Characteristic Generators and Dualization for Tail-Biting Trellises

Heide Gluesing-Luerssen^{*} and Elizabeth A. Weaver University of Kentucky Department of Mathematics 715 Patterson Office Tower Lexington, KY 40506-0027, USA; heide.gl@uky.edu, eaweaver1s@uky.edu

January 3, 2011

Abstract: This paper focuses on dualizing tail-biting trellises, particularly KV-trellises. These trellises are based on characteristic generators, as introduced by Koetter/Vardy (2003), and may be regarded as a natural generalization of minimal conventional trellises, even though they are not necessarily minimal. Two dualization techniques will be investigated: the local dualization, introduced by Forney (2001) for general normal graphs, and a linear algebra based dualization tailored to the specific class of tail-biting BCJR-trellises, introduced by Nori/Shankar (2006). It turns out that, in general, the BCJR-dual is a subtrellis of the local dual, while for KV-trellises these two coincide. Furthermore, making use of both the BCJR-construction and the local dualization, it will be shown that for each complete set of characteristic generators of a code there exists a complete set of characteristic generators of the dual code such that their resulting KV-trellises are dual to each other if paired suitably. This proves a stronger version of a conjecture formulated by Koetter/Vardy.

Keywords: linear block codes, tail-biting trellises, characteristic generators, tail-biting BCJR-trellises, KV-trellises, dualization

MSC (2000): 94B05, 94B12, 68R10, 93B20

1 Introduction

It is well known that for a given linear block code, a tail-biting trellis may be smaller than the minimal conventional trellis with respect to any of the various notions of complexity of a trellis; see the discussion in [10, Sec. III]. Since iterative decoding on tail-biting trellises is well understood (as opposed to decoding on more general graphs with cycles), this has led to an increased interest in the construction of minimal tail-biting trellises; see also [10, 12, 15, 16, 17, 18]. A major breakthrough in this direction has been obtained by Koetter/Vardy [10]. They showed that for each k-dimensional linear block code of length n with full support there exists a list of n characteristic generators, each endowed with a span interval, such that every minimal tail-biting trellis of the code is structurally isomorphic to a product of the elementary trellises of k linearly independent characteristic generators (where structurally

^{*}The author was partially supported by National Science Foundation grant #DMS-0908379

isomorphic means trellis isomorphic, but disregarding the edge labels); see [10, Thm. 5.5] and the adjustments in [8, Prop. III.14, Thm. III.15]. Here minimality may refer to any of the standard complexity notions for tail-biting trellises discussed in [10, Thm. 5.6]. Moreover, the same results show that for each minimal trellis there is a suitable choice of k linearly independent characteristic generators such that the given trellis is isomorphic to the resulting product trellis. This shows a major difference from conventional trellises: while the minimal conventional trellis of a block code is unique up to trellis isomorphism, this is not the case in the tail-biting situation – again, this refers to any minimality notion. Not only may there exist minimal trellises with incomparable state complexity profiles or edge complexity profiles, but even if both these profiles coincide for two minimal trellises, the trellises may not be isomorphic (but they are structurally isomorphic in this situation due to [8, Prop. III.14]).

We will use the term KV-trellises for product trellises based on k linearly independent characteristic generators in the sense described above. From the construction performed in [10], it follows that characteristic generators may be regarded as a generalization of MSGM's or trellis-oriented generator matrices in the realm of conventional trellises, see [14, Def. 6.2] or [11, Sec. IV], or shortest bases in the sense of [5]; see also Lemma 2.6 and Remark 2.7 in the next section. As a consequence, KV-trellises, though not minimal in general, form a natural generalization of minimal conventional trellises, and indeed, these trellises have much nicer properties than more general tail-biting trellises. In the paper [8], two trellis constructions based on generators with span intervals have been investigated: the product construction and the BCJR-construction, introduced by Nori/Shankar [15]. It has been shown that, in general, a BCJR-trellis is smaller than the corresponding product trellis. In fact, the latter can be merged to the former by taking suitable quotients of its state spaces [8, Thm. IV.9]. For KV-trellises, however, these two constructions are isomorphic [8, Thm. IV.11]. As a consequence, KV-trellises are non-mergeable. Another demonstration of the distinctiveness of KV-trellises will be given in this paper. It will be shown that KVtrellises behave significantly nicer under dualization than more general trellises.

We will investigate two dualization techniques for tail-biting trellises. Both lead to trellises representing the dual code. The first construction is a specialization of the local dualization introduced by Forney in [4] for general normal graphs. It amounts to dualizing the transition spaces along with a sign inverter; see also [6] for a different approach based on graphical models. This dualization has been generalized to factor graphs in [13] and recently been recast in the framework of Valiant transforms [1]. Even though the local dualization is a very elegant and convenient construction, for tail-biting trellises it may lead to dual trellises with some undesirable properties; see Example 3.2. The second construction is a simple linear algebra based dualization for BCJR-trellises as introduced by Nori/Shankar [15]. In Section 3 we will see that the BCJR-dual is a subtrellis of the local dual. For KV-trellises, however, these duals coincide. Furthermore, as we will show in Section 4, the dual of a KV-trellis is a KV-trellis again and thus shares all their nice properties.

More specifically, in Section 4 we will prove that for each set of n characteristic generators of a given code $\mathcal{C} \subseteq \mathbb{F}^n$, there exists a set of n characteristic generators of the dual code \mathcal{C}^{\perp} such that the dual of each KV-trellis of \mathcal{C} based on the chosen generators is (isomorphic to) a KV-trellis of \mathcal{C}^{\perp} based on the dual generators. We will construct the list of dual generators explicitly and also show the direct link between the k linearly independent characteristic generators for \mathcal{C} and the n-k dual characteristic generators that give rise to the dual KV-trellis. In fact, this link is easily described because Koetter/Vardy have shown [10, Thm. 5.12] that the characteristic span list of \mathcal{C}^{\perp} is simply obtained by reversing the characteristic spans of \mathcal{C} . The construction of the dual list of characteristic generators is performed as follows. One starts with the BCJR-trellis of \mathcal{C} based on the entire list of chosen characteristic generators. This trellis gives rise to n subtrellises obtained by omitting exactly one characteristic generator. They still represent the code C. The main result of the procedure tells us that the local dual of each such subtrellis contains a cycle that gives rise to a dual characteristic generator whose span is the reversal of the span that had been omitted. All of this leads to the desired n dual characteristic generators. With the principles of local dualization as well as Koetter/Vardy's result of reversed characteristic spans in mind, this approach of constructing dual characteristic generators is quite natural. The details however, carried out in Section 4, become rather technical due to certain linear independence conditions that need to be verified.

Let us close the introduction with introducing the basic notions needed for this paper. Throughout, a tail-biting trellis T = (V, E) of depth n over the finite field \mathbb{F} is a directed edge-labeled graph with the property that the vertex set V partitions into n disjoint sets $V = V_0 \cup V_1 \cup \ldots \cup V_{n-1}$ such that every edge in T that starts in V_i ends in $V_{i+1 \mod n}$. The edges are labeled with field elements from \mathbb{F} . Notice that we compute modulo n on the time axis $\mathcal{I} := \{0, \ldots, n-1\}$. Referring to the fact that a trellis is a state space realization of the code regarded as a dynamical system (behavior) on \mathcal{I} , we call V_i the state space of the trellis at time i, and its elements are the states at that time. The edge set E decomposes into $E = \bigcup_{i=0}^{n-1} E_i$, where E_i is the set of edges starting in V_i and ending in $V_{i+1 \mod n}$. Its elements reflect the present-state to next-state transitions, and therefore the edge sets E_i will be called transition spaces. We identify the elements of E_i (the edges) with the triples consisting of starting state, label, and ending state. Thus, the transition spaces are given by $E_i = \{(v, a, \hat{v}) \mid$ there exists an edge $v \xrightarrow{a} \hat{v}$ where $v \in V_i$, $\hat{v} \in V_{i+1}$, $a \in \mathbb{F}\} \subseteq V_i \times \mathbb{F} \times V_{i+1}$ for $i \in \mathcal{I}$. These spaces have also been called trellis sections [3, 7] or local constraints [4].

A cycle in T is a closed path of length n in the trellis. We always assume that the cycles start and end (at the same state) in V_0 . If $|V_0| = 1$, the trellis is called *conventional*. We call the trellis *reduced* if every state and every edge appear in at least one cycle. The trellis is called *biproper* if any two edges starting at the same vertex or ending at the same vertex are labeled distinctly.

The trellis T is linear if each state space V_i is a vector space over \mathbb{F} and the label code

$$\mathcal{S}(T) = \{ (v_0, \dots, v_{n-1}, c) \in V_0 \times \dots \times V_{n-1} \times \mathbb{F}^n | v_0 \xrightarrow{c_0} v_1 \xrightarrow{c_1} \dots v_{n-1} \xrightarrow{c_{n-1}} v_0 \text{ is a cycle in } T \}$$
(1.1)

is a subspace of $V_0 \times \ldots \times V_{n-1} \times \mathbb{F}^n$ and if the transition spaces E_i are linear subspaces of $V_i \times \mathbb{F} \times V_{i+1}$. If the trellis is reduced, then the linearity of E_i follows from the linearity of $\mathcal{S}(T)$. We say that T represents the code $\mathcal{C} \subseteq \mathbb{F}^n$ if \mathcal{C} equals its edge-label code, that is, $\mathcal{C} = \{(c_0, \ldots, c_{n-1}) \in \mathbb{F}^n \mid \text{there exists a cycle } v_0 \xrightarrow{c_0} v_1 \xrightarrow{c_1} \dots \xrightarrow{c_{n-1}} v_0 \text{ in } T\}.$

Note that if T is linear, then the represented block code is linear. We will only deal with linear block codes and linear trellis representations. The trellis T is called one-toone if distinct cycles in T have distinct edge-label sequences. The state complexity profile (SCP) and edge complexity profile (ECP) of a linear trellis T = (V, E) are defined as $SCP(T) \coloneqq (s_0, \ldots, s_{n-1})$, where $s_i = \dim V_i$, and $ECP(T) \coloneqq (e_0, \ldots, e_{n-1})$, where $e_i = \dim E_i$. Throughout this paper, the notion of minimality for tail-biting trellises refers to any of the orderings discussed by Koetter/Vardy in [10, Sec. III]. While for conventional trellises all these minimality notions coincide, this is not the case for tail-biting trellises. In [10, Thm. 5.5, Thm. 5.6] and [8, Thm. III.15] it has been shown that a minimal trellis (with respect to any of those orderings) is a KV-trellis in the sense of our Definition 2.4. In this paper we will be concerned with KV-trellises, and a specific notion of minimality will not be needed.

Linear trellises T = (V, E) and T' = (V', E') are called *isomorphic* if there exists a bijection $\phi : V \longrightarrow V'$ such that $\phi(V_i) = V'_i$ and $\phi|_{V_i} : V_i \longrightarrow V'_i$ is an isomorphism for all

 $i \in \mathcal{I}$ and $(v, a, w) \in E_i$ if and only if $(\phi(v) a, \phi(w)) \in E'_i$. Obviously, isomorphic trellises represent the same code.

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

$$\mathcal{C} = \operatorname{im} G = \operatorname{ker} H^{\mathsf{T}} \subseteq \mathbb{F}^n \text{ be a } k \text{-dimensional code with support } \mathcal{I} = \{0, \dots, n-1\}, \qquad (1.2)$$

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, 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 rk $G = k \leq r$, 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

$$G = (g_{lj})_{\substack{l=1,\dots,r\\j=0,\dots,n-1}} = \begin{pmatrix} G_0^\mathsf{T} & \dots & G_{n-1}^\mathsf{T} \end{pmatrix} \in \mathbb{F}^{r \times n} \text{ and } H^\mathsf{T} = \begin{pmatrix} H_0 \\ \vdots \\ H_{n-1} \end{pmatrix} \in \mathbb{F}^{n \times (n-k)}.$$
(1.3)

Hence $G_j^{\mathsf{T}} \in \mathbb{F}^r$ and $H_j^{\mathsf{T}} \in \mathbb{F}^{n-k}$ are the columns of G and H, respectively. Finally, in order to avoid extreme cases, we will also assume that \mathcal{C}^{\perp} has support \mathcal{I} . As for the matrices G and H above, we will use the notation M_j^{T} 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.

2 KV-Trellises and the BCJR-Construction

We will begin by briefly recalling the main results about products of (tail-biting) elementary trellises. Thereafter we turn to KV-trellises, the product trellises obtained by choosing k linearly independent characteristic generators of the code as introduced by Koetter/Vardy in [10]. Finally, we will discuss the BCJR-construction of trellises and recall some results from [8] pertaining to the relation between the product- and the BCJR-construction.

Due to the cyclic structure of the time axis \mathcal{I} , the following interval notation has proven to be very convenient. For $a, b \in \mathcal{I}$ we define $[a, b] \coloneqq \{a, a + 1, \ldots, b\}$ if $a \leq b$ and $[a, b] \coloneqq$ $\{a, a + 1, \ldots, n - 1, 0, 1, \ldots, b\}$ if a > b. Moreover, we set $(a, b] \coloneqq [a, b] \setminus \{a\}$. We call the intervals (a, b] and [a, b] conventional if $a \leq b$ and circular otherwise. Notice that $(a, a] = \emptyset$. It is easy to see that $\mathcal{I} \setminus (a, b] = (b, a]$ for all $a \neq b$. Hence the complement of a nonempty conventional interval is circular and vice versa. The following notion from [10, p. 2089] will be crucial for this paper.

