathcal{S}')}$ given in that example has state and transition spaces $V_j = \operatorname{im} N_j'$ and $E_j = \operatorname{im}(N_j', {G_j'}^\top, N_{j+1}')$ where

$$(N_0'|{G_0'}^\top|N_1'|{G_1'}^\top|\ldots|N_3'|{G_3'}^\top|N_0') =$$

$$\left(\begin{array}{cc|c|cc|c|cc|c|cc|c|cc|c|cc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \end{array}\right).$$

In order to compute the local dual T'°, we may again use the standard bilinear form on $\mathbb{F}_2^2$ and let V_j be the dual state spaces. Then

$$(E_0)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right),$$
$$(E_1)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 1 & 1 & 0 & 1 \end{array}\right),$$
$$(E_2)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{array}\right),$$
$$(E_3)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right),$$
$$(E_4)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right).$$

This leads to the trellis T'° in Figure 8.2. The trellis T'° is not edge-reduced because not every edge appears in a cycle. Indeed, the four diagonals in $(E_4)^\circ$, the last section of the trellis, are not part of any cycle in T'°. If we remove these 4 edges, then we obtain the BCJR-trellis $T'^{\perp}$ shown in Figure 8.1 (up to isomorphism).

(c) *Consider the 2-dimensional code*

$$\hat{\mathcal{C}} = \operatorname{im}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \subseteq \mathbb{F}_2^3$$

and choose the span list $\hat{\mathcal{S}} = [(1,2],(0,2]]$. Then the corresponding product trellis $\hat{T} = T_{\hat{G},\hat{\mathcal{S}}}$ has SCP $(0,1,2)$ and ECP $(1,2,2)$ and is shown in Figure 8.2 below. Notice that $\hat{T}$ is a conventional trellis, but not biproper (and thus not minimal). The transition spaces $\hat{E}_j$ of $\hat{T}$ can be read off from the matrix

$$\begin{aligned}&(\hat{M}_0|\hat{G}_0^\top|\hat{M}_1|\hat{G}_1^\top|\hat{M}_2|\hat{G}_2^\top|\hat{M}_0)\\ &= \left(\begin{array}{cc|c|cc|c|cc|c|cc} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right);\end{aligned}$$

see Definition 3.7 for the state space matrices $\hat{M}_j$ of product trellises. According to Theorem 8.3, the local dual $\hat{T}^\circ$ has SCP $(0,1,2)$ and ECP $(1,2,1)$. In order to compute $\hat{T}^\circ$, we observe that the standard bilinear form on $\mathbb{F}_2^2$ induces a non-degenerate form on each $\hat{V}_j = \operatorname{im}\hat{M}_j$, and thus may be used for the computation of the dual spaces $(\hat{E}_j)^\circ$. In particular, we will use $\hat{V}_j$ for the dual state spaces as well. Then we compute

$$(\hat{E}_0)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 0 & 1 & 0 & 1 \end{array}\right),$$
$$(\hat{E}_1)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right),$$
$$(\hat{E}_2)^\circ = \operatorname{im}\left(\begin{array}{cc|c|cc} 1 & 1 & 1 & 0 & 0 \end{array}\right).$$

This leads to the trellis $\hat{T}^\circ$ shown in Figure 8.2. Obviously, not every vertex appears in a cycle and thus the trellis $\hat{T}^\circ$ is not reduced. As a consequence, $\hat{T}^\circ$ is not a product trellis in the sense of Definition 3.7. This phenomenon can easily be explained. Indeed, it is straightforward to show that T is biproper if and only if $\hat{T}^\circ$ is state-trim (that is, each state lies on an edge in both the forward and backward direction).

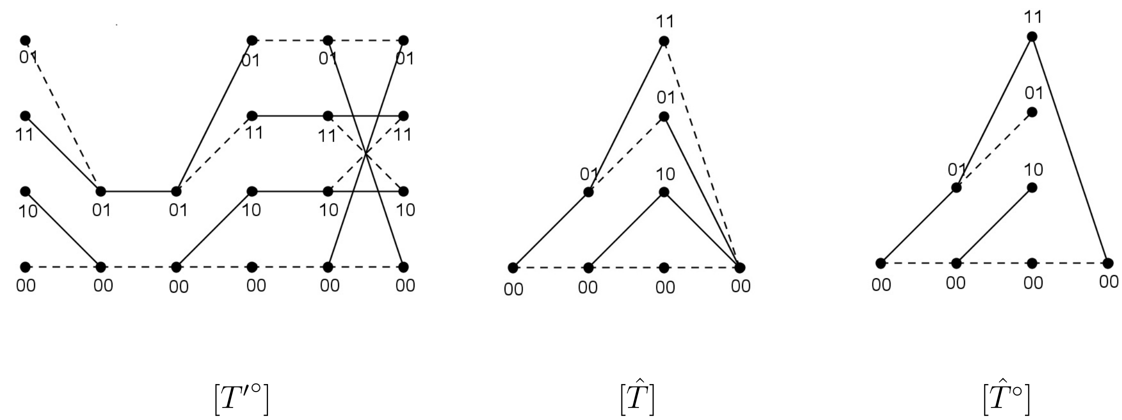


Figure 8.2: Examples of Local Dual Trellises.

The following proposition for connected product trellises will be useful in the remainder of this chapter.

