On the MacWilliams Identity for Convolutional Codes

Heide Gluesing-Luerssen, *Member, IEEE*, and Gert Schneider

***Abstract*—The adjacency matrix associated with a convolutional code collects in a detailed manner information about the weight distribution of the code. A MacWilliams Identity Conjecture, stating that the adjacency matrix of a code fully determines the adjacency matrix of the dual code, will be formulated, and an explicit formula for the transformation will be stated. The formula involves the MacWilliams matrix known from complete weight enumerators of block codes. The conjecture will be proven for the class of convolutional codes where either the code itself or its dual does not have Forney indices bigger than one. For the general case, the conjecture is backed up by many examples, and a weaker version will be established.**

***Index Terms*—Controller canonical form, convolutional codes, MacWilliams identity, weight adjacency matrix, weight distribution.**

I. Introduction

Two of the most famous results in block code theory are MacWilliams' Identity Theorem and Equivalence Theorem [1], [2]. The first one relates the weight enumerator of a block code to that of its dual code. The second one states that two isometric codes are monomially equivalent. The impact of these theorems for practical as well as theoretical purposes is well known, see for instance [3, Chs. 11.3, 6.5, and 19.2] or the classification of constant weight codes in [4, Theorem 7.9.5].

The paramount importance of the weight function in coding theory makes an understanding of weight enumerators, isometries, and, more explicitly, possible versions of the MacWilliams theorems a must for the analysis of any class of codes. For instance, after realizing the relevance of block codes over finite rings, both theorems have seen generalizations to this class of codes, see for instance [5] and [6]. For convolutional codes, the question of a MacWilliams Identity Theorem has been posed already about 30 years ago. In 1977, Shearer and McEliece [7] considered the weight enumerator for convolutional codes as introduced by Viterbi [8]. It is a formal power series in two variables counting the number of irreducible ("atomic") codewords of given weight and length; for the coding-theoretic relevance see, e.g., [8, Sec. VII] and [9, Sec. 4.3]. Unfortunately, a simple example in [7] made clear that a MacWilliams Identity does not exist for these objects. A main step forward has been made in 1992 when Abdel-Ghaffar [10] considered a more refined weight counting object: the weight enumerator state diagram. For unit constraint-length codes he derives a MacWilliams Identity in form of a list of separate formulas relating the labels of this diagram to those of the dual code.

In this paper, we will present a MacWilliams Identity for the class of convolutional codes where either the code or its dual does not have Forney indices bigger than one. Duality of codes will be defined in the module-theoretic way based on the vanishing of the canonical bilinear form on $\mathbb{F}[D]^n$ (in Remark II.7, we will briefly touch upon other notions of duality as well). Our result generalizes not only the block code case, but also Abdel-Ghaffar's transformation for unit constraint-length codes. We will show in Section VI that the list of identities given in [10] can be written in closed form just like in our MacWilliams Identity. In addition to the result just mentioned, we will also formulate an explicit conjecture on a MacWilliams Identity for all classes of convolutional codes. It is backed up by a wealth of examples, and a weaker version will be proven.

The weight counting object in our considerations is the so-called adjacency matrix of the encoder. This matrix has been discussed in detail by McEliece [11], but appears already in different notations earlier in the literature. Indeed, one can show that it basically coincides with the labels of the weight enumerator state diagram as considered in [10]. The adjacency matrix is defined via a state-space description of the encoder as introduced in [12]. In this sense, our approach follows a series of papers where system-theoretic methods have been used successfully in order to investigate convolutional codes, see, for instance, [13]–[16]. The matrix is labeled by the set of all state pairs, and each entry contains the weight enumerator of all outputs associated with the corresponding state pair. The whole matrix contains considerably more detailed information about the code than the weight enumerator discussed above. Indeed, it is well known [11], [17] how to derive the latter from the adjacency matrix. Furthermore, in [17] it has been shown that, after factoring out the group of state-space isomorphisms, the adjacency matrix turns into an invariant of the code, called the generalized adjacency matrix.

The main outline of our arguments is as follows. In the next section, we will introduce two block codes canonically associated with a convolutional code. They are crosswise dual to the corresponding block codes of the dual convolutional code. Later on, this fact will allow us to apply the MacWilliams Identity for block codes suitably. Indeed, in Section III, we will introduce the adjacency matrix Λ and show that its nontrivial entries are given by the weight enumerators of certain cosets of these block codes. The main ingredient for relating Λ with the adjacency matrix of the dual will be a certain transformation matrix $\mathcal{H}$ as it also appears for the *complete* weight enumerator of block codes. This matrix will be studied in Section IV, and a first application

Manuscript received December 15, 2005; revised November 2, 2007.

H. Gluesing-Luerssen is with the Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA (e-mail: heidegl@ms.uky.edu).

G. Schneider is with the Department of Mathematics, University of Groningen, P. O. Box 800, 9700 AV Groningen, The Netherlands (e-mail: schneider@math.rug.nl).

Communicated by G. Zémor, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2008.917664

to the adjacency matrix will be carried out. In Section V, we will be able to show our main results. First, we prove that entrywise application of the block code MacWilliams Identity for the matrix $\mathcal{H}\Lambda^t\mathcal{H}^{-1}$ will result in a matrix that up to reordering of the entries coincides with the adjacency matrix of the dual code. Second, for codes where the dual does not have Forney indices bigger than one, we will show that the reordering of the entries comes from a state-space isomorphism. As a consequence, the resulting matrix is indeed a representative of the generalized adjacency matrix of the dual code. This is exactly the contents of our MacWilliams Identity Theorem.

We end the Introduction with recalling some of the basic notions of convolutional codes. Throughout this paper let

$$\left.\begin{array}{l}\mathbb{F}=\mathbb{F}_q \text{ be a finite field with } q=p^s \text{ elements}\\ \text{where } p \text{ is prime and } s\in\mathbb{N}\end{array}\right\}. \qquad \text{(I.1)}$$

A *k-dimensional convolutional code of length n* is a submodule $\mathcal{C}$ of $\mathbb{F}[D]^n$ of the form

$$\mathcal{C}=\operatorname{im} G := \{uG \mid u\in\mathbb{F}[D]^k\}$$

where G is a *basic* matrix in $\mathbb{F}[D]^{k\times n}$, i.e. there exists some matrix $\tilde{G}\in\mathbb{F}[D]^{n\times k}$ such that $G\tilde{G}=I_k$. In other words, G is noncatastrophic and delay-free. We call G an *encoder* and the number $\delta := \max\{\deg\gamma \mid \gamma \text{ is a } k\text{-minor of } G\}$ is said to be the *degree* of the code $\mathcal{C}$. A code having these parameters is called an (n,k,δ) code. A basic matrix $G\in\mathbb{F}[D]^{k\times n}$ with rows $g_1,\dots,g_k\in\mathbb{F}[D]^n$ is said to be *minimal* if $\sum_{i=1}^k \deg(g_i)=\delta$. For characterizations of minimality see, e. g., [18, Main Theorem] or [19, Theorem A.2]. It is well known [18, p. 495] that each convolutional code $\mathcal{C}$ admits a minimal encoder G. The row degrees $\deg g_i$ of a minimal encoder G are uniquely determined up to ordering and are called the *Forney indices* of the code or of the encoder. It follows that a convolutional code has a constant encoder matrix if and only if the degree is zero. In that case, the code can be regarded as a block code.

The weight of convolutional codewords is defined straightforwardly. We simply extend the ordinary *Hamming weight* $\operatorname{wt}(w_1,\dots,w_n) := \#\{i \mid w_i\neq 0\}$ defined on $\mathbb{F}^n$ to polynomial vectors in the following way. For

$$v=\sum_{j=0}^{N} v^{(j)}D^j \in \mathbb{F}[D]^n$$

where $v^{(j)}\in\mathbb{F}^n$, we put the *weight* of v to be

$$\operatorname{wt}(v)=\sum_{j=0}^{N}\operatorname{wt}(v^{(j)}).$$

Finally, we fix the following notions. For $\delta>0$ we will denote by $e_1,\dots,e_\delta$ the unit vectors in $\mathbb{F}^\delta$. For any matrix $M\in\mathbb{F}^{a\times b}$ we denote by

$$\operatorname{im} M := \{uM \mid u\in\mathbb{F}^a\}$$

and

$$\ker M := \{u\in\mathbb{F}^a \mid uM=0\}$$

the image and kernel, respectively, of the canonical linear mapping $u\mapsto uM$ associated with M. Moreover, for any subset $S\subseteq\mathbb{F}^\ell$ we denote by $\langle S\rangle$ the $\mathbb{F}$-linear subspace generated by S. If $S=\{a_1,\dots,a_t\}$ is finite, we simply write $\langle a_1,\dots,a_t\rangle$ for $\langle S\rangle$. We will also use the notation $\langle a,U\rangle := \langle a\rangle + U$ for any $a\in\mathbb{F}^\ell$ and any linear subspace $U\subseteq\mathbb{F}^\ell$.

II. Preliminaries

The controller canonical form of an encoder is a well-known means of describing convolutional codes. Since our paper is completely based on this description, we will first present the definition of the controller canonical form and thereafter discuss some of the basic properties as needed later on. It also allows us to conveniently introduce the two block codes associated with a convolutional code that are crucial for our investigation.