Definition 2.1 For a vector $c = (c_0, \ldots, c_{n-1}) \in \mathbb{F}^n \setminus \{0\}$ we call any half-open interval (a, b] a span of c if $c_a \neq 0 \neq c_b$ and if the closed interval [a, b] contains the support of c.

Excluding the starting point a from the span does not seem to be intuitive but will be very convenient for our purposes.¹

Let (a, b] be a span of the nonzero vector $c = (c_0, \ldots, c_{n-1}) \in \mathbb{F}^n$. The elementary trellis for the pair (c, (a, b]) is defined as $T_{c,(a,b]} := (V, E)$ with the state spaces and transition spaces given by $V_j = \operatorname{im}(\mu_j) \subseteq \mathbb{F}$ and $E_j = \operatorname{im}(\mu_j, c_j, \mu_{j+1}) \subseteq V_j \times \mathbb{F} \times V_{j+1}$, respectively, where $\mu_j = 1$ for $j \in (a, b]$ and $\mu_j = 0$ otherwise. Thus $V_j = \mathbb{F}$ for $j \in (a, b]$ and $V_j = \{0\}$ otherwise. The

¹It would be more accurate to distinguish between a time axis for the symbols and a time axis for the states, see [5]. Then the span (a, b] of a vector is its active state interval (that is, it is the interval of nonzero states of its corresponding cycle in the elementary trellis), while its active symbol interval is given by [a, b].

trellis $T_{c,(a,b]}$ is linear, reduced, biproper, one-to-one, and represents the 1-dimensional code in \mathbb{F}^n generated by c. The trellis is conventional if and only if (a,b] is a conventional span. Obviously, the SCP and ECP are given by (s_0, \ldots, s_{n-1}) and (e_0, \ldots, e_{n-1}) , respectively, where $s_j = 1$ if $j \in (a,b]$ and $s_j = 0$ if $j \notin (a,b]$, while $e_j = 1$ if $j \in [a,b]$ and $e_j = 0$ if $j \notin [a,b]$.

For the following notion of product trellises recall the definition and basic properties of trellis products $T_1 \times T_2$; see, for instance, [10, p. 2089] and [8, Prop. III.4]).

Definition 2.2 Let $C = \operatorname{im} G$, where $G \in \mathbb{F}^{r \times n}$ has no zero rows. Denote the rows of G by $g_1, \ldots, g_r \in \mathbb{F}^n$ and let $S \coloneqq [(a_l, b_l], l = 1, \ldots, r]$ be a span list for G, that is, $(a_l, b_l]$ is a span (conventional or circular) for the row $g_l, l = 1, \ldots, r$. The product trellis $T_{G,S}$ is defined as the trellis $T_{g_1,(a_1,b_1]} \times \ldots \times T_{g_r,(a_r,b_r]}$. In other words, the state and transition spaces of $T_{G,S}$ are given by $V_j = \operatorname{im} M_j$ and $E_j = \operatorname{im} (M_j, G_j^\mathsf{T}, M_{j+1})$, where, as before, G_j^T denotes the *j*-th column of G and

$$M_j = \begin{pmatrix} \mu_j^1 & \\ & \ddots & \\ & & \mu_j^r \end{pmatrix} \in \mathbb{F}^{r \times r}, \text{ where } \mu_j^l = \begin{cases} 1, & \text{if } j \in (a_l, b_l], \\ 0, & \text{if } j \notin (a_l, b_l]. \end{cases}$$

Let us briefly comment on the subtle differences in the naming of such trellises in the literature. First of all, due to Definition 2.1, we only consider spans such that the vector is nonzero at the endpoints of that span (as opposed to general intervals containing the support of the vector). This way, we immediately exclude certain (but not all) mergeable trellises; see also [10, Lemma 4.3]. This makes our definition more restrictive than the one in [10]. Second, in the paper [15], product trellises of the type above (and where G has full row rank) are called KV-trellises, and the matrix G, along with its span list, is called a KV-product matrix. In the present paper, we will reserve the name KV-trellises for a particular type of product trellis that has been introduced by Koetter/Vardy [10]; see Definition 2.4.

The following properties of product trellises are easy to see. Later on we will only deal with product trellises where the starting points (resp., ending points) of the spans are distinct, and therefore we restrict ourselves to this case in (d) below. One could easily give the formulas for the more general case.

Proposition 2.3 Let the data be as in the previous definition and put $T \coloneqq T_{G,S}$. Then

- (a) T is a linear and reduced trellis.
- (b) T is one-to-one if and only if $\operatorname{rk} G = r$.
- (c) T is biproper if and only if a_1, \ldots, a_r are distinct and b_1, \ldots, b_r are distinct.
- (d) Let a_1, \ldots, a_r be distinct and b_1, \ldots, b_r be distinct. Then the SCP and ECP of T are given by (s_0, \ldots, s_{n-1}) and (e_0, \ldots, e_{n-1}) , respectively, where $s_j = |\{l = 1, \ldots, r \mid j \in (a_l, b_l]\}|$ and

$$e_{j} = \left\{ \begin{array}{cc} s_{j} + 1, & \text{if } j \in \{a_{1}, \dots, a_{r}\} \\ s_{j}, & \text{if } j \notin \{a_{1}, \dots, a_{r}\} \end{array} \right\} = \left\{ \begin{array}{cc} s_{j+1} + 1, & \text{if } j \in \{b_{1}, \dots, b_{r}\} \\ s_{j+1}, & \text{if } j \notin \{b_{1}, \dots, b_{r}\}. \end{array} \right\}$$

PROOF: (a), (b) as well as the formulas in (d) follow easily from the properties of elementary trellises as well as those of trellis products; see also [8, Prop. III.4].

(c) The implication " \Rightarrow " follows from the proof of [10, Cor. 4.5]. The converse has been shown in [8, Thm. III.6].

We now turn to a particular class of product trellises, introduced by Koetter/Vardy [10]. A main result of [10] – and a major breakthrough in the study of minimal tail-biting trellises – is the construction of a list of characteristic generators, collected in a characteristic matrix, from which all minimal trellises can be derived. In [10], this matrix is defined as the outcome of a particular procedure. We will follow the presentation in [8] and define the characteristic matrix in terms of its relevant properties. A justification of this approach versus the one in [10] has been given in [8, Sec. III]

Definition 2.4 Let $C \subseteq \mathbb{F}^n$ be as in (1.2). A characteristic pair of C is defined to be a pair (X, \mathcal{T}) , where

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{F}^{n \times n} \text{ and } \mathcal{T} = [(a_l, b_l], l = 1, \dots, n]$$
(2.1)

have the following properties.

- (i) im X = C, that is, $\{x_1, \ldots, x_n\}$ forms a generating set of C.
- (ii) $(a_l, b_l]$ is a span of x_l for l = 1, ..., n.
- (iii) a_1, \ldots, a_n are distinct and b_1, \ldots, b_n are distinct.
- (iv) For all $j \in \mathcal{I}$, there exist exactly n k row indices, l_1, \ldots, l_{n-k} , such that $j \in (a_{l_i}, b_{l_i}]$ for $i = 1, \ldots, n-k$.

We call X a characteristic matrix of \mathcal{C} and \mathcal{T} the characteristic span list. The rows of X are also called characteristic generators. A trellis $T_{G,S}$ is called a $KV_{(X,\mathcal{T})}$ -trellis of \mathcal{C} if $G \in \mathbb{F}^{k \times n}$ consists of k distinct linearly independent rows of X and S consists of the corresponding k spans in \mathcal{T} . A trellis is called a KV-trellis of \mathcal{C} if it is a $KV_{(X,\mathcal{T})}$ -trellis for some characteristic pair (X,\mathcal{T}) of \mathcal{C} .

The relevance of characteristic pairs and KV-trellises becomes apparent in part (b) below: the class of KV-trellises contains all minimal trellises. As a consequence, one may restrict the study of tail-biting trellises to the important class of KV-trellises. For the following results it is crucial that C has support \mathcal{I} .

- **Theorem 2.5** ([10, Sec. V], [8, Thm. III.15]) (a) The code C has a characteristic pair, and the characteristic span list is, up to ordering, uniquely determined by C.
- (b) For every minimal trellis T of C there exists a characteristic pair (X, \mathcal{T}) such that T is a $KV_{(X,\mathcal{T})}$ -trellis.

The following property shows that characteristic spans are "shortest spans" for each given starting point. This property not only makes it algorithmically easy to find characteristic generators for a given code, see Remark 2.7 below, but will also be crucial later on in order to derive strong properties for KV-trellises.

Lemma 2.6 Let $c \in C$ be a nonzero codeword with (conventional or circular) span (a, b]. Then $(a, \hat{b}] \subseteq (a, b]$, where $(a, \hat{b}]$ is the unique characteristic span starting at a.

PROOF: Notice that by Def. 2.4(iii), there does indeed exist a characteristic span (a, b] starting at a. Let us first consider the case where a = 0. By (iv) of the same definition there exist exactly k characteristic spans not containing 0, thus k conventional spans. One of these spans is $(0, \hat{b}]$. Definition 2.4(iii) tells us that generators with these spans form an MSGM of C in the sense of [14, Def. 6.2, Thm. 6.11]. As a consequence, it has minimal span length. This implies that $(0, \hat{b}] \subseteq (0, b]$, for otherwise we could replace the generator with span $(0, \hat{b}]$ by c and obtain a generator matrix with shorter span length.

If $a \neq 0$, we may use the cyclic left shift σ^a by a units in \mathbb{F}^n . From Definition 2.4 we obtain immediately that if $\mathcal{T} = [(a_1, b_1], \ldots, (a_n, b_n]]$ is the characteristic span list of \mathcal{C} with

generators x_1, \ldots, x_n , then $[(a_1 - a, b_1 - a], \ldots, (a_n - a, b_n - a]]$ is the characteristic span list of the code $\sigma^a(\mathcal{C})$ with generators $\sigma^a(x_1), \ldots, \sigma^a(x_n)$; see also Remark 2.11 further down. Since $\sigma^a(c) \in \sigma^a(\mathcal{C})$ has span (0, b - a], we may now make use of the first case. This leads to the desired result.

Remark 2.7 The above result allows us to set up a greedy algorithm for finding the characteristic spans of a given code (with support \mathcal{I}). Namely, for $a = 0, \ldots, n-1$, let $b_a \in \mathcal{I}$ be such that $(a, b_a]$ is the shortest span among all possible spans starting at time a of the (nonzero) codewords in \mathcal{C} . Then the resulting list $[(a, b_a], a = 0, \ldots, n-1]$ is the characteristic span list of \mathcal{C} . The same can be done by screening the spans by their ending points. Of course, the greedy algorithm will also produce, at the same time, a list of characteristic generators, that is, a characteristic matrix. This generalizes the "shortest basis approach" described by Forney in [5] for very general (conventional) realizations to the tail-biting case.

The following construction of tail-biting trellises has been introduced by Nori/Shankar in [15, Sec. III] and studied in detail in [8, Sec. IV]. We will generalize this slightly by also allowing generator matrices that do not have full rank. For (b) recall that row(D,l) denotes the *l*-th row of the matrix D.

Definition 2.8 Let the code C and the matrices $G \in \mathbb{F}^{r \times n}$, $H \in \mathbb{F}^{(n-k) \times n}$ be as in (1.2) and (1.3).

(a) Let $D \in \mathbb{F}^{r \times (n-k)}$ be any matrix. For $i \in \mathcal{I}$ define the matrices

$$N_0 = D$$
 and $N_i = N_{i-1} + G_{i-1}^{\mathsf{T}} H_{i-1}$ for $i > 0.$ (2.2)

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 $E_i = \operatorname{im} (N_i, G_i^{\mathsf{T}}, N_{i+1})$. It is easy to see that $T_{(G,H,D)}$ is a linear, reduced, and biproper trellis representing the code \mathcal{C} . We call D the displacement matrix for the trellis $T_{(G,H,D)}$.

(b) Let $S := [(a_l, b_l], l = 1, ..., r]$ be a span list of G. Then the trellis $T_{(G,H,S)}$ is defined as $T_{(G,H,D)}$, where

$$D \in \mathbb{F}^{r \times (n-k)} \text{ is such that } \operatorname{row}(D,l) = \sum_{j=a_l}^{n-1} g_{lj} H_j \text{ for } j = 1, \dots, r.$$
(2.3)

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

Notice that if D = 0, then the trellis $T_{(G,H,D)}$ is conventional. If in addition, $r = k = \operatorname{rk} G$, then $T_{(G,H,D)}$ is, in fact, the classical conventional BCJR-trellis of \mathcal{C} , hence minimal, see [14, Sec. IV] and [2]. In general, the rows of the displacement matrix D may be interpreted as the (circular) past of the generators in G. As always in state space realizations, this past is captured in the state at time 0 through which the associated trajectory (cycle) passes. Consequently, the *l*-th rows of the matrices N_j form the sequence of states through which the cycle induced by the *l*-th generator passes. In this sense, the BCJR-trellis is, just like the product trellis, based on r individual generators with spans. The state spaces arise as the sum of the state spaces of each generator (while for product trellises they are the direct sum of the state spaces). As a consequence, BCJR-trellises are not always one-to-one, even if the generator matrix G has full row rank.

The following results have been proven in [8] for the case where $G \in \mathbb{F}^{k \times n}$ has rank k. One can easily verify that the same proofs apply to $G \in \mathbb{F}^{r \times n}$ with $r > k = \operatorname{rk} G$. For the notion of mergeability, we refer to [10, Sec. II.B] or [8, Sec. II].