Proposition 8.5. *Let $T_{G,\mathcal{S}} = (V, E)$ be a product trellis where no span in $\mathcal{S}$ is $\mathcal{I}$. Then T° is one-to-one.*

Proof. Since T° is linear, we must only show that any cycle in T° representing the zero codeword is the all-zero cycle. Let $(\hat{v}_j, 0, \hat{v}_{j+1})_{j=0}^{n-1}$ be a cycle in T°. Since each elementary trellis in the product of T is weakly connected, T is weakly connected as well by Proposition 3.5. Thus for any $v \in V_i$, there exists a path of length n from v to $0 \in V_i$. Let $(v, a_i, v_{i+1}, \ldots, v_0, a_0, v_1, \ldots, v_{i-1}, a_{i-1}, 0)$ be such a path. Then by (8.1), we have that $\langle v, \hat{v}_i\rangle = \langle v_{i+1}, \hat{v}_{i+1}\rangle = \ldots = \langle 0, \hat{v}_i\rangle = 0$. Thus $\hat{v}_i$ is orthogonal to every vertex in V_i. Since $\langle\,\cdot\,,\,\cdot\,\rangle$ is a non-degenerate bilinear form on $V_i \times \hat{V}_i$, we must have that $\hat{v}_i = 0$. Therefore our original cycle in T° is indeed the all-zero cycle. $\square$

We have seen in the previous examples that while the local dual and the BCJR-dual of a given trellis have the same state complexity profile, they are not always isomorphic. However, the following proposition shows that for a trellis T, the BCJR-dual $T^\perp$ is always a subtrellis of the local dual T°.

Proposition 8.6. *[7, Prop. III.4] Let $T = T_{(G,H,D)}$ be as in Definition/Theorem 7.1. Let $\hat{E}_j$ and $(E_j)^\circ$ be the transition spaces of the duals $T^\perp$ and T°, respectively. Then $\hat{E}_j \subseteq (E_j)^\circ$, up to trellis isomorphism.*

In the proof we will construct the local dual based on a suitable choice of dual state space and inner form, which will then make $T^{\perp}$ a true subtrellis of T° and not just an isomorphic copy.

Proof. Let $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im}(N_j,\, G_j^{\top},\, N_{j+1})$ be the state spaces and transition spaces of T, where the matrices N_j are defined as in (7.1). By the very definition of the BCJR-dual, the state spaces of $T^{\perp}$ are given by $\hat{V}_j = \operatorname{im} \hat{N}_j$, where $\hat{N}_j = N_j^{\top}$. Moreover, the transition spaces are given by $\hat{E}_j = \operatorname{im}(\hat{N}_j, H_j^{\top}, \hat{N}_{j+1})$. Notice that the bilinear form $V_j \times \hat{V}_j \longrightarrow \mathbb{F}$, defined as $\langle \alpha N_j, \beta \hat{N}_j \rangle := \alpha N_j \beta^{\mathsf{T}}$, is well-defined and non-degenerate. So we may construct the local dual T° based on this form. Obviously, $\dim \hat{V}_j = \dim V_j$ for all $j \in \mathcal{I}$, and the transition spaces of T° are

$$(E_j)^{\circ} = \left\{ (\beta \hat{N}_j, b, \tilde{\beta} \hat{N}_{j+1}) \in \hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1} \;\middle|\; \begin{array}{l} \alpha N_j \beta^{\mathsf{T}} + \alpha G_j^{\top} b - \alpha N_{j+1} \tilde{\beta}^{\mathsf{T}} = 0 \\ \text{for all } \alpha(N_j, G_j^{\top}, N_{j+1}) \in E_j \end{array} \right\}.$$

Now we see that $\hat{E}_j = \operatorname{im}(\hat{N}_j,\, H_j^{\top}, \hat{N}_{j+1}) \subseteq (E_j)^{\circ}$ since for all $\beta(\hat{N}_j,\, H_j^{\top}, \hat{N}_{j+1}) \in \hat{E}_j$ and $\alpha \in \mathbb{F}^k$ we have $\alpha N_j \beta^{\mathsf{T}} + \alpha G_j^{\top} H_j \beta^{\mathsf{T}} - \alpha N_{j+1} \beta^{\mathsf{T}} = \alpha(N_j + G_j^{\top} H_j - N_{j+1})\beta^{\mathsf{T}} = 0$, due to (7.1). $\square$

Recall from Proposition 8.5 that for a connected product trellis $T = T_{G,\mathcal{S}}$, the local dual T° is one-to-one. If T is also a BCJR-trellis, then the BCJR-dual $T^{\perp}$ is also one-to-one because $T^{\perp}$ is a subtrellis of T°.

Remark 8.7. *In [11, Sec. VII], Koetter and Vardy introduce another dual trellis construction. This dualization process is similar to the product construction. First elementary dual trellises for generators of the code are constructed based on specified spans. The vertex spaces of these trellises are defined in the same way as those of elementary trellises; however, the edge spaces, E_i, are constructed using one of five cases determined by the relationship between i and the span of the generator. These elementary dual trellises are then combined into a larger trellis representing the entire code by taking what they call the intersection product. Through straightforward computations outlined below, one can show that for any linear and reduced trellis (thus for any product trellis), Koetter and Vardy's intersection dual yields the same trellis as Forney's local dual construction. To begin, one can prove that the local dual and the elementary dual trellises coincide for any elementary trellis $T_{g,s}$. Next, it is clear that the intersection dual and the local dual will have the same vertex space since they both retain the vertex spaces of the original trellis. Thus, we only need to show that their edge spaces coincide. Now consider a code with generators $g_1,\ldots,g_r$ and corresponding spans $s_1,\ldots,\, s_r$. We can view the edge spaces, $(E_i)^{\circ}$, of the local dual as $(E_i)^{\circ} = \cap_{l=1}^{r}(E_i^l)^{\circ}$, where $(E_i^l)^{\circ}$ is the i-th edge space of the local dual of T_{g_l,s_l}. One can then show that the intersection dual is a subtrellis of the local dual. Thus, for each time i, the i-th edge space of the intersection dual is a subspace of the i-th edge space of the local dual. Finally, it is simple to prove that for each time i, the i-th edge space of the intersection dual has the same dimension as that of the local dual, and this completes the proof.*

We are now ready to present one of the main results of this section. We will characterize the linear, reduced, nonmergeable, and one-to-one trellises whose BCJR-duals and local duals coincide.