Definition II.1: Let $G\in\mathbb{F}[D]^{k\times n}$ be a basic and minimal matrix with Forney indices

$$\delta_1,\dots,\delta_r > 0 = \delta_{r+1}=\cdots=\delta_k$$

and degree $\delta := \sum_{i=1}^k \delta_i$. Let G have the rows $g_i=\sum_{\nu=0}^{\delta_i} g_{i,\nu}D^\nu$, $i=1,\dots,k$, where $g_{i,\nu}\in\mathbb{F}^n$. For $i=1,\dots,r$ define the matrices

$$A_i=\begin{pmatrix}0 & 1 & & \\ & & \ddots & \\ & & & 1\\ & & & 0\end{pmatrix}\in\mathbb{F}^{\delta_i\times\delta_i}$$

$$B_i=(1\quad 0\quad \cdots\quad 0)\in\mathbb{F}^{\delta_i}$$

$$C_i=\begin{pmatrix}g_{i,1}\\ \vdots\\ g_{i,\delta_i}\end{pmatrix}\in\mathbb{F}^{\delta_i\times n}.$$

Then the *controller canonical form* of G is defined as the matrix quadruple $(A,B,C,E)\in\mathbb{F}^{\delta\times\delta}\times\mathbb{F}^{k\times\delta}\times\mathbb{F}^{\delta\times n}\times\mathbb{F}^{k\times n}$ where

$$A=\begin{pmatrix}A_1 & & \\ & \ddots & \\ & & A_r\end{pmatrix},\quad B=\begin{pmatrix}\bar{B}\\ 0\end{pmatrix}$$

$$\text{with } \bar{B}=\begin{pmatrix}B_1 & & \\ & \ddots & \\ & & B_r\end{pmatrix},$$

$$C=\begin{pmatrix}C_1\\ \vdots\\ C_r\end{pmatrix},\quad E=\begin{pmatrix}g_{1,0}\\ \vdots\\ g_{k,0}\end{pmatrix}=G(0).$$

As is made precise next, the controller canonical form describes the encoding process of the matrix G in form of a state space system.

Remark II.2: It is easily seen [17, Proposition 2.1, Theorem 2.3] that $G(D)=B(D^{-1}I-A)^{-1}C+E$. As a consequence, one has for $u=\sum_{t\geq 0}u_tD^t\in\mathbb{F}[D]^k$ and $v=\sum_{t\geq 0}v_tD^t\in\mathbb{F}[D]^n$

$$v=uG \Longleftrightarrow \left.\begin{cases}x_{t+1}=x_tA+u_tB\\ v_t=x_tC+u_tE\end{cases}\text{for all } t\geq 0\right\}$$

where $x_0 = 0$. We call $\mathbb{F}^\delta$ the *state space* of the encoder G (or of the controller canonical form) and x_t the state of the encoder at time t.

From now on we will assume our data to be as follows.

General Assumption II.3: Let $\mathcal{C} \subseteq \mathbb{F}[D]^n$ be an (n, k, δ) code with minimal encoder matrix $G \in \mathbb{F}[D]^{k\times n}$. Furthermore, assume that the Forney indices of $\mathcal{C}$ are given by $\delta_1, \dots, \delta_r > 0 = \delta_{r+1} = \cdots = \delta_k$ and let (A, B, C, E) be the corresponding controller canonical form.

The two index sets

$$\mathcal{I} := \left\{1, 1+\delta_1, 1+\delta_1+\delta_2, \dots, 1+\sum_{i=1}^{r-1}\delta_i\right\}$$
$$\mathcal{J} := \left\{\delta_1, \delta_1+\delta_2, \dots, \sum_{j=1}^{r}\delta_j = \delta\right\} \qquad \text{(II.1)}$$

will be helpful in the sequel. One easily derives the following properties.

Remark II.4: One has $AB^t = 0$ and $BB^t = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Furthermore, $\operatorname{im} B = \langle e_i \mid i \in \mathcal{I}\rangle$ and

$$\ker B = \operatorname{im}(0_{(k-r)\times r}, I_{k-r}) \subseteq \mathbb{F}^k.$$

Finally

$$(B^tB)_{i,j} = \begin{cases} 1, & \text{if } i = j \in \mathcal{I} \\ 0, & \text{else} \end{cases}$$
$$(A^tA)_{i,j} = \begin{cases} 1, & \text{if } i = j \notin \mathcal{I} \\ 0, & \text{else} \end{cases}$$
$$(AA^t)_{i,j} = \begin{cases} 1, & \text{if } i = j \notin \mathcal{J} \\ 0, & \text{else.} \end{cases}$$

As a consequence, $A^tA + B^tB = I_\delta$.

The following two block codes will play a crucial role throughout the paper.

Definition II.5: For $\mathcal{C}$ as above we define $\mathcal{C}_{\text{const}} := \mathcal{C} \cap \mathbb{F}^n$ as the block code consisting of the constant codewords in $\mathcal{C}$. Moreover, let $\overline{\mathcal{C}} := \operatorname{im}\binom{C}{E} \subseteq \mathbb{F}^n$ and put $\hat{r} \in \{0, \dots, n-k\}$ such that $\dim\overline{\mathcal{C}} = k + \hat{r}$.

The following properties of these codes are easily seen from the controller canonical form.

Remark II.6:

1) Suppose the encoder matrix G is as in Definition II.1. Then
$$\overline{\mathcal{C}} = \operatorname{im}\begin{pmatrix} C \\ E \end{pmatrix} = \langle g_{i,\nu} \mid i = 1, \dots, k, \nu = 0, \dots, \delta_i\rangle.$$

Recalling that two different encoders of $\mathcal{C}$ differ only by a left unimodular transformation, it follows immediately that the block code $\overline{\mathcal{C}}$ does not depend on the choice of the encoder G but rather is an invariant of the code $\mathcal{C}$. Since $\operatorname{rank} E = k$ it is clear that the dimension of $\overline{\mathcal{C}}$ is indeed at least k.

2) One has $\dim\mathcal{C}_{\text{const}} = k - r$ and precisely, with the notation from 1)

$$\begin{aligned}\mathcal{C}_{\text{const}} &= \langle g_i \mid i = r+1, \dots, k\rangle \\ &= (\ker B)E := \{uE \mid u \in \ker B\}.\end{aligned} \qquad \text{(II.2)}$$

Consequently, $\mathcal{C}_{\text{const}} \subseteq \overline{\mathcal{C}}$ and $\operatorname{im} E = \operatorname{im} B^tE \oplus \mathcal{C}_{\text{const}}$.

Let us now turn to dual codes. In this paper, we will use the module-theoretic dual of convolutional codes. Precisely, consider the canonical $\mathbb{F}[D]$-bilinear form

$$\begin{aligned}\beta: \quad \mathbb{F}[D]^n \times \mathbb{F}[D]^n &\longrightarrow \mathbb{F}[D] \\ ((a_1, \dots, a_n), (b_1, \dots, b_n)) &\longmapsto \sum_{j=1}^{n} a_jb_j.\end{aligned}$$

Then the *dual code* of $\mathcal{C}$ is defined as

$$\hat{\mathcal{C}} := \{w \in \mathbb{F}[D]^n \mid \beta(w, v) = 0 \text{ for all } v \in \mathcal{C}\}. \qquad \text{(II.3)}$$

In the sequel, we will also let β denote the canonical bilinear form on $\mathbb{F}^\ell$ for any $\ell \in \mathbb{N}$. In that case we will use the notation

$$U^\perp := \{v \in \mathbb{F}^\ell \mid \beta(v, u) = 0 \text{ for all } u \in U\} \subseteq \mathbb{F}^\ell$$

for the orthogonal of a subspace $U \subseteq \mathbb{F}^\ell$. The different notation $\hat{\mathcal{C}}$ versus $U^\perp$ for the dual of a convolutional code $\mathcal{C}$ versus a block code U is simply to avoid cumbersome notation later. Before proceeding with the theory we wish to comment on our notion of duality.

Remark II.7: In the literature on convolutional coding theory several notions of duality appear. The module theoretic dual as defined in (II.3) above has been considered in, for instance, [20], [7], [21], [19], and [10]. On the other hand, if codes are regarded as subspaces of the sequence space $(\mathbb{F}^n)^{\mathbb{Z}}$, duality is usually based on the $\mathbb{F}$-bilinear form $\mathbb{F}((D))^n \times \mathbb{F}((D))^n \longrightarrow \mathbb{F}$, where

$$\left(\sum_{j\geq L} v_jD^j, \sum_{j\geq L} w_jD^j\right) \longmapsto \sum_{j\geq L} v_jw_j^t.$$

This is being investigated in, for instance, [22] as well as in [23] and [24], where for the latter two articles one has to use the canonical identification of $\mathbb{F}^n$ with its Pontryagin dual (its character group). Furthermore, in the literature regarding convolutional codes as trellis codes, duality is defined via the dual branch group, see, for instance, [25], or [23], where in the latter the branch group is named the space of local constraints. In [23, Theorem 7.3, Corollary 8.2] it is shown that for group codes over finite index sets duality via the branch group is identical to sequence space duality. We strongly believe, supported by plenty of examples, that the same is true for convolutional codes (having an infinite index set), but that needs to be worked out in detail. A crucial point to be noted in this context is that, due to minimality and basicness of the encoder, the associated output trellis is minimal, see [26, Theorems 5.3, 5.5] and [27], which in turn yields that the state space $\mathbb{F}^\delta$ in our Remark II.2 is isomorphic to the canonical state space in [25]. A third notion of duality arises for quantum error-correcting convolutional codes defined over $\mathbb{F}_4$, see [28]. In that case, duality is based on the Hermitian inner product on $\mathbb{F}_4^n$ and extended to sequences as above. Fortunately, all these notions of duality are closely related to module-theoretic duality. Indeed, one can show straightforwardly that for a given convolutional code the module-theoretic dual and the sequence space dual are mutually reciprocal codes [9, Theorem 2.64]. Considered as sequence spaces reciprocity simply means a reversion of the time axis, whereas algebraically, this amounts to the transformation $G(D) \longmapsto \mathrm{diag}(D^{\nu_1},\ldots,D^{\nu_l})G(D^{-1})$ for a minimal basic encoder matrix G with Forney indices $\nu_1,\ldots,\nu_l$. The latter has some strong consequences for the considerations of this paper. Indeed, it is not hard to see how the MacWilliams Identity as conjectured and partly proven in Section V translates to an identity for sequence space duality. Similarly, including the conjugation on $\mathbb{F}_4$, the identity can also be formulated for codes over $\mathbb{F}_4$ with duality based on the Hermitian inner product (and likewise over any finite field admitting an involution). The details will be shown in a forthcoming paper.

Let us now return to module theoretic duality as defined (II.3). It is well known [11, Theorem 7.1] that

$$\mathcal{C} \text{ is an } (n,k,\delta) \text{ code } \Longrightarrow \hat{\mathcal{C}} \text{ is an } (n,n-k,\delta) \text{ code.} \qquad \text{(II.4)}$$

The two block codes from Definition II.5 and the corresponding objects $C_{\hat{\mathcal{C}}}$ and $\hat{\mathcal{C}}_{\mathrm{const}}$ for the dual code $\hat{\mathcal{C}}$ behave as follows under duality.