Theorem 2.9 ([8, Cor. IV.7, Thm. IV.9, Rem. IV.13, Thm. IV.11]) Let G, H, and the span list S be as in Definition 2.8.

- (a) The BCJR-trellis $T_{(G,H,S)}$ is non-mergeable.
- (b) If the product trellis $T_{G,S}$ is non-mergeable, then $T_{G,S}$ is isomorphic to $T_{(G,H,S)}$.
- (c) If the product trellis $T_{G,S}$ has the same SCP as the BCJR-trellis $T_{(G,H,S)}$, then these two trellises are isomorphic. In this case, the starting points of the spans in S are distinct and so are the ending points.

It is worth noting that every product trellis $T_{G,S}$ can be merged to the corresponding BCJRtrellis $T_{(G,H,S)}$, see [8, Thm. IV.9].

Example 2.10 (a) This example appeared already in [8, Ex. IV.12]. It will be revisited later again when discussing dualization techniques. Consider the code $\mathcal{C} = \operatorname{im} G = \ker H^{\mathsf{T}} \subseteq \mathbb{F}_2^5$, where

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

Then S := [(1,3], (3,0], (2,1]] is a span list for G. Making use of the greedy algorithm in Remark 2.7 and checking all 8 codewords, it is easy to see that (1,3] and (2,1] are characteristic spans of C, whereas the codeword $(1,0,0,0,1) \in C$ with span (4,0] shows that (3,0] is not a characteristic span. The BCJR-trellis $T := T_{(G,H,S)}$ has state and transition spaces given by $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im} (N_j, G_j^{\mathsf{T}}, N_{j+1})$, where N_j are given in the matrix

In this matrix we stagger N_j with the columns of G in order to easily read off the transition spaces E_j . The trellis T is shown in the figure below (here and in all other trellises we will denote edges with label one as solid lines and those with zero label as dashed lines). It is straightforward to see that the corresponding product trellis $T_{G,S}$ has the same SCP as T and therefore is isomorphic to T according to Theorem 2.9. As a consequence, T is one-to-one, which we can see also directly from the trellis. Notice that by the above, T is not a KV-trellis.



[Trellis T]

(b) Consider the self-dual code $\mathcal{C} \subseteq \mathbb{F}_2^4$ with $\mathcal{C} = \operatorname{im} G = \ker G^{\mathsf{T}}$, where

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

A characteristic pair is given by

$$X = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \ \mathcal{T} = [(3,0], (2,1], (1,2], (0,3]].$$
(2.5)

This can easily be seen by applying Remark 2.7 to the three nonzero codewords in C. Let us consider the BCJR-trellis $T_{(X,G,\mathcal{T})}$. The matrix

lists the state space matrices N_j staggered with the columns of X. We will not display this trellis as it is not a very useful trellis by itself. For instance, it is not one-to-one. However, we can easily read off the state space matrices and transition spaces for all $KV_{(X,\mathcal{T})}$ -trellises by simply taking the submatrix of S consisting of any two rows for which the corresponding two rows in X are linearly independent. This results in five $KV_{(X,\mathcal{T})}$ -trellises. It is an immediate consequence of Proposition 2.3(d), that no two of these trellises have the same SCP and ECP. In particular, they are pairwise nonisomorphic. The trellises $T_{(X,G,\mathcal{T})}$, based on an entire characteristic matrix, will be used in Section 4 as the starting point of the dualization procedure.

(c) Let $\mathcal{C} = \operatorname{im} G = \operatorname{ker} H^{\mathsf{T}} \subseteq \mathbb{F}_3^4$, where

$$G = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \ H = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

The pair (X, \mathcal{T}) , where

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

is a characteristic pair of C. The matrix X is normalized, that is, each characteristic generator has coordinate 1 at the starting point of its span. It can easily be seen that C has 9 different normalized characteristic matrices with span list T.

The product construction as well as the BCJR-construction behave nicely under the cyclic shift. This is described in the following remark, of which we will make frequent use. Recall that we compute with indices modulo n, which, of course, also applies to the span lists.

Remark 2.11 Denote the rows of G by $g_1, \ldots, g_r \in \mathbb{F}^n$ and let the span list S be as in Definition 2.8(b). Let σ denote the cyclic left shift on \mathbb{F}^n and let $G^* \in \mathbb{F}^{r \times n}$ be the matrix consisting of the shifted rows $\sigma(g_l)$, $l = 1, \ldots, r$. Then $S^* = [(a_l - 1, b_l - 1], l = 1, \ldots, r]$ forms a span list for G^* . The state and transition spaces of the product trellis T_{G^*,S^*} are given by $V_i^* = \operatorname{im} M_{i+1}$ and $E_i^* = \operatorname{im} (M_{i+1}, G_{i+1}^{\mathsf{T}}, M_{i+2})$, where M_i is as in Definition 2.2. Similarly, if

 (X, \mathcal{T}) as in (2.1) is a characteristic pair for \mathcal{C} , then (X^*, \mathcal{T}^*) is a characteristic pair of $\sigma(\mathcal{C})$, where

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

Finally, if D is as in (2.3) and N_i are the state space matrices of $T_{(G,H,S)}$ as in (2.2), then the state space matrices for the BCJR-trellis $T_{(G^*,H^*,S^*)}$ are given by $N_i^* = N_{i+1}$ for $i \in \mathcal{I}$.

In this paper, we will mainly consider BCJR-trellises based on characteristic generators. The following theorem will be crucial later.

Theorem 2.12 Let (X, \mathcal{T}) be a characteristic pair of \mathcal{C} . Let $G \in \mathbb{F}^{r \times n}$ be a selection of r rows of X and $\mathcal{S} := [(a_l, b_l], l = 1, ..., r]$ be the corresponding selection of characteristic spans. Thus, $(a_l, b_l]$ is the span of the *l*-th row of G. Consider the BCJR-trellis $T_{(G,H,\mathcal{S})}$, and let $N_j \in \mathbb{F}^{r \times (n-k)}, j \in \mathcal{I}$, be its state space matrices. Then for every $j \in \mathcal{I}$

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

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

As a consequence, $T_{(G,H,S)}$ is isomorphic to the corresponding product trellis $T_{G,S}$. Therefore it makes sense to call $T_{(G,H,S)}$ a KV-trellis in the case where $r = k = \operatorname{rk} G$.

PROOF: The statement in (1) has been proven in [8, Prop. IV.6] for BCJR-trellises $T_{(G,H,S)}$, where $G \in \mathbb{F}^{k \times n}$ has rank k. It can easily be seen that the same proof applies to the general case for G. As for the second statement, first consider j = 0. Then $j \in (a_l, b_l]$ is equivalent to $(a_l, b_l]$ being a circular span. In the proof of [8, Thm. IV.11] it has been shown that the set $\{ \operatorname{row}(N_0, l) \mid (a_l, b_l] \text{ circular} \}$ is linearly independent (these rows appear in the matrix $Z \in \mathbb{F}^{(n-k)\times(n-k)}$ in [8, Eq. (IV.5)], which is non-singular due to the same proof in [8]). This proves (2) for j = 0. Applying a cyclic left shift by j steps (see Remark 2.11), we obtain the desired result for arbitrary N_j , and this completes the proof of (2).

Finally, the previous results tell us that $\operatorname{rk} N_j = |\{l \mid j \in (a_l, b_l]\}|$. Furthermore, due to Definition 2.4, the starting points of distinct characteristic spans are distinct, and the same is true for the ending points. Therefore we may apply Proposition 2.3(d) and conclude that $T_{(G,H,S)}$ and $T_{G,S}$ have the same SCP. Thus, these trellises are isomorphic due to Theorem 2.9(c). \Box

We close this section with the following technical results pertaining to the BCJR-presentation of KV-trellises, which will be needed later on.

Lemma 2.13 Let $T = T_{(G,H,S)}$ be a KV-trellis of C, that is, $G = (g_{lj}) \in \mathbb{F}^{k \times n}$ has rank k and its span list $S = [(a_l, b_l] | l = 1, ..., k]$ consists of characteristic spans of C. Let $N_j, j \in \mathcal{I}$, be the state space matrices of T. Then

(i) If $j = b_l$ for some l, then $H_j = -g_{lj}^{-1} \operatorname{row}(N_j, l)$

(ii) If $j \notin \{b_1, \ldots, b_k\}$, then $H_j \notin \operatorname{im} N_j$.

As the proof will show, part (i) is true for general BCJR-trellises, and only (ii) needs the particular properties of characteristic spans.

PROOF: Using a cyclic shift, see Remark 2.11, we may assume without loss of generality that j = 0.

(i) Let $0 = b_l$. Then (2.3) implies row $(N_0, l) = \sum_{i=a_l}^{n-1} g_{li}H_i = -g_{l0}H_0$, where the last identity follows from the identity $GH^{\mathsf{T}} = 0$ along with the fact that $(a_l, b_l] = (a_l, 0]$ is the span of the *l*-th row $(g_{l0}, \ldots, g_{l,n-1})$ of G. This establishes (i).

(ii) Let $0 \notin \{b_1, \ldots, b_k\}$ and assume $H_0 = \beta N_0$ for some $\beta \in \mathbb{F}^k$. Using the definition of N_0

in (2.3), this becomes $H_0 = \sum_{l=1}^k \beta_l \sum_{j=a_l}^{n-1} g_{lj} H_j$. Notice that $H_0 \neq 0$, due to our general assumption that the dual code \mathcal{C}^{\perp} has support \mathcal{I} . Therefore with the aid of Theorem 2.12(1), the above may be written as $H_0 = \sum_{l \in \mathcal{L}} \beta_l \sum_{j=a_l}^{n-1} g_{lj} H_j$, where $\mathcal{L} \coloneqq \{l \mid 0 \in (a_l, b_l], \beta_l \neq 0\}$ and $\mathcal{L} \neq \emptyset$. Let $s \in \mathcal{L}$ be such that $a_s = \min\{a_l \mid l \in \mathcal{L}\}$. Then $a_s > 0$ because the condition $0 \in (a_l, b_l]$ implies that the span is circular. Define the vectors

$$\hat{g}_{l} = (\hat{g}_{l0}, \dots, \hat{g}_{l,n-1}), \text{ where } \hat{g}_{lj} = \begin{cases} g_{lj}, & \text{if } j \ge a_{l} \\ 0, & \text{if } j < a_{l}. \end{cases}$$

Then we obtain $H_0 = \sum_{l \in \mathcal{L}} \beta_l \hat{g}_l H^{\mathsf{T}}$. As a consequence, $c \coloneqq \sum_{l \in \mathcal{L}} \beta_l \hat{g}_l - e_0 \in \ker H^{\mathsf{T}} = \mathcal{C}$, where $e_0 \in \mathbb{F}^n$ is the first standard basis vector. Now the definition of \hat{g}_l shows that the codeword c has span $(a_s, 0]$, and thus Lemma 2.6 implies $(a_s, b_s] \subseteq (a_s, 0]$. Since $(a_s, b_s]$ is circular, this in turn yields $b_s = 0$, contradicting our assumption that $0 \notin \{b_1, \ldots, b_k\}$. This proves (ii). \Box

Theorem 2.14 Let $T_{(G,H,S)}$ and N_j be as in Lemma 2.13. Then $\bigcap_{j=0}^{n-1} \text{ im } N_j = \{0\}$.

PROOF: Let $w \in \bigcap_{j=0}^{n-1} \text{ im } N_j$, say $w = \alpha_j N_j$ for some $\alpha_j = (\alpha_{j,1}, \ldots, \alpha_{j,k}) \in \mathbb{F}^k$, $j \in \mathcal{I}$. We will first show that there exists some common $\alpha \in \mathbb{F}^k$ such that $w = \alpha N_j$ for all $j \in \mathcal{I}$.

Note that $0 = \alpha_{j+1}N_{j+1} - \alpha_j N_j = \alpha_{j+1}(N_j + G_j^\mathsf{T}H_j) - \alpha_j N_j$, hence $\alpha_{j+1}G_j^\mathsf{T}H_j = (\alpha_j - \alpha_{j+1})N_j$. Observe that $\alpha_{j+1}G_i^\mathsf{T} \in \mathbb{F}$, thus a scalar.

i) If $\alpha_{j+1}G_j^{\mathsf{T}} = 0$, then $\alpha_j - \alpha_{j+1} \in \ker N_j$, and Theorem 2.12 implies $\alpha_{j,l} = \alpha_{j+1,l}$ for all l such that $j \in (a_l, b_l]$.

ii) If $\alpha_{j+1}G_j^{\mathsf{T}} \neq 0$, then the above yields $H_j = (\alpha_{j+1}G_j^{\mathsf{T}})^{-1}(\alpha_j - \alpha_{j+1})N_j$. In this case Lemma 2.13 implies that $j = b_m$ for some $m \in \{1, \ldots, k\}$. Furthermore, part (i) of that Lemma along with Theorem 2.12(1) and (2) shows that $\alpha_{j,l} = \alpha_{j+1,l}$ for all $l \neq m$ such that $j \in (a_l, b_l]$.

But then i) and ii) together yield

$$\alpha_{j,l} = \alpha_{b_l,l} \text{ for all } l \text{ such that } j \in (a_l, b_l].$$
(2.6)

Define now $\alpha := (\alpha_{b_1,1}, \ldots, \alpha_{b_k,k}) \in \mathbb{F}^k$. Using Theorem 2.12(1) once more, along with (2.6), it is straightforward to see that $\alpha N_j = \alpha_j N_j = w$ for all $j \in \mathcal{I}$.