Lemma 8.8. *Let $T = T_{(G,H,\mathcal{S})}$ be a BCJR-trellis where $\mathcal{S} = [(a_l, b_l] : l = 1, \ldots, r]$, $a_1, \ldots, a_r$ are distinct and $b_1, \ldots, b_r$ are distinct. Let $\mathcal{A} = \{a_1, \ldots, a_r\}$ and $\mathcal{B} = \{b_1, \ldots, b_r\}$. If $T \cong T_{G,\mathcal{S}}$, then SCP($T^\perp$) $= (\xi_0, \ldots, \xi_{n-1})$ and ECP($T^\perp$) $=$ $(\hat{\epsilon}_0, \ldots, \hat{\epsilon}_{n-1})$ satisfy*

(a) $\xi_i \leq \hat{\epsilon}_i \leq \xi_i + 1$ for all $i \in \mathcal{I}$ and $\hat{\epsilon}_i = \xi_i$ if $i \in \mathcal{B}$,

(b) $\xi_{i+1} \leq \hat{\epsilon}_i \leq \xi_{i+1} + 1$ for all $i \in \mathcal{I}$ and $\hat{\epsilon}_i = \xi_{i+1}$ if $i \in \mathcal{A}$.

Proof. As usual, let N_i be the state space matrices of T. Recall that $\xi_i = \operatorname{rk} N_i^\top$ and $\hat{\epsilon}_i = \operatorname{rk}(N_i^\top, H_i^\top, N_{i+1}^\top)$. By the recursive definition of the N_i matrices, we also have that $\hat{\epsilon}_i = \operatorname{rk}(N_i^\top, H_i^\top) = \operatorname{rk}(H_i^\top, N_{i+1}^\top)$. Since $H_i^\top$ is a single column, the inequalities in (a) and (b) follow.
Now let $i \in \mathcal{A}$. Since T is a product trellis by assumption, we may use the formulas in Theorem 3.9. Let $\mathrm{ECP}(T) = (\epsilon_0, \ldots, \epsilon_{n-1})$. Then $\epsilon_i = \xi_i + 1$, and since $\epsilon_i = \operatorname{im}(N_i, G_i^\top, N_{i+1})$, this implies that $G_i^\top$ is not in the column space of N_i, equivalently, $G_i \notin \operatorname{im}(N_i)^\top$. Taking duals gives that there exists $\alpha \in \mathbb{F}^r$ such that $\alpha \in \ker N_i \backslash \ker G_i^\top$. Then $\alpha N_{i+1} = \alpha(N_i + G_i^\top H_i)$ gives that $\alpha N_{i+1} = \alpha G_i^\top H_i$. Since $\alpha G_i^\top \in \mathbb{F}\backslash\{0\}$, we have that H_i is in the column space of $N_{i+1}^\top$. Thus $\xi_{i+1} = \hat{\epsilon}_i$ as desired. The case where $i \in \mathcal{B}$ is similar. □

Lemma 8.9. *Let $T = T_{(G,H,\mathcal{S})}$ be a BCJR-trellis where $\operatorname{rk} G = k$, $\mathcal{S} = [(a_l, b_l] : l = 1, \ldots, k]$, $a_1, \ldots, a_k$ are distinct, and $b_1, ..., b_k$ are distinct. Let N_i, $i \in \mathcal{I}$ be the state space matrices of T. Assume that the following properties are satisfied.*

1. $\operatorname{row}(N_j, l) = 0$ *for all l such that $j \notin (a_l, b_l]$,*
2. *for $j \in \mathcal{I}$ the set $\{\operatorname{row}(N_j, l) : l$ such that $j \in (a_l, b_l]\}$ is linearly independent.*
3. *$j \in \{b_1, \ldots, b_r\} \iff H_j \in \operatorname{im} N_j$.*

Then $\bigcap_{j=0}^{n-1} \operatorname{im} N_j = \{0\}$.

Proof. The proof of Theorem II.13 in [7] proves this statement. It is easy to check that the KV-requirement made in that theorem is not needed. □

Theorem 8.10. *Let $T = T_{(G,H,\mathcal{S})} \cong T_{G,\mathcal{S}}$ where $\operatorname{rk} G = k$, $\mathcal{S} := [(a_l, b_l] : l = 1, \ldots, k]$, the starting points $a_1, \ldots, a_k$ are distinct, the ending points $b_1, \ldots, b_k$ are distinct, and N_i are the state space matrices of $T_{(G,H,\mathcal{S})}$ where $i \in \mathcal{I}$. Then, the following statements are equivalent.*

(a) $T^\perp$ is nonmergeable.

(b) $T^\perp$ is weakly connected.

(c) $T^\perp$ *is the span-based BCJR-trellis* $T_{(\hat{H},G,\hat{\mathcal{S}})}$ *where* $\hat{H}$ *is a parity check matrix for* $\mathcal{C}(T)$ *and* $\hat{\mathcal{S}}$ *is a span set whose starting points are* $\mathcal{I}\backslash\{b_1,\dots,b_k\}$ *and whose ending points are* $\mathcal{I}\backslash\{a_1,\dots,a_k\}$.

(d) $T^\perp \cong T^\circ$

(e) For all $i \in \mathcal{I}$, $[H_i \in \operatorname{im} N_i \iff i \in \{b_1,\dots,b_k\}]$

(f) $\cap_{i=0}^{n-1} \operatorname{im} N_i = \{0\}$

(g) T° *is edge-reduced, that is, every edge in* T° *lies on a cycle.*

Proof. Let $(\xi_0,\dots,\xi_{n-1})$ be the common SCP of the trellises T, $T^\perp$, and T°. Furthermore, denote the ECP of T by $(\epsilon_0,\dots,\epsilon_{n-1})$ and the ECP of $T^\perp$ by $(\hat{\epsilon_0},\dots,\hat{\epsilon}_{n-1})$.