Proposition II.8: One has $(\mathcal{C}_C)^\perp = \hat{\mathcal{C}}_{\mathrm{const}}$. As a consequence, $\hat{\mathcal{C}}$ has exactly $n-k-\hat{r}$ zero Forney indices and $\hat{r}$ nonzero Forney indices. Moreover, $\dim C_{\hat{\mathcal{C}}} = n-k+r$.

Proof: Using the notation and statement of Remark II.6, 1) we obtain

$$\begin{aligned} c\in(\mathcal{C}_C)^\perp &\Longleftrightarrow \beta(c,g_{i,\nu})=0 \text{ for all } i=1,\ldots,k,\nu=0,\ldots,\delta_i\\ &\Longleftrightarrow \beta(c,g_i)=0 \text{ for all } i=1,\ldots,k\\ &\Longleftrightarrow c\in\hat{\mathcal{C}}\cap\mathbb{F}^n=\hat{\mathcal{C}}_{\mathrm{const}}\end{aligned}$$

where the second equivalence uses the fact that c is a constant vector. The consequences are clear from the definition of r and $\hat{r}$. □

Example II.9: Let $q=2, n=5, k=2$, and $\mathcal{C}\subseteq\mathbb{F}_2[D]^5$ be the code generated by the basic and minimal encoder

$$G=\begin{pmatrix}1+D+D^3 & D^2 & D^2 & 1 & D\\ 1 & 1 & 0 & 1 & 0\end{pmatrix}.$$

Thus, $\delta=3,\delta_1=3,\delta_2=0$, and $r=1$. The associated controller canonical form is

$$A=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix},\qquad B=\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix},$$
$$C=\begin{pmatrix}1&0&0&0&1\\0&1&1&0&0\\1&0&0&0&0\end{pmatrix},\quad E=\begin{pmatrix}1&0&0&1&0\\1&1&0&1&0\end{pmatrix}.$$

Using (II.2) we obtain

$$\mathcal{C}_{\mathrm{const}}=\mathrm{im}\,(1\ \ 1\ \ 0\ \ 1\ \ 0)$$

while $\mathcal{C}=\mathbb{F}_2^5$. As a consequence, $\hat{r}=3$. It can easily be checked that the dual code $\hat{\mathcal{C}}$ is generated by the basic and minimal matrix

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

Indeed, it has $\hat{r}=3$ nonzero Forney indices as stated in Proposition II.8. The controller canonical form is given by

$$\hat{A}=0,\qquad \hat{B}=I_3,$$
$$\hat{C}=\begin{pmatrix}0&1&0&1&0\\0&1&1&1&0\\0&0&0&0&1\end{pmatrix},$$
$$\hat{E}=\begin{pmatrix}1&0&0&1&0\\0&0&0&0&1\\0&0&1&0&0\end{pmatrix}.$$

Moreover

$$C_{\hat{\mathcal{C}}}=\mathrm{im}\begin{pmatrix}0&0&1&0&0\\0&0&0&0&1\\0&1&0&1&0\\1&0&0&1&0\end{pmatrix},\quad \hat{\mathcal{C}}_{\mathrm{const}}=\{0\}.$$

This is indeed in compliance with Proposition II.8 since $(\mathcal{C}_C)^\perp=\hat{\mathcal{C}}_{\mathrm{const}}$ and $(C_{\hat{\mathcal{C}}})^\perp=\mathcal{C}_{\mathrm{const}}$.

Let us now return to the general case. Block code theory allows us to apply the MacWilliams transformation to the block codes in Proposition II.8. Before doing so, it will be useful to define the weight enumerator for arbitrary affine sets in $\mathbb{F}^n$ as it will be needed in the following sections. Recall the Hamming weight $\mathrm{wt}(a)$ for $a\in\mathbb{F}^n$.

Definition II.10: Let $\mathbb{C}[W]_{\le n}$ denote the vector space of polynomials over $\mathbb{C}$ of degree at most n. For any affine subspace $S\subseteq\mathbb{F}^n$, we define the *weight enumerator* of S to be the polynomial $\mathrm{we}(S):=\sum_{j=0}^n\alpha_jW^j\in\mathbb{C}[W]_{\le n}$, where $\alpha_j:=\#\{a\in S\mid \mathrm{wt}(a)=j\}$. We also put $\mathrm{we}(\emptyset)=0$.

Recall that the classical MacWilliams Identity for block codes states that for k-dimensional codes $\mathcal{C}\subseteq\mathbb{F}^n=\mathbb{F}_q^n$ one has

$$\mathrm{we}(\mathcal{C}^\perp)=q^{-k}\boldsymbol{H}(\mathrm{we}(\mathcal{C}_C)) \qquad \text{(II.5)}$$

where the MacWilliams transformation $\boldsymbol{H}$; $\mathbb{C}[W]_{\le n}\to\mathbb{C}[W]_{\le n}$ is defined as

$$\boldsymbol{H}(f)(W):=(1+(q-1)W)^nf\left(\tfrac{1-W}{1+(q-1)W}\right). \qquad \text{(II.6)}$$

Observe that the mapping $\boldsymbol{H}$ is $\mathbb{C}$-linear and satisfies $\boldsymbol{H}^2(f)=q^nf$. It should be kept in mind that $\boldsymbol{H}$ depends on the parameters n and q. Since throughout this paper these parameters will be fixed we do not indicate them explicitly.

Let us now return to convolutional codes. Using (II.5) and Proposition II.8 one immediately obtains the following.

Corollary II.11: $q^{k+\hat{r}}\mathrm{we}(\hat{\mathcal{C}}_{\mathrm{const}})=\boldsymbol{H}(\mathrm{we}(\mathcal{C}_C))$.

III. The Adjacency Matrix of a Code

The (weight) adjacency matrix as defined next has been introduced in [11] and studied in detail in [17]. The aim of this section is to survey the structure and redundancies of the adjacency matrix for a given convolutional code. Let the data be as in (I.1) and General Assumption II.3. Recall from Remark II.2 that the controller canonical form leads to a state-space description of the encoding process where the input is given by the coefficients of the message stream while the output is the sequence of codeword coefficients. The following matrix collects for each possible pair of states $(X,Y)\in\mathbb{F}^\delta\times\mathbb{F}^\delta$ the information whether via a suitable input u a transition from X to Y is possible, i.e., whether $Y=XA+uB$ for some u, and if so, collects the weights of all associated outputs $v=XC+uE$.

Difinition III.1: Define $\mathcal{F}:=\mathbb{F}^\delta\times\mathbb{F}^\delta$. The *(weight) adjacency matrix* $\Lambda(G)=(\lambda_{X,Y})\in\mathbb{C}[W]^{q^\delta\times q^\delta}$ is defined to be the matrix indexed by $(X,Y)\in\mathcal{F}$ with the entries

$$\lambda_{X,Y}:=\mathrm{we}(\{XC+uE\,|\,u\in\mathbb{F}^k: Y=XA+uB\})$$

in $\mathbb{C}[W]_{\leq n}$. A pair of states $(X,Y)\in\mathcal{F}$ is called *connected* if $\lambda_{X,Y}\neq 0$, else it is called *disconnected*. The set of all connected state pairs is denoted by $\Delta\subseteq\mathcal{F}$.

Observe that in the case $\delta=0$, the matrices A,B,C do not exist while $E=G$. As a consequence, $\Lambda=\lambda_{0,0}=\mathrm{we}(\mathcal{C}_C)$ is the ordinary weight enumerator of the block code $\mathcal{C}=\{uG\,|\,u\in\mathbb{F}^k\}\subseteq\mathbb{F}^n$.

Example III.2: Let the data be as in Example II.9. In order to explicitly display the adjacency matrices corresponding to G and $\hat{G}$, we need to fix an ordering on the state space $\mathbb{F}_2^3$. Let us choose the lexicographic ordering, that is, we will order the row and column indices according to

$$\begin{aligned}&(0,0,0),(0,0,1),(0,1,0),(0,1,1),\\ &\qquad(1,0,0),(1,0,1),(1,1,0),(1,1,1).\end{aligned}\tag{III.1}$$

Then it is lengthy, but straightforward to see that the matrix $\Lambda(G)$, displayed at the bottom of the page, is the associated adjacency matrix. For instance, in order to compute the entry in the fourth row and second column, put $X:=(0,1,1),Y:=(0,0,1)$. Using the controller canonical form as given in Example II.9 , one observes that $XA+uB=Y$ if and only if $u\in\{(0,0),(0,1)\}$ and thus

$$\begin{aligned}\lambda_{X,Y}&=\mathrm{we}\{XC+uE|u\in\{(0,0),(0,1)\}\}\\ &=\mathrm{we}\{(1,1,1,0,0),(0,0,1,1,0)\}=W^2+W^3.\end{aligned}$$

Likewise, we obtain for the dual code

$$\Lambda(\hat{G})=\begin{pmatrix}1&W&W&W^2&W^2&W^3&W^3&W^4\\ W&W^2&1&W&W^3&W^4&W^2&W^3\\ W^3&W^2&W^4&W^3&W^3&W^2&W^4&W^3\\ W^4&W^3&W^3&W^2&W^4&W^3&W^3&W^2\\ W^2&W^3&W^3&W^4&W^2&W^3&W^3&W^4\\ W^3&W^4&W^2&W^3&W^3&W^4&W^2&W^3\\ W&1&W^2&W&W^3&W^2&W^4&W^3\\ W^2&W&W&1&W^4&W^3&W^3&W^2\end{pmatrix}$$

Later, in Theorem V.8, we will see that these two adjacency matrices determine each other in form of a generalized MacWilliams Identity.

Remark III.3: The adjacency matrix contains very detailed information about the code. First, it is well-known that the classical path weight enumerator of a convolutional code [19, p. 154] can be computed from the adjacency matrix, for details see in [11], [17, Theorem 3.8], and [9, Sec. 3.10]. Second, at the end of Section 3 in [17] it has been outlined that the extended row distances [29] as well as the active burst distances [30] can be recovered from the adjacency matrix. As explained in [29], [30] these parameters are closely related to the error-correcting performance of the code.