Now we have $0 = w - w = \alpha(N_{j+1} - N_j) = \alpha G_j^{\mathsf{T}} H_j$ for all $j \in \mathcal{I}$. Since $H_j \neq 0$ for all $j \in \mathcal{I}$, this leads to $\alpha G_j^{\mathsf{T}} = 0$ for all $j \in \mathcal{I}$ and hence $\alpha G = 0$. As a consequence, the full row rank of G implies $\alpha = 0$, and we arrived at w = 0, as desired.

It is worth noting that the last result is not true for general BCJR-trellises $T_{(G,H,S)}$, even if $G \in \mathbb{F}^{k \times n}$ has rank k and the trellis is isomorphic to the corresponding product trellis. An example is given in Example 2.10(a). The trellis T displayed in that graph satisfies $T_{(G,H,S)} \cong T_{G,S}$ and $G \in \mathbb{F}_2^{3 \times 5}$ has rank 3, but $(0,1) \in \bigcap_{j=0}^4 \operatorname{in} N_j$, where N_j are as in (2.4).

3 Dualizing Trellises

In this section we will investigate two methods of dualizing a given trellis in order to obtain a trellis representing the dual code. One method is what we call local dualization, and it amounts to taking duals of the transition spaces along with a sign inverter. This very elegant and profound method has been introduced by Forney in [4] and applies to all linear (even group) realizations; see also the presentation based on graphical models in [6]. Theorem 3.1 below is simply a special case of it. The second method comes naturally with the BCJRconstruction and has been introduced by Nori/Shankar in [15]. We will show that, in general, the BCJR-dual is a proper subtrellis of the local dual and that for KV-trellises the two coincide. There is yet another notion of trellis duality, introduced by Koetter/Vardy in [10, Ch. VII], based on what they call the intersection product. As it turns out via straightforward computation, this notion is identical to the local dual.

We begin with the local dualization, which in our particular case of tail-biting trellises looks as follows.

Theorem 3.1 Let T = (V, E) be a linear trellis representing the code $C \subseteq \mathbb{F}^n$. 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

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

that is,

$$(E_j)^{\circ:=} \{ (\hat{v}, b, \hat{w}) \in \hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1} | \langle v, \hat{v} \rangle + ab - \langle w, \hat{w} \rangle = 0 \text{ for all } (v, a, w) \in E_j \}.$$
(3.1)

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, $SCP(T^{\circ}) = SCP(T) \coloneqq (s_0, \ldots, s_{n-1})$ and

$$\dim(E_j)^\circ = s_j + s_{j+1} + 1 - \dim E_j \text{ for } j \in \mathcal{I}.$$
(3.2)

We call T° the local dual of T.

One should notice that, via the non-degenerate bilinear form, the space \hat{V}_j is isomorphic to the linear algebra dual of V_j consisting of all linear functionals on V_j . Conversely, the linear algebra dual naturally gives rise to a non-degenerate bilinear form and therefore may serve as dual state space \hat{V}_j . Later on we will make specific choices for \hat{V}_j and the bilinear form, justifying our setting in Theorem 3.1.

It should be noted that the isomorphism class of the trellis T° does not depend on the choice of the spaces \hat{V}_i and the non-degenerate bilinear forms. Indeed, if $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are two such forms on $V_j \times \hat{V}_j$ and $V_j \times \hat{V}_j$, 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.

The theorem is a special case of the local dualization procedure for normal graphs derived in [4, Sec. VII]. However, we think it is worth reproducing Forney's proof for this special case.

PROOF: It is clear that the label code $\mathcal{S}(T^{\circ})$, see (1.1), as well as the transition spaces $(E_j)^{\circ}$ are linear spaces. Hence T° is a linear trellis. Moreover, the bilinear form on $(V_j \times \mathbb{F} \times V_{j+1}) \times (\hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1})$ is non-degenerate, and therefore $\dim(E_j)^{\circ} = \dim(\hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1}) - \dim E_j$, proving (3.2).

It remains to show that T° represents \mathcal{C}^{\perp} . In order to do so, define

$$\mathcal{V} := V_0 \times V_1 \times V_1 \times V_2 \times V_2 \times \ldots \times V_{n-1} \times V_{n-1} \times V_0,$$
$$\hat{\mathcal{V}} := \hat{V}_0 \times \hat{V}_1 \times \hat{V}_1 \times \hat{V}_2 \times \hat{V}_2 \times \ldots \times \hat{V}_{n-1} \times \hat{V}_{n-1} \times \hat{V}_0,$$

(that is, we are replicating each state space). The given bilinear forms on each $V_j \times \hat{V}_j$ naturally give rise to a non-degenerate bilinear form $\langle \cdot, \cdot \rangle$ on $\mathcal{V} \times \hat{\mathcal{V}}$ via

$$\langle (\tilde{v}_0, v_1, \tilde{v}_1, \dots, v_{n-1}, \tilde{v}_{n-1}, v_0), (\tilde{w}_0, w_1, \tilde{w}_1, \dots, w_{n-1}, \tilde{w}_{n-1}, w_0) \rangle = \sum_{j=0}^{n-1} \langle \tilde{v}_j, \tilde{w}_j \rangle + \sum_{j=0}^{n-1} \langle v_j, w_j \rangle.$$

This further extends to a non-degenerate bilinear form

$$(\mathcal{V} \times \mathbb{F}^n) \times (\hat{\mathcal{V}} \times \mathbb{F}^n) \longrightarrow \mathbb{F}, \quad \langle (v, a), (w, b) \rangle = \langle v, w \rangle + ab^{\mathsf{T}}.$$
(3.3)

Recall that the transition spaces E_j are contained in $V_j \times \mathbb{F} \times V_{j+1}$. Hence, the direct product $E_0 \times \ldots \times E_{n-1}$ is in $\mathcal{V} \times \mathbb{F}^n$, if we sort the state and edge labels accordingly. Denoting this obvious permutation by ρ , we therefore have

$$\mathcal{P} \coloneqq \rho(E_0 \times E_1 \times \ldots \times E_{n-1}) \subseteq \mathcal{V} \times \mathbb{F}^n$$

Now we use the replication space

$$\mathcal{R} \coloneqq \{ (v_0, v_1, v_1, v_2, v_2, \dots, v_{n-1}, v_{n-1}, v_0) \mid v_j \in V_j \} \subseteq \mathcal{V}$$

in order to relate edges in the various transition spaces E_j with each other by checking whether ending and starting points coincide. Indeed, the label code $\mathcal{S}(T)$ defined in (1.1) is given by the conditioned space

$$(\mathcal{P} \mid \mathcal{R}) \coloneqq \{(v, a) \in \mathcal{P} \mid v \in \mathcal{R}\}$$

(to be precise, we also have to eliminate for $v \in \mathcal{R}$ one copy of each state to obtain $\mathcal{S}(T)$). As a consequence, the code \mathcal{C} is the projection

$$\mathcal{C} = \pi(\mathcal{P} \mid \mathcal{R}) = \{ a \in \mathbb{F}^n \mid \exists v \in \mathcal{V} \colon (v, a) \in (\mathcal{P} \mid \mathcal{R}) \},\$$

where π denotes the natural projection of $\mathcal{V} \times \mathbb{F}^n$ onto \mathbb{F}^n (this is the identity $\mathcal{C} = (\mathcal{P} | \mathcal{R})|_{I_{\mathcal{A}}}$ in [4, p. 540]). Now we may take duals with respect to $\langle \cdot, \cdot \rangle$. Then Forney's Theorem on Conditioned Code Duality [4, Thm. 7.2] (or straightforward verification) tells us that $\mathcal{C}^{\perp} = [\pi(\mathcal{P} | \mathcal{R})]^{\perp} = \pi(\mathcal{P}^{\perp} | \mathcal{R}^{\perp})$. Here \cdot^{\perp} refers to the bilinear form in (3.3) or a restriction of it to the appropriate subspaces. It remains to clarify the relation between the latter space and the trellis T° . From the very definition of \mathcal{P} and \mathcal{R} and the bilinear forms involved we obtain

$$\mathcal{P}^{\perp} = \rho(E_0^{\perp} \times \ldots \times E_{n-1}^{\perp}),$$

$$\mathcal{R}^{\perp} = \{(w_0, -w_1, w_1, -w_2, w_2, \ldots, -w_{n-1}, w_{n-1}, -w_0) \mid w_j \in \hat{V}_j\},$$

where $E_j^{\perp} = \{(w, b, \tilde{w}) \in \hat{V}_j \times \mathbb{F} \times \hat{V}_{j+1} \mid \langle v, w \rangle + ab + \langle \tilde{v}, \tilde{w} \rangle = 0 \text{ for all } (v, a, \tilde{v}) \in E_j\}$. Observing that $(w_j, b_j, -w_{j+1}) \in E_j^{\perp} \iff (w_j, b_j, w_{j+1}) \in (E_j)^\circ$, we see that

 $\pi(\mathcal{P}^{\perp} \mid \mathcal{R}^{\perp}) = \{ b \in \mathbb{F}^n \mid \text{there exists a cycle in } T^{\circ} \text{ with edge-label sequence } b \}.$

This proves that T° represents \mathcal{C}^{\perp} , as desired.

The following two examples illustrate that in specific cases local dualization may lead to undesirable trellises. While in the first example this is not surprising because the primary trellis is not even proper, the second example is more unexpected. It shows a BCJR-trellis that is isomorphic to the corresponding product trellis and thus non-mergeable (hence biproper) and one-to-one, and yet the local dual is not reduced. In Theorem 3.5 below we will see that this does not happen for KV-trellises. It will be shown that the local dual of a KV-trellis is isomorphic to the BCJR-dual and thus reduced. Furthermore, in Section 4 we will show that this dual is even a KV-trellis of the dual code. **Example 3.2** (1) Consider the 2-dimensional code

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

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

see Definition 2.2 for the state space matrices M_j of product trellises. According to Theorem 3.1, the local dual T° has SCP (0,1,2) and ECP (1,2,1). In order to compute T° , we observe that the standard bilinear form on \mathbb{F}_2^2 induces a non-degenerate form on each $V_j = \operatorname{im} M_j$, and thus may be used for the computation of the dual transition spaces $(E_j)^{\circ}$. In particular, we will use V_j for the dual state spaces as well. Then we compute

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

. .

This leads to the following trellis T°



Obviously, not every vertex appears in a cycle and thus the trellis T° is not reduced. As a consequence, T° is not a product trellis in the sense of Definition 2.2.

(2) Consider Example 2.10(a). The BCJR-trellis $T \coloneqq T_{(G,H,S)}$ given in that example has state and transition spaces $V_j = \operatorname{im} N_j$ and $E_j = \operatorname{im} (N_j, G_j^{\mathsf{T}}, N_{j+1})$, with all matrices displayed in the matrix (2.4). In order to compute the local dual T° , we may again use the standard bilinear form on \mathbb{F}_2^2 and thus let V_j be the dual state space as well. Then

$$(E_0)^{\circ} = \operatorname{im} \begin{pmatrix} 1 & 0 & | & 1 & | & 0 & 0 \\ 0 & 1 & | & 0 & | & 0 & 1 \end{pmatrix}, \qquad (E_3)^{\circ} = \operatorname{im} \begin{pmatrix} 1 & 0 & | & 1 & | & 1 & 0 \\ 0 & 1 & | & 0 & | & 0 & 1 \end{pmatrix}, (E_1)^{\circ} = \operatorname{im} \begin{pmatrix} 0 & 1 & | & 1 & | & 0 & 1 \end{pmatrix}, \qquad (E_4)^{\circ} = \operatorname{im} \begin{pmatrix} 1 & 0 & | & 0 & | & 1 & 1 \\ 0 & 1 & | & 0 & | & 0 & 1 \\ 0 & 1 & | & 1 & | & 0 & 0 \end{pmatrix}.$$

This leads to the first trellis in the following figure.



The trellis T° is not 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 an isomorphic copy of the trellis on the right-hand side, which still represents C^{\perp} . In the next result we will discuss the dualization leading to T^{\perp} . Isomorphic versions of the trellises T° and T^{\perp} appeared already in [8, Ex. IV.12, Rem. V.4].

We now continue with a very simple and natural way of dualizing the trellises defined in Definition 2.8(a). This has been introduced by Nori/Shankar in [15].

Proposition 3.3 ([15, Def. 11]) Let $T = T_{(G,H,D)}$ be as in Definition 2.8(a) and suppose T represents the code C. Then the trellis $T_{(H,G,D^{\mathsf{T}})}$ represents the dual code C^{\perp} . We call $T_{(H,G,D^{\mathsf{T}})}$ the BCJR-dual of the trellis T, denoted by T^{\perp} .

One should bear in mind that even if the trellis T is a BCJR-trellis in the sense of Definition 2.8(b), that is, its displacement matrix is based on a span list, then T^{\perp} is not necessarily a BCJR-trellis in that sense, but only a trellis of the type defined in part (a) of that definition. This can be seen from Example 3.2(2) above. The trellis T given therein is a BCJR-trellis, and one can easily check that the displayed trellis T^{\perp} is indeed its BCJR-dual. Obviously the trellis T^{\perp} is mergeable (merging the states (000) and (011) in V_0 does not create any new cycles), and thus T^{\perp} is not a BCJR-trellis due to Theorem 2.9(a).

For trellises of the form $T = T_{(G,H,D)}$, we now have two ways of dualizing them, both of which result in trellises representing the dual code. By construction, the dual trellises T° and T^{\perp} have the same SCP. In general, however, these trellises are not isomorphic, as we have seen already in Example 3.2(2) above. Next we will show that, just like in the above example, T^{\perp} is a subtrellis of T° .