"(a) ⇔ (b)" Since the dual of a BCJR-trellis is a general BCJR-trellis, this follows from Proposition 7.3.

"(a) ⇒ (c)" Since $T^\perp$ is a subtrellis of T°, we have that $T^\perp$ is one-to-one, and by Definition/Theorem 7.1, we have that $T^\perp$ is linear and reduced. Because Theorem 3.11 states that every linear and reduced trellis is a product trellis, we can also write $T^\perp$ as $T_{\hat{H},\hat{\mathcal{S}}}$ for some $\hat{H}$ such that $\mathcal{C}^\perp = \operatorname{im}\hat{H}$ and $\hat{\mathcal{S}}$ is a span list for $\hat{H}$. Since $T^\perp$ is one-to-one, $\hat{H}$ has full row rank by Theorem 3.9. Since $T^\perp$ is nonmergeable by assumption, we know from (2.1) that $T^\perp$ is biproper. Hence the spans in $\hat{\mathcal{S}}$ are proper, the starting points of the spans are distinct, and so are the ending points (see Remark 3.8). Let $\hat{\mathcal{A}}$ be the set of starting points of $\hat{\mathcal{S}}$, and let $\hat{\mathcal{B}}$ be the set of ending points. We can now use the formulas in Theorem 3.9 on $T^\perp$. Thus, $\hat{\epsilon}_i = \xi_i + I_i^{\hat{\mathcal{A}}} = \xi_{i+1} + I_i^{\hat{\mathcal{B}}}$ for all $i \in \mathcal{I}$, and hence Lemma 8.8 implies that $\hat{\mathcal{A}} \subseteq \mathcal{I}\backslash\{b_1,\dots,b_k\}$ and $\hat{\mathcal{B}} \subseteq \mathcal{I}\backslash\{a_1,\dots,a_k\}$. Since $|\hat{\mathcal{A}}| = |\hat{\mathcal{B}}| = n-k$, we have equality in both cases. Thus $T^\perp = T_{\hat{H},\hat{\mathcal{S}}}$ where $\hat{\mathcal{S}}$ has the desired properties. Since $T^\perp$ is nonmergeable, we have $T^\perp \cong T_{(\hat{H},G,\hat{\mathcal{S}})}$ by Theorem 7.10.

"(c) ⇒ (d)" By Theorem 3.9 and the fact that $\hat{\mathcal{A}} = \mathcal{I}\backslash\{b_1,\dots,b_k\}$ we obtain the following.

$$\hat{\epsilon}_i = \xi_i + I_i^{\hat{\mathcal{A}}} = \xi_i + 1 - I_i^{\mathcal{B}} = \xi_i + 1 + \xi_{i+1} - \epsilon_i$$

Therefore, by Theorem 8.3, we get that $\hat{\epsilon}_i = \dim(E_i)^\circ$ where $(E_i)^\circ$ is the i-th transition space of T°. Hence, $T^\perp \cong T^\circ$ by Proposition 8.6.

"(d) ⇒ (e)" Since $T^\perp \cong T^\circ$ we know that ECP($T^\perp$)=ECP(T°) $= (\hat{\epsilon}_0,\dots,\hat{\epsilon}_{n-1})$. Now, $\hat{\epsilon}_i = \operatorname{rk}(N_i^\top, H_i^\top, N_{i+1}^\top) = \operatorname{rk}(N_i^\top, H_i^\top)$ and thus

$$\hat{\epsilon}_i = \begin{cases} \xi_i & \text{if } H_i^\top \in \operatorname{col}(N_i^\top) \\ \xi_i + 1 & \text{if } H_i^\top \notin \operatorname{col}(N_i^\top) \end{cases},$$

where col() denotes the column space of the matrix. By recognizing that $\operatorname{col}(N_i^\top) = \operatorname{im} N_i$ (up to transposition) and using (8.2), we obtain

$$\epsilon_i = \begin{cases} \xi_{i+1} + 1 & \text{if } H_i \in \operatorname{im}(N_i) \\ \xi_{i+1} & \text{if } H_i \notin \operatorname{im}(N_i) \end{cases}$$

Hence, by Theorem 3.9, we obtain the desired equivalence.

"(e) $\Rightarrow$ (f)" Note that T satisfies the first property of Lemma 8.9 due to Proposition 7.9. By Theorem 3.9, we know that for all $i \in \mathcal{I}$, $\operatorname{row}(M_i, l)$ is zero exactly when $i \notin (a_l, b_l]$ and the non-zero rows of M_i are linearly independent because they are distinct standard basis vectors. Proposition 7.9 implies that if $\operatorname{row}(M_i, l)$ is zero, then $\operatorname{row}(N_i, l)$ is also zero. Now $T_{G,\mathcal{S}} \cong T_{(G,H,\mathcal{S})}$ gives that $\operatorname{rk} M_i = \operatorname{rk} N_i$ for all $i \in \mathcal{I}$. Thus the nonzero rows of N_i are the rows for which $i \in (a_l, b_l]$, and they are linearly independent for all $i \in \mathcal{I}$. Since the third property of Lemma 8.9 is our assumption, part (f) follows.