It is clear from Definition III.1 that the adjacency matrix depends on the chosen encoder G. This dependence, however, can nicely be described. Since we will make intensive use of the notation later on, we introduce the following.

Definition III.4: For any matrix $P\in\mathrm{GL}_\delta(\mathbb{F})$, define $\mathcal{P}(P)\in\mathrm{GL}_{q^\delta}(\mathbb{C})$ by $\mathcal{P}(P)_{X,Y}=1$ if $Y=XP$ and $\mathcal{P}(P)_{X,Y}=0$ else.
Furthermore, let $\Pi:=\{\mathcal{P}(P)\,|\,P\in\mathrm{GL}_\delta(\mathbb{F})\}$ denote the subgroup of all such permutation matrices.

By definition, the matrix $\mathcal{P}(P)$ corresponds to the permutation on the set $\mathbb{F}^\delta$ induced by the isomorphism P. Notice that $\mathcal{P}$ is an isomorphism of groups and basically is the canonical faithful permutation representation of the group $\mathrm{GL}_\delta(\mathbb{F})$. Obviously, we have for any $\Lambda\in\mathbb{C}[W]^{q^\delta\times q^\delta}$ and any $\mathcal{P}:=\mathcal{P}(P)\in\Pi$ the identity

$$(\mathcal{P}\Lambda\mathcal{P}^{-1})_{X,Y}=\Lambda_{XP,YP},\quad\text{for all }(X,Y)\in\mathcal{F}.\tag{III.2}$$

Now we can collect the following facts about the adjacency matrix.

Remark III.5:

a) Using the obvious fact $\mathrm{wt}(\alpha v)=\mathrm{wt}(v)$ for any $\alpha\in\mathbb{F}^*$ and $v\in\mathbb{F}^n$ one immediately has $\lambda_{X,Y}=\lambda_{\alpha X,\alpha Y}$ for all $\alpha\in\mathbb{F}^*$. Hence, $\Lambda(G)$ is invariant under conjugation with permutation matrices that are induced by scalar multiplication on $\mathbb{F}^\delta$, i.e., under conjugation with matrices $\mathcal{P}(P)$ where $P=\alpha I$ for some $\alpha\in\mathbb{F}^*$.
b) In [17, Theorem 4.1] it has been established that if $G_1,G_2\in\mathbb{F}[D]^{k\times n}$ are two minimal encoders of $\mathcal{C}$ then $\Lambda(G_1)=\mathcal{P}\Lambda(G_2)\mathcal{P}^{-1}$ for some $\mathcal{P}\in\Pi$. Hence, the equivalence class of $\Lambda(G)$ modulo conjugation by Π, where G is any minimal encoder, forms an invariant of

$$\Lambda(G)=\begin{pmatrix}1+W^3&0&0&0&W+W^2&0&0&0\\ W+W^2&0&0&0&W+W^2&0&0&0\\ 0&W^2+W^3&0&0&0&W+W^4&0&0\\ 0&W^2+W^3&0&0&0&W^2+W^3&0&0\\ 0&0&W^2+W^3&0&0&0&W^2+W^3&0\\ 0&0&W+W^4&0&0&0&W^2+W^3&0\\ 0&0&0&W^3+W^4&0&0&0&W^3+W^4\\ 0&0&0&W^3+W^4&0&0&0&W^2+W^5\end{pmatrix}.$$

the code. It is called the *generalized adjacency matrix* of $\mathcal{C}$.

c) Combining part b) and part a) we see that the equivalence class of $\Lambda(G)$ is already fully obtained by conjugating $\Lambda(G)$ with matrices $\mathcal{P}(P)$ where P is in the projective linear group $\mathrm{GL}_\delta(\mathbb{F})/\{\alpha I \mid \alpha \in \mathbb{F}^*\}$. This reduces the computational effort when computing examples.

Let us return to Definition III.1. Notice that $(X,Y) \in \mathcal{F}$ is connected if and only if there exists some $u \in \mathbb{F}^k$ such that $(X,Y) = (X, XA + uB)$. Using rank $B = r$, we obtain the following.

Proposition III.6: $\Delta = \mathrm{im}\begin{pmatrix} I & A \\ 0 & B \end{pmatrix}$ is an $\mathbb{F}$-vector space of dimension $\delta + r$.

Later on we will also need the orthogonal of Δ in $\mathcal{F}$. With the help of Remark II.4, it can easily be calculated and is given as follows.

Lemma III.7: The orthogonal space of Δ is given by $\Delta^\perp = \{(X, -XA) \mid X = (X_1, \ldots, X_\delta) \in \mathbb{F}^\delta, X_j = 0 \text{ for } j \in \mathcal{J}\}$.

The next lemma will show that the nontrivial entries $\lambda_{X,Y}$ of the adjacency matrix can be described as weight enumerators of certain cosets of the block code $\mathcal{C}_{\mathrm{const}}$. More precisely, we will relate them to the $\mathbb{F}$-vector space homomorphism

$$\varphi : \mathcal{F} \longrightarrow \mathbb{F}^n, \quad (X,Y) \longmapsto XC + YB^tE. \tag{III.3}$$

Recall the notation $\langle a, U\rangle$ as introduced at the end of Section I.

Lemma III.8: For any state pair $(X,Y) \in \Delta$ we have

$$\lambda_{X,Y} = \mathrm{we}(\varphi(X,Y) + \mathcal{C}_{\mathrm{const}})$$

as well as the equation shown at the bottom of the page.

Proof: First notice that for any $(X,Y) \in \Delta$, the set $\{u \in \mathbb{F}^k \mid Y - XA = uB\}$ is nonempty. Right-multiplying the defining equation of this set with B^t, we get upon use of Remark II.4 that $YB^t = uBB^t$, which says that the first r entries of u are completely determined by Y. This shows

$$\{u \in \mathbb{F}^k \mid Y - XA = uB\} \subseteq YB^t + \mathrm{im}(0, I_{k-r}).$$

From $\mathrm{im}(0, I_{k-r}) = \ker B$, see Remark II.4, we conclude that these two affine subspaces coincide. Hence, using Remark II.6, 2), we obtain

$$\begin{aligned}\lambda_{X,Y} &= \mathrm{we}\left(XC + (YB^t + \ker B)E\right)\\ &= \mathrm{we}\left(\varphi(X,Y) + (\ker B)E\right)\\ &= \mathrm{we}\left(\varphi(X,Y) + \mathcal{C}_{\mathrm{const}}\right).\end{aligned}$$

This shows the first part of the lemma. If $\varphi(X,Y) \in \mathcal{C}_{\mathrm{const}}$, we immediately conclude $\lambda_{X,Y} = \mathrm{we}(\mathcal{C}_{\mathrm{const}})$. Otherwise, we have $\lambda_{X,Y} = \mathrm{we}(\varphi(X,Y) + \mathcal{C}_{\mathrm{const}}) = \mathrm{we}(\alpha(\varphi(X,Y) + \mathcal{C}_{\mathrm{const}})) = \mathrm{we}(\alpha\varphi(X,Y) + \mathcal{C}_{\mathrm{const}})$ for all $\alpha \in \mathbb{F}^*$. Moreover

$$\langle \varphi(X,Y), \mathcal{C}_{\mathrm{const}}\rangle = \bigcup_{\alpha \in \mathbb{F}} (\alpha\varphi(X,Y) + \mathcal{C}_{\mathrm{const}})$$

where due to $\varphi(X,Y) \notin \mathcal{C}_{\mathrm{const}}$ this union is disjoint. From this the last assertion can be deduced. □

The lemma shows that the mapping φ and the block code $\mathcal{C}_{\mathrm{const}}$ along with the knowledge of Δ fully determine $\Lambda(G)$. Moreover, to find out how many state pairs $(X,Y) \in \Delta$ are mapped to $\mathcal{C}_{\mathrm{const}}$, we will slightly modify the mapping φ.

Lemma III.9: The homomorphism

$$\Phi : \Delta \longrightarrow \mathcal{C}_C/\mathcal{C}_{\mathrm{const}}, \quad (X,Y) \longmapsto \varphi(X,Y) + \mathcal{C}_{\mathrm{const}}$$

is well-defined, surjective, and satisfies

a) $\ker\Phi = \{(X,Y) \in \Delta \mid \varphi(X,Y) \in \mathcal{C}_{\mathrm{const}}\}$;

b) $\dim\ker\Phi = \delta - \hat{r}$, where $\hat{r}$ is as in Definition II.5.

Proof: The well-definedness of Φ simply follows from $\mathrm{im}\varphi \subseteq \mathcal{C}$. As for the surjectivity notice that any row of $\begin{pmatrix} C \\ E\end{pmatrix}$ that is not in $\mathcal{C}_{\mathrm{const}}$ is a row of the matrix $\begin{pmatrix} C \\ BB^tE\end{pmatrix}$, see also Remark II.6(2). Moreover, by Remark II.4 we have

$$\mathrm{im}\begin{pmatrix} C \\ BB^tE\end{pmatrix} = \mathrm{im}\begin{pmatrix} I & A \\ 0 & B\end{pmatrix}\begin{pmatrix} C \\ B^tE\end{pmatrix} = \varphi(\Delta)$$

where the latter follows from the definition of the mapping φ along with Proposition III.6. All this implies the surjectivity of Φ. Now part a) is trivial. The surjectivity together with $\dim\Delta = \delta + r$ yields part b) since $\dim\mathcal{C} = k + \hat{r}$ while $\dim\mathcal{C}_{\mathrm{const}} = k - r$. □

Let us illustrate the results so far by the previous example.

Example III.10: Consider again the data from Example II.9 and III.2. We can observe the following properties of the two adjacency matrices.