Proposition 3.4 Let $T = T_{(G,H,D)}$ be as in Definition 2.8(a). 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 $E_j = \operatorname{im} (N_j, G_j^{\mathsf{T}}, N_{j+1})$ be the state spaces and transition spaces of T, where the matrices N_j are defined as in (2.2). By the very definition of the BCJRdual, the state spaces of T^{\perp} are given by $\hat{V}_j = \operatorname{im} \hat{N}_j$, where $\hat{N}_j = N_j^{\mathsf{T}}$. Notice that the bilinear form $V_j \times \hat{V}_j \longrightarrow \mathbb{F}$, defined as $\langle \alpha N_j, \beta \hat{N}_j \rangle \coloneqq \alpha N_j \beta^{\mathsf{T}}$, 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

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

Now we see that $\hat{E}_j = \operatorname{im}(\hat{N}_j, H_j^{\mathsf{T}}, \hat{N}_{j+1}) \subseteq (E_j)^\circ$ since for all $\beta(\hat{N}_j, H_j^{\mathsf{T}}, \hat{N}_{j+1}) \in \hat{E}_j$ and $\alpha \in \mathbb{F}^k$ we have $\alpha N_j \beta^{\mathsf{T}} + \alpha G_j^{\mathsf{T}} H_j \beta^{\mathsf{T}} - \alpha N_{j+1} \beta^{\mathsf{T}} = \alpha (N_j + G_j^{\mathsf{T}} H_j - N_{j+1}) \beta^{\mathsf{T}} = 0$, due to (2.2). \Box

Now we are ready to show the main result of this section.

Theorem 3.5 If T is a KV-trellis of \mathcal{C} , then T^{\perp} is isomorphic to T° .

One should keep in mind that at this point it is not clear whether T^{\perp} , hence T° , is a KV-trellis of the dual code. This is indeed the case, as we will prove in Section 4.

PROOF: By Theorem 2.12 we may represent T as a BCJR-trellis $T = T_{(G,H,S)}$, where $G \in \mathbb{F}^{k \times n}$ has rank k and $S = [(a_l, b_l], l = 1, ..., k]$ is a span list of G consisting of characteristic spans of C. Put $\mathcal{A} = \{a_1, \ldots, a_k\}$ and $\mathcal{B} = \{b_1, \ldots, b_k\}$. Let $N_0 = D$ and N_j be as in (2.3), (2.2). By Proposition 3.3, the trellis T^{\perp} represents C^{\perp} and has state and transition spaces $\hat{V}_j := \operatorname{im} \hat{N}_j$ and $\hat{E}_j := \operatorname{im} (\hat{N}_j, H_j^{\mathsf{T}}, \hat{N}_{j+1})$, where $\hat{N}_j = N_j^{\mathsf{T}}$. In light of Proposition 3.4 it suffices to show that $\dim(E_j)^\circ = \dim \hat{E}_j$ for all $j \in \mathcal{I}$.

Denote the SCP and ECP of T by (s_0, \ldots, s_{n-1}) and (e_0, \ldots, e_{n-1}) , respectively, and let $e_j^{\circ} = \dim(E_j)^{\circ}$. Since T is isomorphic to the corresponding product trellis $T_{G,S}$, the formulas in Proposition 2.3(d) apply. Thus we have $e_j = s_{j+1}$ if $j \notin \mathcal{B}$ and $e_j = s_{j+1} + 1$ if $j \in \mathcal{B}$. Using (3.2) this leads to

$$e_j^\circ = s_j + 1 \text{ if } j \notin \mathcal{B} \text{ and } e_j^\circ = s_j \text{ if } j \in \mathcal{B}.$$
 (3.4)

On the other hand, the recursion in (2.2) implies that

$$\hat{e}_j \coloneqq \dim(\hat{E}_j) = \operatorname{rk}(\hat{N}_j, H_j^{\mathsf{T}}, \hat{N}_{j+1}) = \operatorname{rk}(\hat{N}_j, H_j^{\mathsf{T}}).$$
(3.5)

Using that $\hat{N}_j = N_j^{\mathsf{T}}$ has rank s_j , we conclude that $\hat{e}_j = s_j$ iff $H_j \in \operatorname{im} N_j$ and $\hat{e}_j = s_j + 1$ otherwise. But then Lemma 2.13 shows that $\hat{e}_j = s_j$ iff $j \in \mathcal{B}$ and $\hat{e}_j = s_j + 1$ iff $j \notin \mathcal{B}$. A comparison with (3.4) establishes $\dim(E_j)^\circ = \dim \hat{E}_j$ for all $j \in \mathcal{I}$. So the trellises are isomorphic.

The following example shows that $T^{\perp} \cong T^{\circ}$ may be true even if T is not a KV-trellis.

Example 3.6 Let $\mathcal{C} = \operatorname{im} G = \ker H^{\mathsf{T}} \subseteq \mathbb{F}_2^6$, 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 S = [(1,5], (2,4], (3,1]] for G. One can easily verify that (1,5] and (2,4] are characteristic spans of C, but (3,1] is not (there exists a codeword with span (3,0]). Hence the product trellis $T_{G,S}$ is not a KV-trellis. By straightforwardly computing the data for $T_{G,S}$ and the corresponding BCJR-trellis $T := T_{(G,H,S)}$, one obtains that both trellises have SCP (1,1,1,2,3,2). Hence they are isomorphic due to Theorem 2.9(c). Their ECP is (1,2,2,3,3,2). The displacement matrix of $T_{(G,H,S)}$ 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 $T_{(H,G,\hat{S})}$ of \mathcal{C}^{\perp} , where $\hat{S} = [(2,4], (3,0], (0,5]]$ is the chosen span list for the rows of H. Its displacement matrix turns out to be N_0^{T} . As a consequence, $T^{\perp} = T_{(G,H,S)}^{\perp} = T_{(G,H,N_0)}^{\perp} = T_{(H,G,N_0^{\mathsf{T}})} = T_{(H,G,\hat{S})}$. One can also easily verify that $T_{(H,G,\hat{S})} \cong T_{H,\hat{S}}$. Now it is easy to check that both T^{\perp} and T° have ECP (2,1,2,3,3,2). Thus Proposition 3.4 yields $T^{\perp} \cong T^{\circ}$.

4 A Duality for Characteristic Matrices

In this section we will restrict ourselves to KV-trellises. Recall from Theorem 3.5 that for these trellises the local dual and the BCJR-dual coincide, and hence we may use these dualizations interchangeably. In [10, p. 2097], Koetter/Vardy formulated the conjecture that if (X, \mathcal{T}) and $(Y, \hat{\mathcal{T}})$ are characteristic pairs of \mathcal{C} and \mathcal{C}^{\perp} , respectively, then for every a $\mathrm{KV}_{(X,\mathcal{T})}$ -trellis there is a $KV_{(Y,\hat{T})}$ -trellis with the same state complexity profile. They actually restricted the characteristic matrices to a specific choice, the lexicographically first ones, and also made the correspondence more precise. In [8, Ex. III.13, Ex. V.2], however, it has been shown that the conjecture is not true in this generality, and in particular not for the lexicographically first characteristic matrices of \mathcal{C} and \mathcal{C}^{\perp} . In this section we will prove the following reformulated, but much stronger, version of the Koetter/Vardy conjecture: for each characteristic pair (X,\mathcal{T}) of \mathcal{C} there exists a characteristic pair $(Y,\hat{\mathcal{T}})$ of \mathcal{C}^{\perp} such that the dual of a $\mathrm{KV}_{(X,\mathcal{T})}$ trellis is a $KV_{(Y,\hat{T})}$ -trellis. Since, due to Theorem 3.1, the local dual and the primary trellis have the same SCP, this result covers indeed Koetter/Vardy's original conjecture. We will also explicitly construct the dual characteristic matrix Y. The details of our result also extend a theorem in [8, Thm. V.3], which showed how the characteristic span list of the dual of a minimal trellis of \mathcal{C} looks (recall from Theorem 2.5(b) that minimal trellises are KV-trellises).

A first step toward formulating the duality conjecture had been done in [10]. Our general assumption that both codes \mathcal{C} and \mathcal{C}^{\perp} have support \mathcal{I} is crucial for this result and the rest of this section.

Theorem 4.1 ([10, Thm. 5.12]) If $\mathcal{T} = [(a_l, b_l], l = 1, ..., n]$ is the characteristic span list of \mathcal{C} , then the characteristic span list of \mathcal{C}^{\perp} is given by $[(b_l, a_l], l = 1, ..., n]$.

As a consequence, one easily derives from Proposition 2.3(d) and Definition 2.4(iv) that if one selects k linearly independent characteristic generators of C with spans, say, $[(a_l, b_l], l = 1, ..., k]$ and n-k characteristic generators of C^{\perp} that do not have spans $(b_l, a_l], l = 1, ..., k$, then the resulting product trellises have the same SCP; this has been proven in [10, Prop. 5.13]. However, it is not guaranteed that those n - k dual characteristic generators are linearly independent and thus generate C^{\perp} . Indeed, this is not the case in general; see [8, Ex. III.13, Ex. V.2]. Furthermore, even if those dual generators are linearly independent, they may not give rise to a KV-trellis dual to the KV-trellis of C; see Example 4.10 below. Despite these obstacles, the result in [10, Prop. 5.13] indicates how the dual should look: if the dual of a KV-trellis of C is a KV-trellis of C^{\perp} , then its span list has to be given by the reversed complementary spans in the above sense. This is indeed the only option for the dual characteristic span list because in [8, Prop. III.14] it has been shown that different selections of n-k characteristic spans lead to non-isomorphic trellises. All of this leads to the following useful notation.

Definition 4.2 Let (X, \mathcal{T}) and $(Y, \hat{\mathcal{T}})$ be characteristic matrices of the codes \mathcal{C} and \mathcal{C}^{\perp} , respectively. A pair (\tilde{X}, \tilde{Y}) is called a *dual selection* of (X, Y) if $\tilde{X} \in \mathbb{F}^{k \times n}$ is a submatrix

of X and $\tilde{Y} \in \mathbb{F}^{(n-k)\times n}$ is a submatrix of Y such that their corresponding span lists $\mathcal{S} \subset \mathcal{T}$ and $\hat{\mathcal{S}} \subset \hat{\mathcal{T}}$ satisfy $\hat{\mathcal{S}} = \hat{\mathcal{T}} \setminus [(b,a] \mid (a,b] \in \mathcal{S}].$

In this section we will prove the following.

Theorem 4.3 Let (X, \mathcal{T}) be a characteristic pair of \mathcal{C} , and let $\hat{\mathcal{T}}$ be the characteristic span list of \mathcal{C}^{\perp} . Then there exists a characteristic matrix Y of \mathcal{C}^{\perp} such that each dual selection (\tilde{X}, \tilde{Y}) of (X, Y) satisfies the following properties.

- (1) $\operatorname{rk} \tilde{X} = k \iff \operatorname{rk} \tilde{Y} = n k.$
- (2) Let $\operatorname{rk} \tilde{X} = k$ and let $S \subset \mathcal{T}$ and $\hat{S} \subset \hat{\mathcal{T}}$ be the characteristic span lists of \tilde{X} and \tilde{Y} , respectively. Then the KV-trellises $T_{\tilde{X},S}$ and $T_{\tilde{Y},\hat{S}}$ are dual to each other.

Recall that in the situation (2) the trellises $T_{\tilde{X},S}$ and $T_{\tilde{Y},\hat{S}}$ represent \mathcal{C} and \mathcal{C}^{\perp} , respectively. We will show this result by explicitly constructing the dual matrix Y. Here is an outline of the quite technical procedure. Fix a characteristic matrix X of $\mathcal{C} = \ker H^{\mathsf{T}}$ and consider the BCJR-trellis $T_{(X,H,\mathcal{T})}$ based on the entire matrix X. For each $m = 1, \ldots, n$ consider the subtrellis T_m of T generated by all characteristic generators except the m-th one. Let the omitted generator have span $(a_m, b_m]$. Then one can find a cycle in the local dual $(T_m)^\circ$ with span $(b_m, a_m]$, and hence the associated edge-label sequence is a characteristic generator of \mathcal{C}^{\perp} . This is carried out in Proposition 4.6 (with the special case where $b_m = 0$ appearing in Lemma 4.5). Collecting all of these dual characteristic generators results in a characteristic matrix Y for \mathcal{C}^{\perp} . This is the matrix Y that will satisfy Theorem 4.3. Unfortunately, it is not a priori clear whether the pair (X, Y) satisfies the dual rank condition in part (1) of Theorem 4.3. Its somewhat technical proof is given in Proposition 4.9. For this step, the BCJR-representation of the trellises turns out to be crucial as it provides us with a close link between states and dual codewords. Once the dual rank condition is established, Theorem 4.3(2) is essentially a consequence of the construction. Indeed, pick a subset $\mathcal{K} \subset \mathcal{I}$ such that $|\mathcal{K}| = k$, and let \tilde{X} consist of the characteristic generators in X with span list $\mathcal{S} = [(a_l, b_l] \mid l \in \mathcal{K}].$ Then the KV-trellis $T_{(\tilde{X}, H, \mathcal{S})}$ is a subtrellis of T_m for each $m \notin \mathcal{K}$, and therefore each characteristic generator in Y with span in $[(b_m, a_m] \mid m \notin \mathcal{K}]$ appears in the local dual of $T_{(\tilde{X},H,S)}$. As a consequence, Theorem 4.3(1) tells us that if $rk \tilde{X} = k$, then these n - k generators give rise to the entire local dual of $T_{(\tilde{X},H,S)}$. The only detail that needs attention is the choice of the dual state spaces and the bilinear form. But this comes with the construction of Y, and it is easy to verify that it is indeed non-degenerate, see Proposition 4.11. Finally, in Theorem 4.12 we will make the isomorphism, hidden in (2) of Theorem 4.3, explicit by using suitable BCJR-representations.