"(f) $\Rightarrow$ (b)" Recall from Proposition 2.2 that $T^\perp = (\hat{V}, \hat{E})$ is weakly connected if for all $v \in \hat{V}_0$ there exists a path of length n from v to $0 \in \hat{V}_0$. We will show the existence of such a path. Let $v \in \operatorname{im}(N_0)^\top = \hat{V}_0$. Then $v = \alpha(N_0)^\top$ for some $\alpha \in \mathbb{F}^{n-k}$. Taking duals of the statement in (f) gives $\sum_{i=0}^{n-1} \ker(N_i)^\top = \mathbb{F}^{n-k}$. Thus we may write $\alpha = \sum_{i=0}^{n-1} \alpha^{(i)}$ where $\alpha^{(i)} \in \ker(N_i)^\top$ for all $i \in \mathcal{I}$. Then $\sum_{i=r}^{n-1} \alpha^{(i)}(N_r)^\top = \sum_{i=r+1}^{n-1} \alpha^{(i)}(N_r)^\top$ for all $r = 0, ..., n-1$. Because $v = (\alpha - \alpha^{(0)})(N_0)^\top = \sum_{i=1}^{n-1} \alpha^{(i)}(N_0)^\top$ and $\alpha^{(n-1)}(N_{n-1})^\top = 0$, we get that the edges

$$\left(\sum_{i=r}^{n-1} \alpha^{(i)}((N_{r-1})^\top, {H_{r-1}}^\top, (N_r)^\top) \right)_{r=1,\dots,n-1}$$

form a path from v to $0 \in V_{n-1}$. This path can then be extended to the desired path.

"(d) $\Rightarrow$ (g)" This is clear since $T^\perp$ is edge-reduced.

"(g) $\Rightarrow$ (d)" Since Proposition 8.6 gives that $\hat{E}_j \subseteq (E_j)^\circ$ for $j \in \mathcal{I}$, it remains to show that every edge in T° is an edge in $T^\perp$. Consider a given edge in T°. Since T° is edge reduced, this edge in T° is part of a cycle whose edge-label sequence is a codeword in $\mathcal{C}^\perp$. This codeword is also represented by a cycle in $T^\perp$. By Proposition 8.6, we know that this cycle also appears in T°. Since T° is one-to-one, this implies that these cycles coincide, and thus the given edge in T° is also in $T^\perp$.

□

Corollary 8.11. *KV-trellises satisfy the properties (a)-(g) of Theorem 8.10. In particular, we have that the local dual and the BCJR-dual coincide.*

Proof. Note that by Theorem 7.12, KV-trellises satisfy the hypothesis of Theorem 8.10 as well as part (e). □

While Theorem 8.10 applies to KV-trellises, this theorem does indeed apply to a larger class of trellises. Part (a) of the following example, which appears as Example III.6 in [7], provides such a trellis.

Example 8.12. *(a) Let $\mathcal{C} = \operatorname{im} G = \ker H^\top \subseteq \mathbb{F}_2^6$, where*

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \; H = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Consider the span list $\mathcal{S} = [(1,5],(2,4],(3,1]]$ for G. One can easily verify that $(1,5]$ and $(2,4]$ are characteristic spans of $\mathcal{C}$, but $(3,1]$ is not (there exists a codeword with span $(3,0]$). Hence the product trellis $T_{G,\mathcal{S}}$ is not a KV-trellis. By straightforwardly computing the data for $T_{G,\mathcal{S}}$ and the corresponding BCJR-trellis $T := T_{(G,H,\mathcal{S})}$, one obtains that both trellises have SCP $(1,1,1,2,3,2)$. Hence they are isomorphic due to Theorem 7.10. Their ECP is $(1,2,2,3,3,2)$. The displacement matrix of $T_{(G,H,\mathcal{S})}$ is given by

$$N_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let us now consider the BCJR-trellis $T_{(H,G,\hat{\mathcal{S}})}$ of $\mathcal{C}^\perp$, where $\hat{\mathcal{S}} = [(2,4],(3,0],(0,5]]$ is the chosen span list for the rows of H. Its displacement matrix turns out to be $N_0^\top$. As a consequence, $T^\perp = T_{(G,H,\mathcal{S})}{}^\perp = T_{(G,H,N_0)}{}^\perp = T_{(H,G,N_0^\top)} = T_{(H,G,\hat{\mathcal{S}})}$. One can also easily verify that $T_{(H,G,\hat{\mathcal{S}})} \cong T_{H,\hat{\mathcal{S}}}$. Now it is easy to check that both $T^\perp$ and T° have ECP $(2,1,2,3,3,2)$. Thus Proposition 8.6 yields $T^\perp \cong T^\circ$.

(b) The condition that $\operatorname{rk} G = k$ is also a crucial assumption, and plays a part in the proof of "(e) $\Rightarrow$ (f)", see the proof of Theorem II.13 in [7]. In the case where $\operatorname{rk} G = r > k$, the following example shows that we may have $T^\perp \cong T^\circ$ even if $T^\perp$ is not connected. Let $\mathcal{C}$ be as in Example 7.11 (b), that is, $\mathcal{C} = \operatorname{im} G = \ker H^\top \subseteq \mathbb{F}_2^3$, where

$$G = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \; H = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

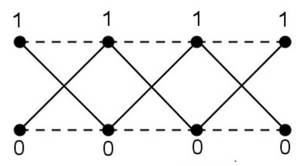
Consider the span list $\mathcal{S} = [(0,2],(1,0]]$ for G. Then the span-based BCJR-trellis $T = T_{(G,H,\mathcal{S})}$ has state and transition spaces given by $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im}(N_j, G_j^\top, N_{j+1})$, where

$$(N_0|G_0^\top|N_1|G_1^\top|N_2|G_2^\top|N_0) = \left(\begin{array}{c|c|c|c|c|c|c} 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{array}\right).$$

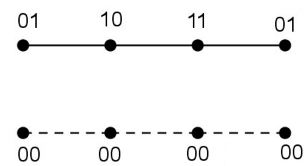
The trellises T and $T^\perp$ are shown in Figure 8.3. Using (8.2), it is easy to see that ECP(T°)=ECP($T^\perp$), and thus the two trellises are isomorphic. However, one can see that this dual trellis is not connected. In this case, the trellis T is not isomorphic to the product trellis $T_{G,\mathcal{S}}$ (which is shown in Figure 3.2). The trellis T is in fact isomorphic to the product trellis based on the matrix

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

along with the span set $\mathcal{S}' = [(0,1], (1,2], (2,0]]$. It is easy to see that the matrix G' does not have full row rank.