1) The matrix $\Lambda(G)$ has exactly $2^4 = 16$ nonzero entries, while $\Lambda(\hat{G})$ has exactly $2^6 = 64$ nonzero entries. This is in compliance with Proposition III.6 applied to $\mathcal{C}$ as well as $\hat{\mathcal{C}}$.
2) Each nonzero entry of $\Lambda(G)$ is the sum of two monomials, while each entry of $\Lambda(\hat{G})$ is a monomial. This also follows from the first part of Lemma III.8 since $\#\mathcal{C}_{\mathrm{const}} = 2$ while $\hat{\mathcal{C}}_{\mathrm{const}} = \{0\}$.
3) There are four entries in $\Lambda(\hat{G})$ that are equal to 1. This also follows from application of Lemma III.9 and Lemma III.8 to the dual code: we obtain $2^{\delta - r} = 4$ times the case $\hat{\lambda}_{X,Y} = \mathrm{we}(\hat{\mathcal{C}}_{\mathrm{const}}) = 1$ while the second case appearing in Lemma III.8, being the difference of the weight enumerators of two block codes, never contains the monomial $1 = W^0$. Along the same line of arguments one can also explain that $\lambda_{0,0}$ is the only entry of $\Lambda(G)$ containing the monomial $1 = W^0$.

$$\lambda_{X,Y} = \begin{cases} \mathrm{we}(\mathcal{C}_{\mathrm{const}}), & \text{if } \varphi(X,Y) \in \mathcal{C}_{\mathrm{const}} \\ \frac{1}{q-1}\left(\mathrm{we}(\langle \varphi(X,Y), \mathcal{C}_{\mathrm{const}}\rangle) - \mathrm{we}(\mathcal{C}_{\mathrm{const}})\right), & \text{else.}\end{cases}$$

As a consequence of Lemma III.9, one has

$$\varphi(\Delta)+\mathcal{C}_{\rm const}=\mathcal{C}. \tag{III.4}$$

We are now prepared to clarify some more redundancies in the adjacency matrix of $\mathcal{C}$.

Proposition III.11: Let $\Delta^*\subseteq\Delta$ be any subspace such that $\Delta=\Delta^*\oplus\ker\Phi$. Moreover, define $\Delta^-:=\langle(0,e_i)\,|\,i\notin\mathcal{I}\rangle\subseteq\mathcal{F}$. Then we get the following.

a) $\Delta\oplus\Delta^-=\mathcal{F}$, hence, $\Delta^*\oplus\ker\Phi\oplus\Delta^-=\mathcal{F}$.
b) For any state pairs $(X,Y)\in\Delta^-$ and $(X',Y')\in\Delta$ one has $\lambda_{X+X',Y+Y'}=0$ if and only if $(X,Y)\neq 0$.
c) For any state pairs $(X,Y)\in\Delta^*$ and $(X',Y')\in\ker\Phi$ one has $\lambda_{X+X',Y+Y'}=\lambda_{X,Y}$.

Proof: Part a) $\Delta\cap\Delta^-=\{0\}$ follows from $e_i\notin\operatorname{im}B$ for $i\notin\mathcal{I}$. The rest is clear since $\dim\Delta^-=\delta-r=2\delta-\dim\Delta$. part b) is obvious from the first direct sum in part a) and the definition of Δ. As for part c), notice that by linearity and Lemma III.9, part a) $\varphi(X,Y)-\varphi(X+X',Y+Y')\in\mathcal{C}_{\rm const}$. Hence, $\varphi(X,Y)+\mathcal{C}_{\rm const}=\varphi(X+X',Y+Y')+\mathcal{C}_{\rm const}$ and the result follows from Lemma III.8. □

It is worth noting that the converse of Proposition III.11, part c), that is, $[\lambda_{X,Y}=\lambda_{\tilde{X},\tilde{Y}}\Longrightarrow(X,Y)-(\tilde{X},\tilde{Y})\in\ker\Phi]$, is in general not true as different affine sets may well have the same weight enumerator. Moreover, notice that the results above are obviously true for any direct complement of Δ in $\mathcal{F}$. Our particular choice of Δ^- will play an important role due to the following corollary.

Corollary III.12: One has $\varphi|_{\Delta^-}=0$ and the space $\mathcal{C}$ satisfies $\mathcal{C}=\bigcup_{(X,Y)\in\Delta^*}(\varphi(X,Y)+\mathcal{C}_{\rm const})$ with the union being disjoint.

Proof: The first part follows directly from the definition of all objects involved. The inclusion "$\supseteq$" of the second statement is obvious. For the converse, let $XC+uE\in\mathcal{C}$ for some $(X,u)\in\mathbb{F}^{\delta+k}$. Using that $\operatorname{im}E=\operatorname{im}B^tE+\mathcal{C}_{\rm const}$, see Remark II.6, 2), this yields $XC+uE=XC+YB^tE+a$ for some $Y\in\mathbb{F}^\delta$ and $a\in\mathcal{C}_{\rm const}$. Hence, $XC+uE\in\varphi(X,Y)+\mathcal{C}_{\rm const}$, where $(X,Y)\in\mathcal{F}$. Now $\varphi|_{\Delta^-}=0$ and Lemma III.9, part a) imply that without loss of generality $(X,Y)\in\Delta^*$. The disjointness of the union follows from $\Delta^*\cap\ker\Phi=\{0\}$ with the same lemma. □

We will conclude this section by computing the sum over all entries of the adjacency matrix of a convolutional code in order to demonstrate how the terminology developed above facilitates this task. The result will be needed later on for proving Theorem IV.7.

Proposition III.13: The entries of the adjacency matrix satisfy

$$\sum_{(X,Y)\in\Delta^*}\lambda_{X,Y}=\operatorname{we}(C_{\mathcal{C}})$$

as well as

$$\sum_{(X,Y)\in\mathcal{F}}\lambda_{X,Y}=\sum_{(X,Y)\in\Delta}\lambda_{X,Y}=q^{\delta-\hat{r}}\operatorname{we}(\mathcal{C}_c).$$

Proof: Using Lemma III.8 and Corollary III.12, we obtain

$$\begin{aligned}&\sum_{(X,Y)\in\Delta^*}\lambda_{X,Y}\\ &\quad=\sum_{(X,Y)\in\Delta^*}\operatorname{we}(\varphi(X,Y)+\mathcal{C}_{\rm const})\\ &\quad=\operatorname{we}(\mathcal{C}_C).\end{aligned}$$

Next notice that $\sum_{(X,Y)\in\mathcal{F}}\lambda_{X,Y}=\sum_{(X,Y)\in\Delta}\lambda_{X,Y}$ as any disconnected state pair (X,Y) satisfies $\lambda_{X,Y}=0$. Hence with Proposition III.11, part c) and Lemma III.9, part b) we get

$$\begin{aligned}\sum_{(X,Y)\in\Delta}\lambda_{X,Y}&=\sum_{(\bar{X},\bar{Y})\in\ker\Phi}\ \sum_{(X,Y)\in\Delta^*}\lambda_{X+\bar{X},Y+\bar{Y}}\\ &=\sum_{(\bar{X},\bar{Y})\in\ker\Phi}\ \sum_{(X,Y)\in\Delta^*}\lambda_{X,Y}\\ &=\sum_{(\bar{X},\bar{Y})\in\ker\Phi}\operatorname{we}(\mathcal{C}_C)=q^{\delta-\hat{r}}\operatorname{we}(\mathcal{C}_C).\end{aligned}$$

It is straightforward to verify the second result of this proposition for Examples II.9/III.2.

IV. The MacWilliams Matrices

Recall the notation from (I.1) and fix some $\delta\in\mathbb{N}$. In this section, we will define a set of complex matrices that are essential for our transformation formula as discussed in the next section, and we will collect some of their properties. To define the matrices, we will use complex-valued characters on $\mathbb{F}^\delta$, i.e., group homomorphisms $(\mathbb{F}^\delta,+)\longrightarrow(\mathbb{C}^*,\cdot)$. It is a well-known fact [31, Theorem 5.5], that using a fixed primitive pth root of unity $\zeta\in\mathbb{C}$, the character group on $\mathbb{F}^\delta$ is given as $\{\zeta^{\tau(\beta(X,\cdot))}\,|\,X\in\mathbb{F}^\delta\}$, where $\tau:\mathbb{F}\longrightarrow\mathbb{F}_p, a\longmapsto\sum_{i=0}^{s-1}a^{p^i}$ is the usual trace form and β is the canonical bilinear form on $\mathbb{F}^\delta$. It will be convenient to define

$$\theta_X:=\zeta^{\tau(\beta(X,\cdot))}:\mathbb{F}^\delta\longrightarrow\mathbb{C}^*,\quad\text{for all } X\in\mathbb{F}^\delta. \tag{IV.1}$$

For easier reference, we list the following properties.

Remark IV.1:

a) The character θ_X is nontrivial if and only if $X\neq 0$. This follows from the fact that for $X\neq 0$ we have $\#\operatorname{im}\beta(X,\cdot)=q$, while the $\mathbb{F}_p$-linear and surjective mapping τ satisfies $\#\ker\tau=p^{s-1}=\frac{q}{p}$.
b) By standard results on characters [31, Theorem 5.4], we obtain $\sum_{Y\in\mathbb{F}^\delta}\theta_X(Y)=0$ if $X\neq 0$ and $\sum_{Y\in\mathbb{F}^\delta}\theta_0(Y)=q^\delta$.
c) For all $X,Y\in\mathbb{F}^\delta$ and all $P\in\mathrm{GL}_\delta(\mathbb{F})$ we have $\theta_X(Y)=\theta_Y(X)$ and $\theta_{XP}(Y)=\theta_X(YP^t)$.
d) For all $X,Y,Z_1,Z_2\in\mathbb{F}^\delta$ one has $\theta_X(Z_1)\theta_Y(Z_2)=\theta_{(X,Y)}(Z_1,Z_2)$, where the latter is defined on $\mathbb{F}^{2\delta}$ analogously to(IV.1) . More precisely, $\theta_{(X,Y)}(Z_1,Z_2):=\zeta^{\tau(\beta((X,Y),(Z_1,Z_2)))}$, where β also denotes the canonical bilinear form on $\mathbb{F}^{2\delta}$,

Definition IV.2: Let $\zeta \in \mathbb{C}^*$ be a fixed primitive pth root of unity. For $P \in \mathrm{GL}_\delta(\mathbb{F})$ we define the P-*MacWilliams matrix* as

$$\mathcal{H}(P) := q^{\frac{-\delta}{2}}(\theta_{XP}(Y))_{(X,Y)\in\mathcal{F}} \in \mathbb{C}^{q^\delta \times q^\delta}.$$

For simplicity, we also put $\mathcal{H} := \mathcal{H}(I)$. For $\delta = 0$ we simply have $\mathcal{H} = 1$.