Throughout this section we will use the following notation.

Notation 4.4 Let $C = \ker H^{\mathsf{T}} \subseteq \mathbb{F}^n$ be a k-dimensional code as in (1.2) and (1.3). Recall that we assume that both C and C^{\perp} have support \mathcal{I} . Furthermore, let (X, \mathcal{T}) be a fixed characteristic pair of C, and write $X = (x_{lj})_{l,j=0,\ldots,n-1}$, where the rows of X are sorted such that $\mathcal{T} = [(a_l, l], l = 0, \ldots, n-1]$. Hence we index the characteristic generators by $0, \ldots, n-1$. Since C^{\perp} has support \mathcal{I} , no characteristic span is empty and therefore $\mathcal{I} \setminus (a_l, l] = (l, a_l]$ for all $l \in \mathcal{I}$. Let $X_0^{\mathsf{T}}, \ldots, X_{n-1}^{\mathsf{T}}$ be the columns of X (we will always denote the j-th column of a matrix M by M_j^{T}). Let $T = T_{(X,H,\mathcal{T})}$ be the associated BCJR-trellis. The state space matrices of T will be denoted by $N_j \in \mathbb{F}^{n \times (n-k)}$, $j \in \mathcal{I}$. Recall that the BCJR-construction is completely row-wise, that is, the l-th rows of the matrices N_j solely depend on the generator with span $(a_l, l]$. In the following, we will consider various submatrices of N_j and X. In all of these instances, we will index the rows of the submatrices by the ending point l of the span $(a_l, l]$ they correspond to. In this situation, Theorem 2.12 reads as

- (1) $\operatorname{row}(N_i, l) = 0$ for all l such that $j \notin (a_l, l]$,
- (2) the set {row(N_j, l) | l such that $j \in (a_l, l]$ } is linearly independent for all $j \in \mathcal{I}$.

We start with the following technical lemma. It results in a characteristic generator $c \in C^{\perp}$ with characteristic span starting at time 0 and which is normalized in such a way that $c_0 = 1$.

Lemma 4.5 Let $G \in \mathbb{F}^{(n-1)\times n}$ be the matrix consisting of all except the first row of X and let $S = [(a_l, l], l = 1, ..., n - 1]$ be its span list. Consider the BCJR-trellis $T_{(G,H,S)}$ with state space matrices $Q_j \in \mathbb{F}^{(n-1)\times(n-k)}, j \in \mathcal{I}$.

Then there exists a unique vector $v \in \mathbb{F}^{n-k}$ such that $G_0^{\mathsf{T}} = Q_1 v^{\mathsf{T}}$. For $j \in \mathcal{I}$ put

$$w_j = \begin{cases} v, & \text{if } j \in (0, a_0], \\ 0, & \text{otherwise} \end{cases}$$

and define the dual codeword $c = (c_0, \ldots, c_{n-1}) \coloneqq vH \in \mathcal{C}^{\perp}$. Then

- (i) the dual codeword c has span $(0, a_0]$ and $c_0 = 1$.
- (ii) $Q_j v^{\mathsf{T}} = 0$ for all $j \notin (0, a_0]$.
- (iii) $Q_j w_j^\mathsf{T} + G_j^\mathsf{T} c_j Q_{j+1} w_{j+1}^\mathsf{T} = 0$ for all $j \in \mathcal{I}$.

Observe that due to Definition 2.4(iv), there exist k conventional characteristic spans (that is, spans not containing 0). Since $0 \in (a_0, 0]$, this implies that G contains all k characteristic generators with conventional spans. In other words, G contains an MSGM of C. Therefore im G = C, and the trellis $T_{(G,H,S)}$ does indeed represent C.

PROOF: By definition, $G = (x_{lj})_{\substack{l=1,\dots,n-1 \ j=0,\dots,n-1}}$. First notice that statement (iii) follows from the previous parts due to the identities $c_j = vH_j^{\mathsf{T}} = H_jv^{\mathsf{T}}$ and $Q_jw_j^{\mathsf{T}} = Q_jv^{\mathsf{T}}$ along with the recursion

$$Q_{j+1} = Q_j + G_j^\mathsf{T} H_j \text{ for } j \in \mathcal{I},$$

$$(4.1)$$

see (2.2). Thus it remains to show the existence of v and properties (i) and (ii). Let us first collect some properties of the matrices Q_j . By definition of BCJR-trellises, they are the submatrices of N_j obtained by omitting the first row, where N_j is as in Notation 4.4. Due to property (iv) of Definition 2.4, each index $j \in \mathcal{I}$ is contained in exactly n - k of the characteristic spans of \mathcal{C} . With the aid of Notation 4.4(1) and (2), we may therefore conclude that

$$\operatorname{rk} Q_{j} = \begin{cases} n-k, & \text{if } j \in (0, a_{0}], \\ n-k-1, & \text{if } j \notin (0, a_{0}]. \end{cases}$$
(4.2)

Moreover,

$$\operatorname{rk}(Q_{j}, G_{j}^{\mathsf{T}}) = \operatorname{rk}(G_{j}^{\mathsf{T}}, Q_{j+1}) = \operatorname{rk}(Q_{j}, G_{j}^{\mathsf{T}}, Q_{j+1}) = \begin{cases} \operatorname{rk}Q_{j}, & \text{if } j = a_{0}, \\ \operatorname{rk}Q_{j} + 1, & \text{if } j \neq a_{0}, \\ \operatorname{rk}Q_{j+1}, & \text{if } j = 0, \\ \operatorname{rk}Q_{j+1} + 1, & \text{if } j \neq 0. \end{cases}$$
(4.3)

The first two identities are a consequence of the recursion (4.1). As for the 4 cases of the last identity, recall that by Theorem 2.12, the trellis $T_{(G,H,S)}$ is isomorphic to the product trellis $T_{G,S}$. Therefore, we may apply Proposition 2.3(d). Since the set of starting points is $\mathcal{I} \setminus \{a_0\}$ and the set of ending points is $\mathcal{I} \setminus \{0\}$, we immediately obtain the 4 cases above.

Now (4.3) and (4.2) imply the existence of a unique $v \in \mathbb{F}^{n-k}$ such that $G_0^{\mathsf{T}} = Q_1 v^{\mathsf{T}}$. This results in $G_0^{\mathsf{T}} = (Q_0 + G_0^{\mathsf{T}} H_0)v^{\mathsf{T}}$, and thus $G_0^{\mathsf{T}}(1 - H_0v^{\mathsf{T}}) = Q_0v^{\mathsf{T}}$. But the second case in (4.3) shows that G_0^{T} is not in the column space of Q_0 , and therefore we may conclude

that $c_0 = H_0 v^{\mathsf{T}} = 1$ and $Q_0 v^{\mathsf{T}} = 0$. If $a_0 = n - 1$, then all of this shows that the span of c is contained in (0, n - 1]. Since Theorem 4.1 gives that the latter is a characteristic span of \mathcal{C}^{\perp} , Lemma 2.6 implies that the span of c equals (0, n - 1]. This proves (i) and (ii) for $a_0 = n - 1$. Let us now assume $a_0 < n - 1$. From the very definition of BCJR-trellises we know that

$$\operatorname{row}(Q_0, l) = \sum_{j=a_l}^{n-1} x_{lj} H_j, \ l = 1, \dots, n-1;$$

for the row indexing see Notation 4.4. Recall that $\{a_1, \ldots, a_{n-1}\} = \{0, \ldots, n-1\} \setminus \{a_0\}$. Hence there exists some l for which $a_l = n - 1$. Since $x_{l,a_l} \neq 0$, the identity $\operatorname{row}(Q_0, l)v^{\mathsf{T}} = 0$ yields $H_{n-1}v^{\mathsf{T}} = 0$. Now we can proceed recursively to obtain $c_j = H_jv^{\mathsf{T}} = 0$ for all $j = n-1, n-2, \ldots, a_0+1$. Hence the span of c is contained in $(0, a_0]$, and again Lemma 2.6 yields equality. This establishes (i).

As for (ii) we also proceed backwards. From $Q_0v^{\mathsf{T}} = 0$ and $H_{n-1}v^{\mathsf{T}} = 0$ we obtain $Q_{n-1}v^{\mathsf{T}} = (Q_0 - G_{n-1}^{\mathsf{T}}H_{n-1})v^{\mathsf{T}} = 0$ and successively $Q_jv^{\mathsf{T}} = 0$ for $j = n - 1, n - 2, \dots, a_0 + 1$, as desired. This completes the proof.

Using the behavior of the characteristic matrix and BCJR-trellises under the cyclic shift as summarized in Remark 2.11, the lemma above generalizes to any characteristic span. This leads to the following result.

Proposition 4.6 Fix $m \in \mathcal{I}$ and denote by $X^m \in \mathbb{F}^{(n-1)\times n}$ and $N_j^m \in \mathbb{F}^{(n-1)\times (n-k)}$ the submatrices of X and N_j , respectively, where the *m*-th row has been omitted. Denote the columns of X^m by $(X^m)_0^{\mathsf{T}}, \ldots, (X^m)_{n-1}^{\mathsf{T}}$. Then there exists a unique vector $v_m \in \mathbb{F}^{n-k}$ with the following properties:

- (a) $(X^m)_m^{\mathsf{T}} = N_{m+1}^m v_m^{\mathsf{T}},$
- (b) the dual codeword $c^m = (c_0^m, \ldots, c_{n-1}^m) := v_m H \in \mathcal{C}^{\perp}$ has span $(m, a_m]$ and satisfies $c_m^m = 1$,

(c)
$$N_j^m v_m^{\mathsf{I}} = 0$$
 for $j \notin (m, a_m]$,

(d) $N_j^m w_{m,j}^\mathsf{T} + (X^m)_j^\mathsf{T} c_j^m - N_{j+1}^m w_{m,j+1}^\mathsf{T} = 0$ for all $j \in \mathcal{I}$, where $w_{m,j} \in \mathbb{F}^{n-k}$ is defined as

$$w_{m,j} = \begin{cases} v_m, & \text{if } j \in (m, a_m], \\ 0, & \text{otherwise} \end{cases}$$

As a consequence, the pair $(Y, \hat{\mathcal{T}})$, where

$$Y = \begin{pmatrix} c^{0} \\ \vdots \\ c^{n-1} \end{pmatrix}, \ \hat{\mathcal{T}} = [(m, a_{m}], m = 0, \dots, n-1],$$
(4.4)

is a characteristic pair of \mathcal{C}^{\perp} .

PROOF: For m = 0 this is exactly the statement of Lemma 4.5: Part (a) translates into $G_0^{\mathsf{T}} = Q_1 v^{\mathsf{T}}$ in that lemma and (b), (c), (d) are (i), (ii), (iii) of that lemma, respectively. For m > 0 we may apply the left cyclic shift σ^m by m units on the characteristic generators. As detailed in Remark 2.11, this results in a characteristic pair (X^*, \mathcal{T}^*) for the code $\sigma^m(\mathcal{C}) = \operatorname{im} G^* = \ker H^{*\mathsf{T}}$ with BCJR-trellis $T_{(X^*, H^*, \mathcal{T}^*)}$. Its state space matrices are given by $N_j^* = N_{j+m}$ for $j \in \mathcal{I}$. This shifts the span $(a_m, m]$ to the span $(a_m - m, 0]$ and we may use Lemma 4.5 again. Applying now the inverse shift σ^{n-m} leads to the statements in (a) – (d). The consequence about $(Y, \hat{\mathcal{T}})$ is clear due to Theorem 4.1.

One should observe that the identities in Proposition 4.6(d) indicate that the dual codeword c^m and the states $w_{m,j}, j \in \mathcal{I}$, give rise to a cycle in the local dual of the trellis $T_{(X^m,H,\mathcal{T}^m)}$, where \mathcal{T}^m is the characteristic span list of X^m ; see also (3.1). We will make this precise later on after specifying the dual state spaces and the bilinear form as needed for the local dualization.