$[T = T_{(G,H,\mathcal{S})}]$



$[T^{\perp} \cong T^{\circ}]$

Figure 8.3: A span-based BCJR-trellis and its dual

While Theorem 8.10 does apply to KV-trellises, we can get an even nicer result for this class of trellises. Through the following lemma and theorem, we will show that not only are the BCJR and local duals of a KV-trellis isomorphic, the resulting trellis is also a KV-trellis of the dual code.

Lemma 8.13. *Let $\hat{\mathcal{S}} = [(\hat{a}_l, \hat{b}_l] : l = 1, \ldots, k]$ be a set of k distinct characteristic spans of $\mathcal{C}$, and set $\hat{\mathcal{A}} = \{\hat{a}_1, \ldots, \hat{a}_k\}$, and $\hat{\mathcal{B}} = \{\hat{b}_1, \ldots, \hat{b}_k\}$. Suppose that $T = T_{G,\mathcal{S}}$ is a product trellis for $\mathcal{C}$ with $\mathcal{S} = [(a_l, b_l] : l = 1, \ldots, k]$ where the a_l's are distinct and the b_l's are distinct, and set $\mathcal{A} = \{a_1, \ldots, a_k\}$, and $\mathcal{B} = \{b_1, ..., b_k\}$. If ECP($T$)-SCP($T$)=$I^{\hat{\mathcal{A}}}$, ECP($T$)-$\sigma$(SCP($T$))=$I^{\hat{\mathcal{B}}}$, and $|\{l \mid 0 \in (\hat{a}_l, \hat{b}_l]\}| = |\{l \mid 0 \in (a_l, b_l]\}|$, then $\mathcal{S} = \hat{\mathcal{S}}$, and thus, T is a KV-trellis.*

Proof. By Theorem 3.9, we have that $I^{\mathcal{A}} =$ ECP(T)-SCP(T)=$I^{\hat{\mathcal{A}}}$ and $I^{\mathcal{B}} =$ ECP(T)-σ(SCP(T))=$I^{\hat{\mathcal{B}}}$. Thus, we obtain $\mathcal{A} = \hat{\mathcal{A}}$ and $\mathcal{B} = \hat{\mathcal{B}}$. Without loss of generality, let $a_l = \hat{a}_l$ for all $l = 1, \ldots, k$. Then, there exists a permutation $\phi : \{1, \ldots, k\} \to \{1, \ldots, k\}$ such that $[(a_l, b_{\phi(l)}] : l = 1, \ldots, k] = \hat{\mathcal{S}}$. If this permutation ϕ is not the identity map, then $b_j < b_{\phi(j)}$ for some $j \in \{1, \ldots, k\}$. Thus if $(a_j, b_j]$ and $(a_j, b_{\phi(j)}]$ are both conventional or both circular, we get that $(a_j, b_j] \subsetneq (a_j, b_{\phi(j)}]$. Since $(a_j, b_{\phi(j)}] \in \hat{\mathcal{S}}$ is a characteristic span, this is a contradiction. Thus we must have $b_j < a_j < b_{\phi(j)}$, and the permutation ϕ takes the circular span $(a_j, b_j] \in \mathcal{S}$ to the conventional span $(a_j, b_{\phi(j)}] \in \hat{\mathcal{S}}$. Now, the condition $|\{l \mid 0 \in (\hat{a}_l, \hat{b}_l]\}| = |\{l \mid 0 \in (a_l, b_l]\}|$ implies that $\mathcal{S}$ and $\hat{\mathcal{S}}$ have the same number of circular spans, and hence they have the same number of conventional spans as well. Thus, ϕ must also take a conventional span in $\mathcal{S}$, say $(a_m, b_m]$, to a circular span in $\hat{\mathcal{S}}$. In this case, we get that $(a_m, b_m] \subsetneq (a_m, b_{\phi(m)}]$, which is a contradiction because $(a_m, b_{\phi(m)}]$ is a characteristic span. Therefore, we must have that ϕ is the identity map, and thus $\mathcal{S} = \hat{\mathcal{S}}$. $\square$

Theorem 8.14. *Suppose that $\mathcal{T} := [(a_l, b_l] : l = 1, ..., n]$ is the set of characteristic spans of $\mathcal{C}$, and $T = T_{(G,H,S)} \cong T_{G,S}$ is a KV-trellis for the code $\mathcal{C}$ where $\mathcal{S} := [(a_l, b_l] : l = 1, ..., k]$. Then the dual trellis $T' = T^{\perp} \cong T^{\circ}$ is a KV-trellis for $\mathcal{C}^{\perp}$ based on the span set $\hat{\mathcal{S}} = [(b_l, a_l] : l = k+1, \ldots, n]$.*

Proof. By Corollary 8.11, T satisfies the hypothesis and (a)-(g) of Theorem 8.10. Set $[(b_l, a_l] : l = k+1, \ldots, n] = [(\hat{a}_l, \hat{b}_l]; l = 1, \ldots, n-k]$, $\mathcal{A} = \{a_1, \ldots, a_k\}$, and $\mathcal{B} = \{b_1, \ldots, b_k\}$. Then $\hat{\mathcal{A}} = \mathcal{I} - \hat{\mathcal{B}}$ and $\hat{\mathcal{B}} = \mathcal{I} - \hat{A}$ as in Theorem 8.10. Since T' is a BCJR-trellis, we have that T' is linear, reduced, and biproper by Definition/Theorem 7.1. Since T is a KV-trellis, by Corollary 8.11, we also know that T' is connected. Additionally, since T is connected, we have that T' is one-to-one by Proposition 8.5. Thus, T' is a product trellis of the form $T_{H',\mathcal{S}'}$ where $\mathcal{C}^\perp = \operatorname{im} H'$, $\operatorname{rk} H' = n-k$, $\mathcal{S}' = [(a'_l, b'_l] : l = 1, ..., n-k]$, the starting points of the spans in $\mathcal{S}'$ are distinct, and the ending points of the spans in $\mathcal{S}'$ are distinct.