Remark IV.3: Notice that the MacWilliams matrices depend on δ. Since this parameter will be fixed throughout our paper (except for the examples and Remark IV.6), we will not explicitly denote this dependence. Moreover, the matrices depend on the choice of the primitive root ζ. This dependence, however, can easily be described. Suppose ζ_1 and ζ_2 are two primitive pth roots of unity and let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the corresponding I-MacWilliams matrices. Then $\zeta_1^d = \zeta_2$ for some $0 < d < p$ and, using the $\mathbb{F}_p$-linearity of τ, it is easy to check that $\mathcal{H}_2 = \mathcal{P}(dI)\mathcal{H}_1 = \mathcal{H}_1\mathcal{P}(d^{-1}I)$. Making use of Remark III.5, part a) this results in $\mathcal{H}_2\Lambda^t\mathcal{H}_2^{-1} = \mathcal{H}_1\Lambda^t\mathcal{H}_1^{-1}$ and $\mathcal{H}_2\Lambda\mathcal{H}_2 = \mathcal{H}_1\Lambda\mathcal{H}_1$. Since all later expressions will be one of these forms, our results later on do not depend on the choice of ζ.

Obviously, the matrix $\mathcal{H}$ is symmetric. Moreover, all MacWilliams matrices are invertible since the q^δ different characters are linearly independent in the vector space of $\mathbb{C}$-valued functions on $\mathbb{F}^\delta$. However, the inverse of these matrices can easily be calculated. Recall the matrices $\mathcal{P}(P)$ from Definition III.4.

Lemma IV.4: One has $\mathcal{H}^2 = \mathcal{P}(-I)$ and, hence, $\mathcal{H}^4 = I$. Furthermore

$$\mathcal{H}(P) = \mathcal{P}(P)\mathcal{H} = \mathcal{H}\mathcal{P}((P^t)^{-1}), \quad \text{for all } P \in \mathrm{GL}_\delta(\mathbb{F}).$$

In particular, the inverse of a MacWilliams matrix is a MacWilliams matrix again.

Proof: For the computation of $\mathcal{H}^2$, fix a pair $(X,Y) \in \mathcal{F}$. Then, upon using the rules in Remark IV.1, part b) and part c)

$$\begin{aligned}(\mathcal{H}^2)_{X,Y} &= q^{-\delta}\sum_{Z\in\mathbb{F}^\delta}\theta_X(Z)\theta_Z(Y)\\ &= q^{-\delta}\sum_{Z\in\mathbb{F}^\delta}\theta_{X+Y}(Z)\\ &= \left\{\begin{matrix}1, & \text{if Y=-X}\\ 0, & \text{else}\end{matrix}\right\} = \mathcal{P}(-I)_{X,Y}.\end{aligned}$$

The rest of the lemma can be checked in the same way using again Remark IV.1, part c). □

Example IV.5: Let $p = q = 2$ and $\delta = 3$. Then $\zeta = -1$ and with respect to the lexicographic ordering (III.1) on $\mathbb{F}_2^3$ we obtain

$$\mathcal{H} = \frac{1}{\sqrt{8}}\begin{pmatrix}1&1&1&1&1&1&1&1\\1&-1&1&-1&1&-1&1&-1\\1&1&-1&-1&1&1&-1&-1\\1&-1&-1&1&1&-1&-1&1\\1&1&1&1&-1&-1&-1&-1\\1&-1&1&-1&-1&1&-1&1\\1&1&-1&-1&-1&-1&1&1\\1&-1&-1&1&-1&1&1&-1\end{pmatrix}.$$

Remark IV.6: It should be mentioned that the MacWilliams matrices as presented here appear already in classical block code theory in the context of complete weight enumerators. Given a block code $C \subseteq \mathbb{F}^n$, the complete weight enumerator is defined as

$$\mathrm{cwe}(C) := \sum_{(c_1,\ldots,c_n)\in C}\prod_{i=1}^n X_{c_i} \in \mathbb{C}[X_a \mid a \in \mathbb{F}].$$

Obviously, we obtain the ordinary weight enumerator $\mathrm{we}(C)$ from $\mathrm{cwe}(C)$ by putting $X_0 = 1$ and $X_a = W$ for all $a \in \mathbb{F}^*$. Let now $\delta = 1$ and $\mathcal{H} \in \mathbb{C}^{q\times q}$ be the corresponding MacWilliams matrix. Then $\mathcal{H}$ is the standard matrix interpretation of the $\mathbb{C}$-vector space automorphism $h : \langle X_a \mid a \in \mathbb{F}\rangle_\mathbb{C} \longrightarrow \langle X_a \mid a \in \mathbb{F}\rangle_\mathbb{C}$ defined via

$$h(X_a) = q^{-\frac{1}{2}}\sum_{b\in\mathbb{F}}\zeta^{\tau(ab)}X_b.$$

Extending h to a $\mathbb{C}$-algebra-homomorphism on $\mathbb{C}[X_a \mid a \in \mathbb{F}]$ it is well known [3, Ch. 5.6, Theorem 10] that the complete weight enumerators of a k-dimensional block code $C \subseteq \mathbb{F}_q^n$ and its dual satisfy the MacWilliams Identity $\mathrm{cwe}(C^\perp) = q^{-k+\frac{n}{2}}h(\mathrm{cwe}(C))$. At this point, it is not clear to us why the MacWilliams matrix appears in the seemingly unrelated contexts of complete weight enumerators for block codes and adjacency matrices for convolutional codes.

In the next section, we will investigate a conjecture concerning a MacWilliams Identity Theorem for the adjacency matrices of convolutional codes and their duals. It states that for the data as in General Assumption II.3 and for any $P \in \mathrm{GL}_\delta(\mathbb{F})$ the matrix $q^{-k}\boldsymbol{H}(\mathcal{H}(P)\Lambda(G)^t\mathcal{H}(P)^{-1})$ is a representative of the generalized adjacency matrix of $\hat{\mathcal{C}}$ (in the sense of Remark III.5, part b)), see Conjecture V.2. Using Lemma IV.4 and the fact $\mathcal{H}^t = \mathcal{H}$, one easily observes that $\mathcal{H}\Lambda(G)\mathcal{H} = \mathcal{P}(-I)(\mathcal{H}\Lambda(G)^t\mathcal{H}^{-1})^t$. Therefore, the matrix $\mathcal{H}\Lambda(G)\mathcal{H}$ will be particularly helpful and will be studied first. Let, as usual, the data be as in General Assumption II.3 and Definition III.1 and remember $\hat{r}$ from Proposition II.8. Put

$$\ell_{X,Y} := (\mathcal{H}\Lambda(G)\mathcal{H})_{X,Y} \text{ for } (X,Y) \in \mathcal{F}. \tag{IV.2}$$

The entries $\ell_{X,Y}$ can be described explicitly. In the sequel, we will use for any pair $(X,Y) \in \mathcal{F}$ the short notation $(X,Y)^\perp := \langle(X,Y)\rangle^\perp$ to denote the orthogonal space in $\mathcal{F}$. The following result will be crucial for the MacWilliams Identity Conjecture as studied in the next section.

Theorem IV.7: Let $(X,Y) \in \mathcal{F}$. Then

$$\ell_{X,Y} = \begin{cases}0, & \text{if } (X,Y) \notin (\ker\Phi)^\perp\\ q^{-\hat{r}}\mathrm{we}(\mathcal{C}_C), & \text{if } (X,Y) \in \Delta^\perp\end{cases}$$

whereas for $(X,Y) \in (\ker\Phi)^\perp\backslash\Delta^\perp$

$$\ell_{X,Y} = \frac{1}{q^\delta(q-1)}\left(q\sum_{(Z_1,Z_2)\in(X,Y)^\perp}\lambda_{Z_1,Z_2} - q^{\delta-\hat{r}}\mathrm{we}(\mathcal{C}_C)\right).$$

Furthermore

$$\ell_{X+U,Y+V} = \ell_{X,Y}, \quad \text{for all } (U,V) \in \Delta^\perp.$$

The last statement can be regarded as a counterpart to part c) of Proposition III.11. In fact, both these invariance properties will be needed to derive a correspondence between the matrix $\mathcal{H}\Lambda(G)\mathcal{H}$ and the adjacency matrix of the dual code later on in Section V.

Proof: Fix $(X,Y)\in\mathcal{F}$.

1) We begin with proving the identity

$$q^{\delta}\ell_{X,Y}=\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}-\frac{1}{q-1}\sum_{(Z_1,Z_2)\notin(X,Y)^{\perp}}\lambda_{Z_1,Z_2}. \quad \text{(IV.3)}$$

Using Remark IV.1, part d), we have

$$\begin{aligned}q^{\delta}\ell_{X,Y}&=\sum_{Z_1,Z_2\in\mathbb{F}^{\delta}}\theta_X(Z_1)\lambda_{Z_1,Z_2}\theta_{Z_2}(Y)\\&=\sum_{(Z_1,Z_2)\in\mathcal{F}}\theta_{(X,Y)}(Z_1,Z_2)\lambda_{Z_1,Z_2}.\end{aligned}$$

This proves (IV.3) for $(X,Y)=(0,0)$. Thus, let $(X,Y)\neq(0,0)$. Choose $(V_1,V_2)\in\mathcal{F}$ such that $\mathcal{F}=(X,Y)^{\perp}\oplus\langle(V_1,V_2)\rangle$. This allows, recalling Remark III.5, part a), to further simplify $\ell_{X,Y}$. Indeed

$$\begin{aligned}q^{\delta}\ell_{X,Y}&=\sum_{\alpha\in\mathbb{F}}\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\theta_{(X,Y)}(Z_1+\alpha V_1,Z_2+\alpha V_2)\lambda_{Z_1+\alpha V_1,Z_2+\alpha V_2}\\&=\sum_{\alpha\in\mathbb{F}}\theta_{(X,Y)}(\alpha V_1,\alpha V_2)\times\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1+\alpha V_1,Z_2+\alpha V_2}\\&=\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}+\sum_{\alpha\in\mathbb{F}^*}\theta_{(X,Y)}(\alpha V_1,\alpha V_2)\times\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1+V_1,Z_2+V_2}.\end{aligned}$$

Since $(X,Y)\neq(0,0)$, the character $\alpha\mapsto\theta_{(X,Y)}(\alpha V_1,\alpha V_2)$ is nontrivial on $\mathbb{F}$ and thus

$$\begin{aligned}q^{\delta}\ell_{X,Y}&=\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}-\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1+V_1,Z_2+V_2}\\&=\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}-\frac{1}{q-1}\sum_{(Z_1,Z_2)\notin(X,Y)^{\perp}}\lambda_{Z_1,Z_2}\end{aligned} \quad \text{(IV.4)}$$

where the last identity is again derived from Remark III.5, part a) considering that $(X,Y)^{\perp}$ is a subspace of $\mathcal{F}$. This completes the proof of (IV.3).