Example 4.7 Consider the self-dual code from Example 2.10(b) with characteristic matrix X given in (2.5) and characteristic span list $\mathcal{T} = [(3,0], (2,1], (1,2], (0,3]]$ ordered as required in Notation 4.4. The matrix, displaying all information about the trellis $T_{(X,G,\mathcal{T})}$, is given by

$$S = (N_0 | X_0^{\mathsf{T}} | N_1 | X_1^{\mathsf{T}} | N_2 | X_2^{\mathsf{T}} | N_3 | X_3^{\mathsf{T}} | N_0) = \begin{pmatrix} 1 & 0 & | & 1 & | & 0 & 0 & | & 0 & | & 0 & | & 0 & | & 0 & | & 1 & | & 1 & 0 \\ 1 & 1 & 0 & | & 1 & 1 & | & 1 & | & 0 & 0 & | & 1 & 1 & 1 \\ 0 & 0 & | & 0 & | & 0 & | & 1 & | & 1 & | & 1 & | & 0 & | & 0 & | & 0 & | \\ 0 & 0 & | & 1 & | & 0 & | & 1 & | & 1 & | & 1 & | & 0 & | & 0 & | & 0 & | \\ 0 & 0 & | & 1 & | & 0 & | & 1 & | & 1 & | & 1 & | & 0 & | & 1 & | & 0 & | \\ \end{pmatrix}.$$

Omitting the *m*-th row from S, where m = 0, ..., 3, we see that

$$(X^{0})_{0}^{\mathsf{T}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = N_{1}^{0} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ (X^{1})_{1}^{\mathsf{T}} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = N_{2}^{1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ (X^{2})_{2}^{\mathsf{T}} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = N_{3}^{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ (X^{3})_{3}^{\mathsf{T}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = N_{0}^{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Hence $v_0 = (11)$, $v_1 = (01)$, $v_2 = (10)$, and $v_3 = (11)$. Recalling that H = G due to selfduality, we obtain $c^0 = v_0 G = (1001)$, $c^1 = v_1 G = (0110)$, $c^2 = v_2 G = (1111)$, and $c^3 = v_3 G = (1001)$. Thus, the characteristic pair $(Y, \hat{\mathcal{T}})$ for $\mathcal{C}^{\perp} = \mathcal{C}$, resulting from this procedure, is given by

$$Y = \begin{pmatrix} c^{0} \\ c^{1} \\ c^{2} \\ c^{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \hat{\mathcal{T}} = [(0,3], (1,2], (2,1], (3,0]].$$

One should notice that even though the code C is self-dual, the matrix Y is not identical to the characteristic matrix X of C in (2.5), with which we started the procedure! As mentioned earlier, the identity in Proposition 4.6(d) gives rise to cycles in certain local duals, namely $(w_{m,0}, c_0^m, w_{m,1}, c_1^m, w_{m,2}, c_2^m, w_{m,3}, c_3^m, w_{m,0})$. The matrix

contains these four cycles. Proposition 4.6(d) now states that for m = 0, ..., 3, the *m*-th row of \hat{S} is in the local dual of the trellis determined by the three rows of S not having index m. Using the definition of the local dual in (3.1) this can also be verified directly.

Now we can formulate the main result of this section.

Theorem 4.8 Let (X, \mathcal{T}) be as in Notation 4.4 and Y as in Proposition 4.6. Then the pair (X, Y) satisfies Theorem 4.3.

The proof consists of several steps. We begin with Part (1) of Theorem 4.3. It is worth pointing out that for the proof of this result the BCJR-description of the trellises involved is crucial. It directly links the "dual states" v_m to the associated dual codewords c^m , making the second equivalence in (4.5) below obvious. **Proposition 4.9** Let v_m and $(Y, \hat{\mathcal{T}})$ be as in Proposition 4.6, and let $\mathcal{K} \cup \hat{\mathcal{K}} = \mathcal{I}$ be an index partition such that $|\mathcal{K}| = k$. Let $\tilde{X} \in \mathbb{F}^{k \times n}$ be the submatrix of X consisting of the generators with spans in the list $\mathcal{S} := [(a_l, l] \mid l \in \mathcal{K}]$, while $\tilde{Y} \in \mathbb{F}^{(n-k) \times n}$ is the submatrix of Y consisting of the dual generators with spans in $\hat{\mathcal{S}} := [(m, a_m] \mid m \in \hat{\mathcal{K}}]$. Then (\tilde{X}, \tilde{Y}) is a dual selection of (X, Y) in the sense of Definition 4.2 and

$$\operatorname{rk} \tilde{X} = k \Longleftrightarrow \{ v_m \mid m \in \hat{\mathcal{K}} \} \text{ is linearly independent } \Longleftrightarrow \operatorname{rk} \tilde{Y} = n - k.$$

$$(4.5)$$

PROOF: By Proposition 4.6(b) we have $\tilde{Y} = (v_m H)_{m \in \hat{\mathcal{K}}}$. Using that $\operatorname{rk} H = n - k$, this immediately establishes the second equivalence in (4.5).

As for the first equivalence, we need some preparation. Recall the matrixes N_j from Notation 4.4, and define $\tilde{N}_j \in \mathbb{F}^{k \times (n-k)}$ as the submatrices of N_j consisting of the rows with indices in \mathcal{K} . Consider also the matrices X^m and N_j^m as defined in Proposition 4.6. Since \mathcal{K} and $\hat{\mathcal{K}}$ are disjoint, we observe that for each $m \in \hat{\mathcal{K}}$ the matrices \tilde{X} and \tilde{N}_j are the submatrices of X^m and N_j^m , respectively, consisting of the rows with indices in \mathcal{K} . Denote the columns of \tilde{X} by $\tilde{X}_0^{\mathsf{T}}, \ldots, \tilde{X}_{n-1}^{\mathsf{T}}$. Then Proposition 4.6(a) yields

$$\tilde{X}_m^{\mathsf{T}} = \tilde{N}_{m+1} v_m^{\mathsf{T}} \text{ for all } m \in \hat{\mathcal{K}}.$$
(4.6)

Now we can prove the equivalence.

" \Leftarrow " First suppose that rk $\tilde{X} < k$, and let $\alpha \tilde{X} = 0$ for some $\alpha \in \mathbb{F}^k \setminus \{0\}$. We will index the vector α by $l \in \mathcal{K}$ to be consistent with our indexing of the rows of \tilde{X} . With the aid of the recursion $\tilde{N}_{j+1} = \tilde{N}_j + \tilde{X}_j^{\mathsf{T}} H_j$, we compute $\alpha \tilde{N}_{j+1} = \alpha (\tilde{N}_j + \tilde{X}_j^{\mathsf{T}} H_j) = \alpha \tilde{N}_j$ for all $j \in \mathcal{I}$. Thus, $\alpha \tilde{N}_0 = \alpha \tilde{N}_1 = \ldots = \alpha \tilde{N}_{n-1} =: w$. Suppose w = 0, and hence $\alpha \in \bigcap_{j=0}^{n-1} \ker \tilde{N}_j$. But then Notation 4.4(2) tells us that the identity $\alpha \tilde{N}_j = 0$ leads to $\alpha_l = 0$ for each $l \in \mathcal{K}$ such that $j \in (a_l, l]$. Since no span $(a_l, l]$ is empty, this results in $\alpha = 0$, contradicting our assumption. Hence $w \neq 0$. Next, along with (4.6), the identity $\alpha \tilde{X} = 0$ implies $\alpha \tilde{N}_{m+1} v_m^{\mathsf{T}} = 0$ and hence $wv_m^{\mathsf{T}} = 0$ for all $m \in \hat{\mathcal{K}}$. This shows that the matrix $(v_m^{\mathsf{T}}, m \in \hat{\mathcal{K}}) \in \mathbb{F}^{(n-k) \times (n-k)}$ is singular, and therefore the set $\{v_m \mid m \in \hat{\mathcal{K}}\}$ is linearly dependent.

" \Longrightarrow " Assume rk $\tilde{X} = k$. Then $\mathcal{C} = \operatorname{im} \tilde{X} = \operatorname{ker} H^{\mathsf{T}}$, and the matrices \tilde{N}_j are the state space matrices of the BCJR-trellis $T_{(\tilde{X},H,\mathcal{S})}$. Let $w(v_m^{\mathsf{T}}, m \in \hat{\mathcal{K}}) = 0$ for some $w \in \mathbb{F}^{n-k}$. We first show that $w \in \operatorname{im} N_j^m$ for all $m \in \hat{\mathcal{K}}$ and all $j \in \mathcal{I}$. In order to do so, notice that the rank condition in (4.2) generalizes to N_j^m as rk $N_j^m = n - k$ if $j \in (m, a_m]$ and rk $N_j^m = n - k - 1$ if $j \notin (m, a_m]$. Thus, if $j \in (m, a_m]$, then $\operatorname{im} N_j^m = \mathbb{F}^{n-k}$, and hence $w \in \operatorname{im} N_j^m$. If $j \notin (m, a_m]$, then rk $N_j^m = n - k - 1$ and Proposition 4.6(c) yields $\operatorname{im} N_j^m = (\operatorname{im} v_m)^{\perp}$ in \mathbb{F}^{n-k} . Thus, $w \in \operatorname{im} N_j^m$, as desired.

Now we have $w = \alpha_j^m N_j^m$ for all $m \in \hat{\mathcal{K}}$ and some $\alpha_j^m \in \mathbb{F}^{n-1}$. Then Notation 4.4(1) tells us that for every $m \in \hat{\mathcal{K}}$ we have $w \in \text{span}\{\text{row}(N_j, l) \mid l \text{ such that } j \in (a_l, l], l \neq m\}$. From this and the linear independence of the nonzero rows in N_j , see Notation 4.4(2), we conclude that $w \in \text{span}\{\text{row}(N_j, l) \mid l \notin \hat{\mathcal{K}} \text{ and } j \in (a_l, l]\}$. But the latter space is im \tilde{N}_j , and, since $j \in \mathcal{I}$ was arbitrary, this proves that $w \in \bigcap_{j=0}^{n-1} \text{im } \tilde{N}_j$. Using that $T_{(\tilde{X}, H, S)}$ is a KV-trellis, we may apply Theorem 2.14 and conclude that w = 0. As a consequence, $\{v_m \mid m \in \hat{\mathcal{K}}\}$ is linearly independent. This concludes the proof of the remaining equivalence in (4.5).

So far we have proven the dual rank property in Theorem 4.3(1). It is worth noting that this does not guarantee that part (2) of that theorem is also satisfied. This is due to the fact that KV-trellises based on the same span selection but on different generators need not be isomorphic. Indeed, we have the following example.

Example 4.10 Consider again the code $C \subseteq \mathbb{F}_3^4$ from Example 2.10(c) with the characteristic matrix X as given there. The dual code has characteristic span list $\hat{\mathcal{T}} = [(1,0], (0,1], (3,2], (2,3]]$ and the matrices

$$Y_1 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

are both characteristic matrices of \mathcal{C}^{\perp} . Both are normalized, that is, the generators have coordinate 1 at the starting point of their span. Proposition 4.6, applied to the characteristic matrix X of \mathcal{C} , produces the matrix Y_1 . It is easy to see that both pairs (X, Y_i) , i = 1, 2, satisfy the dual rank condition (1) of Theorem 4.3 for all dual selections. It is a bit more tedious to show that only Y_1 satisfies part (2) of Theorem 4.3 for all full rank dual selections. Indeed, the two KV-trellises resulting from the first two rows of Y_1 and Y_2 are not isomorphic², and only the trellis resulting from Y_1 is the dual of the KV-trellis of \mathcal{C} corresponding to the last two rows of X.

Thus, it remains to prove (2) of Theorem 4.3 in order to complete the proof of Theorem 4.8. Precisely, if $rk \tilde{X} = k$ in Proposition 4.9, then the pairs (\tilde{X}, S) and (\tilde{Y}, \hat{S}) give rise to KV-trellises of C and C^{\perp} , respectively, and we aim to show that these trellises are duals of each other. Since for KV-trellises, the BCJR-dual and the local dual are isomorphic, we may use either approach. The initial idea in Proposition 4.6 was motivated by local dualization, and therefore we will use that construction. While we do have a description of the state spaces V_j of $T_{(\tilde{X},H,S)}$, we do not yet have dual state spaces \hat{V}_j and a non-degenerate bilinear form on $V_j \times \hat{V}_j$ which are needed to establish a local duality as in Theorem 3.1. Example 4.7 shows that the standard bilinear form on \mathbb{F}^{n-k} restricted to V_j is in general degenerate, and therefore $\hat{V}_j = V_j$ along with this bilinear form is not an option for the local dualization: take, for instance, $\mathcal{K} = \{0, 1\}$; then $V_1 = \operatorname{im}(1, 1)$. On the other hand, the example and the construction preceding it indicate what the dual state spaces should be: the spaces generated by the "dual states" v_m . All this leads to the following result.

Proposition 4.11 Let \mathcal{K} , $\hat{\mathcal{K}}$, (\tilde{X}, \mathcal{S}) , and $(\tilde{Y}, \hat{\mathcal{S}})$ be as in Proposition 4.9. Assume $\operatorname{rk} \tilde{X} = k$, and let $V_j := \operatorname{im} \tilde{N}_j$, where $\tilde{N}_j \in \mathbb{F}^{k \times (n-k)}$ are the state space matrices of the KV-trellis $T := T_{(\tilde{X}, H, \mathcal{S})}$. For $j \in \mathcal{I}$ define the matrices $P_j \in \mathbb{F}^{(n-k) \times (n-k)}$ via

$$\operatorname{row}(P_j, m) = w_{m,j} = \left\{ \begin{array}{l} v_m, & \text{if } j \in (m, a_m] \\ 0, & \text{otherwise} \end{array} \right\} \text{ for } m \in \hat{\mathcal{K}}.$$

$$(4.7)$$

Put $\hat{V}_j := \operatorname{im} P_j$ and $\hat{E}_j = \operatorname{im} (P_j, \tilde{Y}_j^{\mathsf{T}}, P_{j+1})$ for $j \in \mathcal{I}$, where \tilde{Y}_j^{T} denotes the *j*-th column of \tilde{Y} . Then

(1) $\operatorname{rk} \tilde{N}_j = \operatorname{rk} P_j = \operatorname{rk} \tilde{N}_j P_j^{\mathsf{T}}$ for all $j \in \mathcal{I}$. As a consequence, dim $V_j = \dim \hat{V}_j$ for all $j \in \mathcal{I}$ and the bilinear form

$$V_j \times \hat{V}_j \longrightarrow \mathbb{F}, \quad (\alpha \tilde{N}_j, \beta P_j) \longmapsto \alpha \tilde{N}_j P_j^\mathsf{T} \beta^\mathsf{T}$$

$$(4.8)$$

is non-degenerate.