Note that by Lemma 6.5, the set $\hat{\mathcal{S}}$ consists of characteristic spans of $\mathcal{C}^\perp$. Because $T' = T^\perp$, the proof of "(a)$\Rightarrow$ (c)" in Theorem 8.10 gives that ECP(T')-SCP(T')$=I^{\hat{\mathcal{A}}}$ and ECP(T')-σ(SCP(T'))$=I^{\hat{\mathcal{B}}}$. Now, let SCP(T)$=(\xi_0, \ldots, \xi_{n-1})$, and note that by Theorem 3.9 we have that ξ_0 is equal to the number of circular spans in $\mathcal{S}$. By Proposition 5.3, we know that the total number of circular spans in $\mathcal{T}$ is equal to $n-k$. Thus, the number of circular spans in $\mathcal{S}$ plus the number of conventional spans in $\hat{\mathcal{S}}$ is equal to $n-k$. Since the total number of spans in $\hat{\mathcal{S}}$ is also equal to $n-k$, we have that the number of circular spans in $\mathcal{S}$ is equal to the number of circular spans in $\hat{\mathcal{S}}$. Since we know that SCP(T')=SCP(T), we obtain that $\mathcal{S}'$ and $\hat{\mathcal{S}}$ contain the same number of circular spans or $|\{l \mid 0 \in (a'_l, b'_l]\}| = |\{l \mid 0 \in (\hat{a}_l, \hat{b}_l]\}|$. Thus by Lemma 8.13, we get that $\mathcal{S}' = \hat{\mathcal{S}}$. $\square$

Theorem 8.14 settles a conjecture posed by Koetter and Vardy in Chapter V of [11].

We conclude this thesis with a brief discussion of which types of minimality are preserved under dualization. So far we have defined minimality in Definition 2.3 and product minimality in Definition 6.9. We would like to introduce one more notion of minimality, defined below.

Definition 8.15. *A trellis $T = (V, E)$ for a code $\mathcal{C}$ is called edge-minimal if there does not exist a linear trellis $T' = (V', E')$ for $\mathcal{C}$ such that $|E'_i| \leq |E_i|$ for all $i \in \mathcal{I}$ and $|E'_j| < |E_j|$ for some $j \in \mathcal{I}$.*

Now Koetter and Vardy showed in [11, Theorem 5.6] that all tail-biting trellises that are minimal with respect to any of the orders we have introduced (as well as several others) are KV-trellises. Thus, for these trellises, the local and BCJR-dual trellises are isomorphic, and we will simply refer to the dual trellis. By Theorem 8.14, this dual trellis is a KV-trellis. Since minimality and product-minimality are based solely on the SCP of the trellis, it is clear that the dual of a (product-)minimal trellis of $\mathcal{C}$ is a (product-)minimal trellis of $\mathcal{C}^\perp$. However, in the following example, we can see that edge-minimality is not preserved under dualization.

Example 8.16. *Let $\mathcal{C} := \operatorname{im} G \subseteq \mathbb{F}_2^{3\times 5}$, where*

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

The characteristic spans for $\mathcal{C}$ are $(0,3], (1,2], (3,4], (2,0]$, and $(4,1]$. We have 10110 and 11010 as the characteristic generators with the span $(0,3]$, 01100 with span $(1,2]$, 00011 with span $(3,4]$, 10110 and 10101 with span $(2,0]$, and 11001 with span $(4,1]$. Thus, it is possible to construct a KV-trellis for $\mathcal{C}$ by choosing any three of these spans and corresponding linearly independent characteristic generators. Out of the 10 possible KV-trellises (up to structural isomorphism) for $\mathcal{C}$, five are edge-minimal. For example, one can easily verify that the trellis $T_{G_1,\mathcal{S}}$, where

$$G_1 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \text{ and } \mathcal{S} = [(2,0],(3,4],(4,1]]$$

is an edge-minimal trellis with ECP of (2,1,1,2,3).

Since we are dealing with KV-trellises, dualizing using either the local dual or the BCJR-dual will result in a KV-trellis for $\mathcal{C}^\perp$. Thus there are also 10 KV-trellises (up to structural isomorphism) for $\mathcal{C}^\perp$. However, it is easy to check that only three of the KV-trellises for $\mathcal{C}^\perp$ are edge-minimal. For example, consider the dual of the trellis of $T_{G_1,\mathcal{S}}$. It has an ECP of (2,1,1,2,2); however, it is not an edge-minimal trellis for $\mathcal{C}^\perp$ since the KV-trellis for $\mathcal{C}$ with spans $(1,4]$ and $(3,0]$ has an ECP of (1,1,1,2,2).