2) Now we will prove each case of the second assertion separately. First, let $(X,Y)\notin(\ker\Phi)^{\perp}$. This implies $\ker\Phi\not\subseteq(X,Y)^{\perp}$. Hence, (V_1,V_2) above can be chosen in $\ker\Phi$ and therefore Proposition III.11, part c) along with (IV.4) yields $\ell_{X,Y}=0$.

3) Let $(X,Y)\in\Delta^{\perp}$, which implies $\Delta\subseteq(X,Y)^{\perp}$ and, therefore, $(\mathcal{F}\backslash(X,Y)^{\perp})\cap\Delta=\emptyset$. Using (IV.3) together with the fact that $\lambda_{Z_1,Z_2}=0$ for $(Z_1,Z_2)\notin\Delta$ one gets

$$\begin{aligned}q^{\delta}\ell_{X,Y}&=\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}\cap\Delta}\lambda_{Z_1,Z_2}-\frac{1}{q-1}\sum_{(Z_1,Z_2)\in(\mathcal{F}\backslash(X,Y)^{\perp})\cap\Delta}\lambda_{Z_1,Z_2}\\&=\sum_{(Z_1,Z_2)\in\Delta}\lambda_{Z_1,Z_2}=q^{\delta-\hat{r}}\mathrm{we}(\mathcal{C}_C)\end{aligned}$$

where the last identity follows from Proposition III.13.

4) Finally, for the last case, let $(X,Y)\in(\ker\Phi)^{\perp}\backslash\Delta^{\perp}$. Thus, $\ker\Phi\subseteq(X,Y)^{\perp}$, but $\Delta\not\subseteq(X,Y)^{\perp}$. Again, using (IV.3) together with Proposition III.1, part 3), we obtain

$$\begin{aligned}\ell_{X,Y}&=\frac{1}{q^{\delta}(q-1)}\left((q-1)\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}-\sum_{(Z_1,Z_2)\notin(X,Y)^{\perp}}\lambda_{Z_1,Z_2}\right)\\&=\frac{1}{q^{\delta}(q-1)}\left(q\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}-q^{\delta-\hat{r}}\mathrm{we}(\mathcal{C}_C)\right).\end{aligned}$$

5) It remains to show $\ell_{X+U,Y+V}=\ell_{X,Y}$ for $(U,V)\in\Delta^{\perp}$. Since $\Delta^{\perp}\subseteq(\ker\Phi)^{\perp}$, the statement is obvious in the first two cases of $\ell_{X,Y}$. For the remaining case, notice that in the expression for $\ell_{X,Y}$ we have

$$\begin{aligned}\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}}\lambda_{Z_1,Z_2}&=\sum_{(Z_1,Z_2)\in(X,Y)^{\perp}\cap\Delta}\lambda_{Z_1,Z_2}\\&=\sum_{(Z_1,Z_2)\in(X+U,Y+V)^{\perp}\cap\Delta}\lambda_{Z_1,Z_2}\end{aligned}$$

for any $(U,V)\in\Delta^{\perp}$. This completes the proof. □

At this point, it is possible to derive a formula for the MacWilliams transformation $\boldsymbol{H}$ as defined in (II.6) applied to the entries $\ell_{X,Y}$. It will play a central role in the next section. Recall from (II.3) the notation $\hat{\mathcal{C}}$ for the dual code of $\mathcal{C}$. Notice from (IV.2) that the polynomials $\ell_{X,Y}$ are in $\mathbb{C}[W]_{\leq n}$ so that indeed the mapping $\boldsymbol{H}$ can be applied.

Proposition IV.8: Let $(X,Y)\in\mathcal{F}$. Then

$$q^{-k}\boldsymbol{H}(\ell_{X,Y})=\begin{cases}0, & \text{if } (X,Y)\notin(\ker\Phi)^{\perp}\\ \mathrm{we}(\hat{\mathcal{C}}_{\mathrm{const}}), & \text{if } (X,Y)\in\Delta^{\perp}\end{cases}$$

whereas for $(X,Y)\in(\ker\Phi)^{\perp}\backslash\Delta^{\perp}$ we have

$$q^{-k}\boldsymbol{H}(\ell_{X,Y})=\frac{1}{q-1}\Big(\mathrm{we}\langle\hat{\mathcal{C}}_{\mathrm{const}},c(X,Y)\rangle)-\mathrm{we}(\hat{\mathcal{C}}_{\mathrm{const}})\Big)$$

with $c(X,Y)$ being any element in

$$[\varphi((X,Y)^{\perp}\cap\Delta^*)+\mathcal{C}_{\mathrm{const}}]^{\perp}\backslash\hat{\mathcal{C}}_{\mathrm{const}}.$$

Proof: Use the form of $\ell_{X,Y}$ as given in Theorem IV.7. The first case is immediate as $\boldsymbol{H}(0) = 0$. The second case is exactly Corollary II.11. The third case requires more work. Thus, let $(X,Y) \in (\ker\Phi)^\perp \backslash \Delta^\perp$, hence, $\ker\Phi \subseteq (X,Y)^\perp$, but $\Delta \nsubseteq (X,Y)^\perp$. As a consequence, $(X,Y)^\perp \cap \Delta$ is a hyperplane of Δ and $\ker\Phi$ is contained in $(X,Y)^\perp \cap \Delta$. Using the direct complement Δ^* of $\ker\Phi$ in Δ, as introduced in Proposition III.11, we obtain

$$(X,Y)^\perp \cap \Delta = ((X,Y)^\perp \cap \Delta^*) \oplus \ker\Phi$$

and $(X,Y)^\perp \cap \Delta^*$ is a hyperplane in Δ^*. With the help of Proposition III.11, part c) and Lemma III.9, part b) we get

$$\begin{aligned}\sum_{(Z_1,Z_2)\in(X,Y)^\perp} \lambda_{Z_1,Z_2} &= \sum_{(Z_1,Z_2)\in(X,Y)^\perp\cap\Delta} \lambda_{Z_1,Z_2}\\ &= q^{\delta-\hat{r}} \sum_{(Z_1,Z_2)\in(X,Y)^\perp\cap\Delta^*} \lambda_{Z_1,Z_2}.\end{aligned}$$

By Lemma III.8, $\lambda_{Z_1,Z_2} = \mathrm{we}(\varphi(Z_1,Z_2)+\mathcal{C}_{\mathrm{const}})$ and these cosets are pairwise disjoint for $(Z_1,Z_2) \in (X,Y)^\perp \cap \Delta^*$, see Corollary III.12. Therefore, we obtain

$$\begin{aligned}&\sum_{(Z_1,Z_2)\in(X,Y)^\perp} \lambda_{Z_1,Z_2}\\ &= q^{\delta-\hat{r}} \sum_{(Z_1,Z_2)\in(X,Y)^\perp\cap\Delta^*} \mathrm{we}\left(\varphi(Z_1,Z_2)+\mathcal{C}_{\mathrm{const}}\right)\\ &= q^{\delta-\hat{r}}\mathrm{we}(H(X,Y))\end{aligned}$$

where

$$\begin{aligned}H(X,Y) &:= \bigcup_{(Z_1,Z_2)\in(X,Y)^\perp\cap\Delta^*} (\varphi(Z_1,Z_2)+\mathcal{C}_{\mathrm{const}})\\ &= \varphi((X,Y)^\perp\cap\Delta^*)+\mathcal{C}_{\mathrm{const}}.\end{aligned}$$

We will show next that $H(X,Y)^\perp = \langle\hat{\mathcal{C}}_{\mathrm{const}}, c(X,Y)\rangle$ for some element $c(X,Y)$. In order to do so, we need to compute the dimension of $H(X,Y)$. Since $\ker\varphi \cap \Delta^* = \{0\}$, we have $\dim\varphi((X,Y)^\perp \cap \Delta^*) = \dim((X,Y)^\perp \cap \Delta^*) = \dim\Delta^* - 1$. Furthermore, Lemma III.9, part a) shows that $\varphi(\Delta^*)\cap\mathcal{C}_{\mathrm{const}} = \{0\}$. As a consequence

$$\dim H(X,Y) = \dim \Delta^* - 1 + \dim \mathcal{C}_{\mathrm{const}} = k + \hat{r} - 1.$$

This implies $\dim H(X,Y)^\perp = n-k-\hat{r}+1 = \dim\hat{\mathcal{C}}_{\mathrm{const}}+1$. Furthermore, $H(X,Y) \subseteq \mathrm{im}\varphi + \mathcal{C}_{\mathrm{const}} = \mathcal{C}$ and along with Proposition II.8 this yields $\hat{\mathcal{C}}_{\mathrm{const}} \subseteq H(X,Y)^\perp$. All this shows that there exists some $c(X,Y) \in H(X,Y)^\perp \backslash \hat{\mathcal{C}}_{\mathrm{const}}$ such that

$$H(X,Y)^\perp = \langle\hat{\mathcal{C}}_{\mathrm{const}}, c(X,Y)\rangle.$$

Now we can compute $q^{-k}\boldsymbol{H}(\ell_{X,Y})$. From Theorem IV.7 we derive

$$\begin{aligned}&q^{-k}\boldsymbol{H}(\ell_{X,Y})\\ &= q^{-k}\boldsymbol{H}\left(\frac{1}{q^\delta(q-1)}\right.\\ &\qquad\left.\times\left(q\sum_{(Z_1,Z_2)\in(X,Y)^\perp}\lambda_{Z_1,Z_2} - q^{\delta-\hat{r}}\mathrm{we}(\mathcal{C}_C)\right)\right)\\ &= \frac{1}{q-1}q^{-\hat{r}-k}\left(q\boldsymbol{H}(\mathrm{we}(H(X,Y))) - \boldsymbol{H}(\mathrm{we}(\mathcal{C}_C))\right)\\ &= \frac{q^{-\hat{r}-k}}{q-1}\Big(q\cdot q^{k+\hat{r}-1}\mathrm{we}\big(\langle\hat{\mathcal{C}}_{\mathrm{const}}, c(X,Y)\rangle\big)\\ &\qquad - q^{k+\hat{r}}\mathrm{we}(\hat{\mathcal{C}}_{\mathrm{const}})\Big)\end{aligned}$$

where the last identity is again due to (II.5) and Corollary II.11. This proves the desired result. □