(2) The trellis $\hat{T} = (\hat{V}, \hat{E})$, where $\hat{V} = \bigcup_{j=0}^{n-1} \hat{V}_j$ and $\hat{E} = \bigcup_{j=0}^{n-1} \hat{E}_j$, is a linear trellis representing \mathcal{C}^{\perp} and $\hat{T} = T^{\circ}$, where we dualize with respect to the bilinear form (4.8). Moreover, the trellis \hat{T} is isomorphic to the $KV_{(Y,\hat{T})}$ -trellis $T_{\tilde{Y},\hat{S}}$.

²In the first trellis, the unique cycle representing the codeword (1112) $\in C^{\perp}$ passes through the zero state at time 1, but this is not the case in the second trellis.

PROOF: (1) Recall that by property (iv) of Definition 2.4 every index j is contained in exactly n-k characteristic spans. Thus, with the aid of Notation 4.4(1) and (2), and $|\hat{\mathcal{K}}| = n-k$ we compute

$$\operatorname{rk} \tilde{N}_{j} = |\{l \in \mathcal{K} \mid j \in (a_{l}, l]\}| = n - k - |\{m \in \hat{\mathcal{K}} \mid j \in (a_{m}, m]\}|$$
$$= |\{m \in \hat{\mathcal{K}} \mid j \notin (a_{m}, m]\}| = |\{m \in \hat{\mathcal{K}} \mid j \in (m, a_{m}]\}|.$$

But the last quantity is exactly $\operatorname{rk} P_j$ due to Proposition 4.9. It remains to show $\operatorname{rk} \tilde{N}_j = \operatorname{rk} \tilde{N}_j P_j^{\mathsf{T}}$. We will do this by showing $\operatorname{ker} \tilde{N}_j P_j^{\mathsf{T}} = \operatorname{ker} \tilde{N}_j$. Clearly, we have " \supseteq ". For the converse, let $\alpha \tilde{N}_j P_j^{\mathsf{T}} = 0$ for some $\alpha \in \mathbb{F}^k$. Then $\alpha \tilde{N}_j v_m^{\mathsf{T}} = 0$ for all $m \in \hat{\mathcal{K}}$ such that $j \in (m, a_m]$. Along with Proposition 4.6(c) this leads to $\alpha \tilde{N}_j v_m^{\mathsf{T}} = 0$ for all $m \in \hat{\mathcal{K}}$. But due to Proposition 4.9 the matrix $(v_m \mid m \in \hat{\mathcal{K}}) \in \mathbb{F}^{(n-k) \times (n-k)}$ is non-singular and thus we conclude $\alpha \tilde{N}_j = 0$. This concludes the proof of the rank identities. The non-degeneracy of the bilinear form follows from $\operatorname{ker} \tilde{N}_j P_j^{\mathsf{T}} = \operatorname{ker} \tilde{N}_j$ and $\operatorname{ker} P_j \tilde{N}_j^{\mathsf{T}} = \operatorname{ker} P_j$.

(2) We first show $\hat{T} = T^{\circ}$. In order to do so, we have to prove $\hat{E}_j = (\tilde{E}_j)^{\circ}$, where $\tilde{E}_j = im(\tilde{N}_j, \tilde{X}_j^{\mathsf{T}}, \tilde{N}_{j+1})$ are the transition spaces of T, and $(\tilde{E}_j)^{\circ}$ are the duals in the sense of (3.1). The rows of $(P_j, \tilde{Y}_j^{\mathsf{T}}, P_{j+1})$ are given by $(w_{m,j}, c_j^m, w_{m,j+1})$ for $m \in \hat{\mathcal{K}}$. Since $(\tilde{N}_j, \tilde{X}_j^{\mathsf{T}}, \tilde{N}_{j+1})$ is a submatrix of $(N_j^m, (X^m)_j^{\mathsf{T}}, N_{j+1}^m)$ for all $m \in \hat{\mathcal{K}}$, Proposition 4.6(d) shows that $\hat{E}_j \subseteq (\tilde{E}_j)^{\circ}$. For the converse we will show that the two spaces have the same dimension. Let $s_j \coloneqq rk \tilde{N}_j = rk P_j$ and $e_j = \dim \tilde{E}_j$. Then Theorem 3.1 tells us that $e_j^{\circ} \coloneqq \dim (\tilde{E}_j)^{\circ} = s_j + s_{j+1} + 1 - e_j$. Moreover, being a KV-trellis, T is isomorphic to the product trellis $T_{\tilde{X},S}$. Therefore, Proposition 2.3(d) tells us that $e_j = s_{j+1}$ if $j \in \hat{\mathcal{K}}$ and $e_j = s_{j+1} + 1$ if $j \in \mathcal{K}$. Thus

$$e_j^{\circ} = s_j + 1$$
 if $j \in \hat{\mathcal{K}}$ and $e_j^{\circ} = s_j$ if $j \in \mathcal{K}$.

Using that dim $\hat{E}_j = \operatorname{rk}(P_j, \tilde{Y}_j^{\mathsf{T}}, P_{j+1}) \ge \operatorname{rk} P_j = s_j$ for all j, we obtain dim $\hat{E}_j \ge e_j^{\circ}$ for $j \in \mathcal{K}$. If $j \notin \mathcal{K}$ then $j \in \hat{\mathcal{K}}$ and $\operatorname{row}(P_j, j) = 0$ due to (4.7). Since the j-th entry of \tilde{Y}_j^{T} is given by $e_j^j = 1$, this yields dim $\hat{E}_j = \operatorname{rk}(P_j, \tilde{Y}_j^{\mathsf{T}}, P_{j+1}) \ge \operatorname{rk} P_j + 1 = s_j + 1 = e_j^{\circ}$. Along with $\hat{E}_j \subseteq (\tilde{E}_j)^{\circ}$, all of this shows $\hat{E}_j = (\tilde{E}_j)^{\circ}$, and therefore $\hat{T} = T^{\circ}$. Consequently, \hat{T} represents \mathcal{C}^{\perp} .

It remains to show that \hat{T} is isomorphic to the $\mathrm{KV}_{(Y,\hat{T})}$ -trellis $T_{\tilde{Y},\hat{S}}$. For this remember that the latter has state spaces im M_j and transition spaces im $(M_j, \tilde{Y}_j^{\mathsf{T}}, M_{j+1})$, where M_j is the state space matrix as in Definition 2.2 based on the span list $[(m, a_m] \mid m \in \hat{\mathcal{K}}]$. As a consequence, the rows of M_j with index m such that $j \in (m, a_m]$ are linearly independent while all other rows are zero. Comparing this with (4.7) and making use of Proposition 4.9, we see that $\alpha M_j \mapsto \alpha P_j$ induces a well-defined isomorphism between im M_j and $\hat{V}_j = \mathrm{im} P_j$. Now the very definitions of the transition spaces of \hat{T} and $T_{\tilde{Y},\hat{S}}$ show that this gives rise to a trellis isomorphism. This concludes the proof.

This establishes the proof of Theorem 4.8.

Having the dual pairing (X, Y) of characteristic matrices allows us to express the duality of their KV-trellises in terms of BCJR-representations. Indeed, let (\tilde{X}, \tilde{Y}) be a dual selection of (X, Y) with span lists S and \hat{S} as in Proposition 4.9 and such that $\operatorname{rk} \tilde{X} = k$, thus $\operatorname{rk} \tilde{Y} =$ n - k. Then \tilde{Y} and \tilde{X} are parity check matrices of C and C^{\perp} , respectively, and may be used for the BCJR-representations of the codes. This results in the KV-trellises $T_{(\tilde{X},\tilde{Y},S)}$ of C and $T_{(\tilde{Y},\tilde{X},\hat{S})}$ of C^{\perp} . By definition of the dual matrix Y, we have $\tilde{Y} = UH$, where $U = (v_m)_{m \in \hat{\mathcal{K}}}$. Moreover, U is non-singular due to Proposition 4.9. Thus, the trellis $T_{(\tilde{X},\tilde{Y},S)}$ is isomorphic to the trellis $T_{(\tilde{X},H,S)}$; see also [8, Rem. IV.4]. As a consequence, the trellis \hat{T} of Proposition 4.11(2), being the local dual of $T_{(\tilde{X},H,S)}$, is isomorphic to the local dual of $T_{(\tilde{X},\tilde{Y},S)}$, which in turn is isomorphic to the BCJR-dual $T_{(\tilde{X},\tilde{Y},S)}^{\perp}$, due to Theorem 3.5. A tedious, but straightforward matrix computation shows that this trellis satisfies the following symmetry.

Theorem 4.12 Let the data be as in Proposition 4.9 and let $\operatorname{rk} \tilde{X} = k$. Then

$$T^{\perp}_{(\tilde{X},\tilde{Y},\mathcal{S})} = T_{(\tilde{Y},\tilde{X},\hat{\mathcal{S}})}.$$

The situation described in Example 4.10 may be used to show that this symmetry is not true for full rank dual selections of arbitrary pairs (X, Y) of characteristic matrices, even if they satisfy the dual rank condition.

Let us close the paper with the following remarks pertaining to specific classes of codes. First, if $\mathcal{C} \subseteq \mathbb{F}^n$ is a cyclic code, then it is easy to see that \mathcal{C} has only one characteristic matrix (up to scalar factors). It is given by the *n* cyclic shifts of the generator polynomial, see also [9, Lem. 2]. As a consequence, this matrix and the corresponding characteristic matrix for the dual code satisfy Theorem 4.3. The dual rank condition of this theorem has already been proven in [9, Thm. 6].

Second, let us return to self-dual codes. Notice that for the self-dual code in Example 2.10(b), the pair (X, X), where X is as in (2.5), does not satisfy the dual rank condition of Theorem 4.3(1). It is easy to check that X is not the lexicographically first characteristic matrix of that code (the lexicographically first characteristic matrix is obtained by choosing for each span the lexicographically first codeword having that span, and where the lexicographic ordering starts at the starting point of that span). In the proof of [9, Thm. 1] it was stated that for any self-dual code, the pair (X, X), where X is the lexicographically first characteristic matrix, satisfies the dual rank condition of Theorem 4.3(1). While this is indeed the case for the particular code in Example 2.10(b), this is, unfortunately, not true in general.

Example 4.13 Consider the extended [8,4,4]-Hamming code $\mathcal{C} \subseteq \mathbb{F}_2^8$ generated by

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

Then

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$
$$\mathcal{T} = [(5,0], (4,1], (7,2], (6,3], (1,4], (0,5], (3,6], (2,7]]]$$

form a characteristic pair, and X is the lexicographically first characteristic matrix of C. It is easy to see that the pair (X, X) does not satisfy the dual rank condition in Theorem 4.3(1): the rows with spans (4,1], (7,2], (6,3], (3,6] are linearly independent, whereas this is not the case for the rows with spans (5,0], (4,1], (7,2], (0,5] (these are the spans that are not the reversed spans of the first list).

References

- A. Al-Bashabsheh and Y. Mao. Valiant transform of Forney graphs. Preprint 2010. arXiv: 1004.3833v2.
- [2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Inform. Theory*, IT-20:284–287, 1974.
- [3] 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.
- [4] G. D. Forney, Jr. Codes on graphs: Normal realizations. *IEEE Trans. Inform. Theory*, IT-47:520–548, 2001.
- [5] G. D. Forney, Jr. Minimal realizations of linear systems: The "shortest basis" approach. Preprint 2009. To appear in *IEEE Trans. Inform. Theory.* arXiv: 0910.4336v3.
- [6] G. D. Forney, Jr. Codes on graphs: Duality and MacWilliams identities. Preprint 2010. To appear in *IEEE Trans. Inform. Theory.* arXiv: 0911.5508v4.
- [7] G. D. Forney, Jr. and M. D. Trott. The dynamics of group codes: State spaces, trellis diagrams, and canonical encoders. *IEEE Trans. Inform. Theory*, IT-39:1491–1513, 1993.
- [8] H. Gluesing-Luerssen and E. Weaver. Linear tail-biting trellises: Characteristic generators and the BCJR-construction. Preprint 2010. To appear in *IEEE Trans. Inform. Theory.* arXiv: 1003.4539v1.
- [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. The structure of tail-biting trellises: Minimality and basic principles. IEEE Trans. Inform. Theory, IT-49:2081–2105, 2003.
- [11] F. R. Kschischang and V. Sorokine. On the trellis structure of block codes. *IEEE Trans. Inform. Theory*, IT-41:1924–1937, 1995.
- [12] S. Lin and R. Y. Shao. General structure and construction of tail-biting trellises for linear block codes. In *Proceedings of the 2000 IEEE International Symposium on Information Theory*, page 117, 2000.
- [13] Y. Mao and F. R. Kschischang. On factor graphs and the Fourier transform. *IEEE Trans. Inform. Theory*, IT-51:1635–1649, 2005.
- [14] R. J. McEliece. On the BCJR Trellis for linear block codes. IEEE Trans. Inform. Theory, IT-42:1072–1092, 1996.
- [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] Y. Shany and Y. Be'ery. Linear tail-biting trellises, the square root bound, and applications for Reed-Muller codes. *IEEE Trans. Inform. Theory*, IT-46:1514–1523, 2000.
- [17] Q. Yang and Z. Qin. An algorithm for computing characteristic matrices of group codes over finite abelian groups. In The 4th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2008), Dalian 2008, 2008.
- [18] J. Zhou and T. Ohtsuki. Isomorphic constructions of tail-biting trellises for linear block codes. In Proceedings of the 2nd International Conference on Communications and Networking in China (ChinaCom'07), Shanghai 2007, pages 585 – 589, 2008.