Appendix: Maple Procedure

The following code is a procedure for Maple that returns the number of normalized characteristic matrices for a given code, $\mathcal{C} = \operatorname{im} G$, as well as the set of codewords that fit each characteristic span as normalized characteristic generators. To run this procedure, one must first load the package linalg. It also calls on several procedures written by Heide Gluesing-Luerssen which are described below. Prior to running the procedure the finite field $\mathbb{F}_p$ is computed and stored as F. This is then used in the procedure VecSpace() to generate the subspace $\mathbb{F}_p^k$ which is stored as Fk. The procedure NormMat() ensures that the entries of a specified matrix are remain in our finite field, CharMat() produces a characteristic pair for the given code, and SortA() sorts the characteristic spans in increasing order by starting point.

```
NumCharMat:=proc(G)

global p, F;
local C, CM, A, B, i, j, l, Code, k, n, Fk, N;

k:=rowdim(G):
n:=coldim(G):
Fk:=VecSpace(F,k):
Code:=[]:
for i from 1 to nops(Fk) do
    c||i:=NormMat(Fk[i]&*G):
    Code:=[op(Code),evalm(c||i)]
od:
CM:=CharMat(G):
C:=SortA(CM):
A:=[seq(C[2][j,1],j=1..n)]:
B:=[seq(C[2][j,2],j=1..n)]:
for i from 1 to n do
    C||(i-1):=[]:
    if A[i]<=B[i] then
        for j from 1 to nops(Code) do
            if (Code[j][A[i]+1]=1 and Code[j][B[i]+1]<>0 and
                seq(Code[j][l],l=1..A[i])=seq(0,l=1..A[i])
                and seq(Code[j][l],l=B[i]+2..n)=seq(0,l=B[i]+2..n))
            then C||(i-1):=[op(C||(i-1)),evalm(Code[j])]
            else
            fi:
        od:
    else
        for j from 1 to nops(Code) do
```

```
            if (Code[j][A[i]+1]=1 and Code[j][B[i]+1]<>0 and
                seq(Code[j][l],l=B[i]+2..A[i])=seq(0,l=B[i]+2..A[i]))
            then C||(i-1):=[op(C||(i-1)),evalm(Code[j])]
            else
            fi:
        od:
    fi:
od:
for i from 1 to n do
    N||(i-1):=nops(C||(i-1)) od:
N:=1:
for i from 1 to n do
    N:=N*N||(i-1) od:
RETURN(N,transpose(matrix(2,n,[[seq([A[i],B[i]],i=1..n)],
[seq(C||(i-1),i=1..nn)]])));

end proc:
```

Bibliography

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

[2] A. R. Calderbank, G. D. Forney, Jr., and A. Vardy. Minimal tail-biting trellises: The Golay code and more. *IEEE Trans. Inform. Theory*, IT-45:1435–1455, 1999.

[3] D. Conti and N. Boston. Factoring linear trellises. In *Proceedings of the International Zurich Seminar on Communications*, pages 83–86, 2012.

[4] G. D. Forney, Jr. Coset codes II: Binary lattices and related codes. *IEEE Trans. Inform. Theory*, IT-34:1152–1187, 1988.

[5] G. D. Forney, Jr. Codes on graphs: Normal realizations. *IEEE Trans. Inform. Theory*, IT-47:520–548, 2001.

[6] G. D. Forney, Jr. and H. Gluesing-Luerssen. Codes on graphs: Observability, controllability and local reducibility. Preprint 2012. Submitted. ArXiv: cs.IT/1203.3115v1.

[7] H. Gluesing-Luerssen and E. Weaver. Characteristic generators and dualization for tail-biting trellises. *IEEE Trans. Inform. Theory*, IT-57:7418–7430, 2011.

[8] H. Gluesing-Luerssen and E. Weaver. Linear tail-biting trellises: Characteristic generators and the BCJR-construction. *IEEE Trans. Inform. Theory*, IT-57:738–751, 2011.

[9] H. Kan and H. Shen. A relation between the characteristic generators of a linear code and its dual. *IEEE Trans. Inform. Theory*, IT-51:1199–1202, 2005.

[10] R. Koetter and A. Vardy. On the theory of linear trellises. In M. Blaum, P. G. Farrel, and H. C. A. van Tilborg, editors, *Information, Coding and Mathematics*, pages 323–354. Kluwer, Boston, 2002.

[11] R. Koetter and A. Vardy. The structure of tail-biting trellises: Minimality and basic principles. *IEEE Trans. Inform. Theory*, IT-49:2081–2105, 2003.

[12] F. R. Kschischang and V. Sorokine. On the trellis structure of block codes. *IEEE Trans. Inform. Theory*, IT-41:1924–1937, 1995.

[13] R. J. McEliece. On the BCJR Trellis for linear block codes. *IEEE Trans. Inform. Theory*, IT-42:1072–1092, 1996.

[14] D. J. Muder. Minimal trellises for block codes. *IEEE Trans. Inform. Theory*, IT-34:1049–1053, 1988.

[15] A. V. Nori and P. Shankar. Unifying views of tail-biting trellis constructions for linear block codes. *IEEE Trans. Inform. Theory*, IT-52:4431–4443, 2006.

[16] A. Vardy. Trellis structure of codes. In V. S. Pless and W. C. Huffman, editors, *Handbook of Coding Theory, Vol. 2*, pages 1989–2117. Elsevier, Amsterdam, 1998.

Vita

Elizabeth A. Weaver
Born in Quincy, IL
July 3, 1984

Education

- **University of Kentucky**, Lexington, Kentucky
  - M.A., Mathematics, July 2008
- **Southeast Missouri State University**, Cape Girardeau, Missouri
  - B.S., Summa cum Laude, Mathematics, May 2006
  - B.A., Summa cum Laude, Music, May 2006

Teaching Experience

- **Teaching assistant:**
  - University of Kentucky, August 2006 - May 2012

Awards and Fellowships

- **University of Kentucky**, Lexington, Kentucky
  - Wimberly C. Royster Teaching Assistant Award, 2010
  - Max Steckler Graduate Fellowship, 2009 and 2011
  - Daniel R. Reedy Quality Achievement Fellowship, Fall 2006- Spring 2009
- **Southeast Missouri State University**, Cape Girardeau, Missouri
  - Southeast Missouri State Governor's Scholarship, Fall 2002 - Spring 2006
  - Robert C. Byrd Honors Scholarship, Fall 2002 - Spring 2006
  - Missouri Federation of Music Clubs Award, Spring 2004

Publications

- *Characteristic Generators and Dualization for Tail-Biting Trellises*, with Heide Gluesing-Luerssen, IEEE Transactions on Information Theory, Vol. 57, pp 7418-7430, November 2011
- *Linear Tail-biting Trellises: Characteristic Generators and the BCJR-Construction*, with Heide Gluesing-Luerssen, IEEE Transactions on Information Theory, Vol. 57, pp 738-751, February 2011.