The last proposition together with Lemma III.8 reveals an immediate resemblance of the entries $q^{-k}\boldsymbol{H}(\ell_{X,Y})$ to that of any given adjacency matrix of the dual code of $\mathcal{C}$. Indeed, first notice that both matrices have the same number of zero entries since $\#(\ker\Phi)^\perp = q^{\delta+\hat{r}}$ is exactly the number of connected state pairs of the dual code. Moreover, Proposition IV.8 tells us that $q^{-k}\boldsymbol{H}(\ell_{X,Y})$ has $\#\Delta^\perp = q^{\delta-r}$ entries equal to we $(\hat{\mathcal{C}}_{\mathrm{const}})$. Applying Lemmas III.8 and III.9 to the dual code $\hat{\mathcal{C}}$ we see that the adjacency matrix of the dual code has the same number of entries equal to we $(\hat{\mathcal{C}}_{\mathrm{const}})$. The remaining entries also have an analogous form. All this indicates that there might be a strong relation between $q^{-k}(\boldsymbol{H}(\ell_{X,Y}))$ and the adjacency matrix of the dual code. This will be formulated in a precise conjecture in the next section and proven for a specific class of codes. The difficulty for proving this will be, among other things, that we need a concrete description of the mapping $(\ker\Phi)^\perp\backslash\Delta^\perp \longrightarrow \mathcal{F}, (X,Y) \mapsto c(X,Y)$ as used in the last part of Proposition IV.8.

V. A MacWilliams Identity for Convolutional Codes

In this section, we will formulate the MacWilliams identity and prove it for a particular class of codes. Let again the data be as in (I.1) and General Assumption II.3. Denote the associated adjacency matrix $\Lambda(G)$ simply by Λ. Furthermore, let $\hat{\mathcal{C}}$ be the dual code. We fix the following notation.

General Assumption V.1: Let $\hat{\mathcal{C}}$ have encoder matrix $\hat{G} \in \mathbb{F}[D]^{(n-k)\times n}$ and let the corresponding controller canonical form be denoted by $(\hat{A},\hat{B},\hat{C},\hat{E})$. Moreover, let the associated adjacency matrix be written as $\hat{\Lambda} =: (\hat{\lambda}_{X,Y})$ and let $\hat{\Delta}$ be the space of connected state pairs for $\hat{\mathcal{C}}$. Finally, we define the mappings $\hat{\varphi}$ and $\hat{\Phi}$ for the code $\hat{\mathcal{C}}$ analogously to (III.3) and Lemma III.9, and the spaces $\hat{\Delta}^-$ and $\hat{\Delta}^*$ analogously to Proposition III.11. Recall from Proposition II.8 that $\hat{\mathcal{C}}$ has $\hat{r}$ nonzero Forney indices.

We know from (II.4) that $\mathcal{C}$ and $\hat{\mathcal{C}}$ both have degree δ and thus the adjacency matrices Λ and $\hat{\Lambda}$ are both in $\mathbb{C}[W]^{q^\delta\times q^\delta}$.

As a consequence, we have all results of Section III literally available in a ^ version, and we will make frequent use of them.

Notice that duality implies $G\hat{G}^t = 0$. From Remark II.2, we know that

$$G(D) = B(D^{-1}I - A)^{-1}C + E$$
$$\hat{G}(D) = \hat{B}(D^{-1}I - \hat{A})^{-1}\hat{C} + \hat{E}.$$

Since E, $\hat{E}$ both have full row rank this implies

$$\operatorname{im} E = \ker \hat{E}^t. \tag{V.1}$$

Now we can formulate our conjecture. Recall the definition of the MacWilliams matrix $\mathcal{H}$ from Definition IV.2.

Conjecture V.2: The matrix $q^{-k}\boldsymbol{H}(\mathcal{H}\Lambda^t\mathcal{H}^{-1})$, where $\boldsymbol{H}$ is applied entrywise to the given matrix, is a representative of the generalized adjacency matrix of $\hat{\mathcal{C}}$. In other words, there exists some $P \in \mathrm{GL}_\delta(\mathbb{F})$ such that

$$\hat{\lambda}_{X,Y} = q^{-k}\boldsymbol{H}((\mathcal{H}\Lambda^t\mathcal{H}^{-1})_{XP,YP}), \quad \text{for all } (X,Y) \in \mathcal{F}. \tag{V.2}$$

Recall from Remark III.5, part b) that the adjacency matrices for two different minimal encoders of $\hat{\mathcal{C}}$ differ by conjugation with a suitable matrix $\mathcal{P}(P) \in \Pi$. This explains the presence of the matrix $P \in \mathrm{GL}_\delta(\mathbb{F})$ above. Of course, P depends on the chosen encoders G and $\hat{G}$. It is worth mentioning that in the case $\delta = 0$, identity (V.2) immediately leads to the MacWilliams identity for block codes as given in (II.5). Notice also that, due to Lemma IV.4 and (III.2), the conjecture implies the same statement if we replace $\mathcal{H}$ by an arbitrary Q MacWilliams matrix $\mathcal{H}(Q)$.

The conjecture is backed up by many numerical examples. A proof, however, is still open for the general case. As a first step, a somewhat weaker result will be proven in Theorem V.5. Thereafter, we will fully prove the conjecture for codes where $\delta = \hat{r}$ or $\delta = r$. In that case, we will even be able to precisely tell which transformation matrix $P \in \mathrm{GL}_\delta(\mathbb{F})$, depending on G and $\hat{G}$, to choose for (V.2) to be true. We need the following lemma. It still applies to the general situation.

Lemma V.3: Let

$$M := \begin{pmatrix} \hat{C}C^t & \hat{C}(B^tE)^t \\ \hat{B}^t\hat{E}C^t & 0 \end{pmatrix} \in \mathbb{F}^{2\delta\times 2\delta}.$$

Then

a) $\operatorname{im} M \subseteq (\ker \Phi)^\perp$;
b) $\ker \hat{\Phi} \oplus \hat{\Delta}^- \subseteq \ker M$;
c) $\operatorname{im} M \cap \Delta^\perp = \{0\}$;
d) M is injective on $\hat{\Delta}^*$;
e) $\operatorname{rank} M = r + \hat{r}$ and $\mathcal{M} \oplus \Delta^\perp = (\ker \Phi)^\perp$, where $\mathcal{M}$ is given by $\mathcal{M} := \operatorname{im} M = \{(X,Y)M \mid (X,Y) \in \hat{\Delta}^*\}$.

Proof: First notice that by (V.1)

$$M^t = \begin{pmatrix} C \\ B^tE \end{pmatrix} (\hat{C}^t \ (\hat{B}^t\hat{E})^t)$$

and therefore

$$\begin{aligned}\beta((X',Y'),(X,Y)M) &= (X',Y')M^t(X,Y)^t \\ &= \beta(\varphi(X',Y'),\hat{\varphi}(X,Y))\end{aligned} \tag{V.3}$$

for all state pairs $(X,Y),(X',Y') \in \mathcal{F}$. Remember also that $\hat{\varphi}(X,Y) \in C_{\hat{\mathcal{C}}}$ for all $(X,Y) \in \mathcal{F}$.

a) Follows from (V.3) since for $(X',Y') \in \ker\Phi$ we have $\varphi(X',Y') \in \mathcal{C}_{\text{const}} = (C_{\hat{\mathcal{C}}})^\perp$.

b) If $(X,Y) \in \ker\hat{\Phi}\oplus\hat{\Delta}^-$, then $\hat{\varphi}(X,Y) \in \hat{\mathcal{C}}_{\text{const}}$ by Corollary III.12 and Lemma III.9, part a). Thus, $\hat{\varphi}(X,Y) \in (\mathcal{C}_C)^\perp$ while $\varphi(X',Y') \in \mathcal{C}$ for all $(X',Y') \in \mathcal{F}$. Now (V.3) along with the regularity of the bilinear form β shows $(X,Y)M = (0,0)$.

c) Let $(X,Y)M \in \Delta^\perp$. Then by (V.3) we have $\hat{\varphi}(X,Y) \in \varphi(\Delta)^\perp$. Since also $\hat{\varphi}(X,Y) \in C_{\hat{\mathcal{C}}} = (\mathcal{C}_{\text{const}})^\perp$, we obtain from (III.4) and Proposition II.8 that $\hat{\varphi}(X,Y) \in \hat{\mathcal{C}}_{\text{const}}$. But then $(X,Y) \in \ker\hat{\Phi}$ and (b) implies $(X,Y)M = (0,0)$.

d) Let $(X,Y)M = 0$ for some $(X,Y) \in \hat{\Delta}^*$. Similarly to item c), we obtain by use of (V.3) and (III.4)

$$\begin{aligned}\hat{\varphi}(X,Y) \in (\operatorname{im}\varphi)^\perp \cap C_{\hat{\mathcal{C}}} &= (\operatorname{im}\varphi)^\perp \cap (\mathcal{C}_{\text{const}})^\perp \\ &= (\operatorname{im}\varphi + \mathcal{C}_{\text{const}})^\perp = \hat{\mathcal{C}}_{\text{const}}.\end{aligned}$$

But this means that $(X,Y) \in \ker\hat{\Phi}$ and the assumption $(X,Y) \in \hat{\Delta}^*$ finally yields $(X,Y) = (0,0)$.

e) The assertion on the rank follows from items d) and b) since $\dim\hat{\Delta}^* = r+\hat{r}$ and $\dim(\ker\hat{\Phi}\oplus\hat{\Delta}^-) = 2\delta-(r+\hat{r})$. The rest is immediate from the above and $\dim\,(\ker\Phi)^\perp - \dim\Delta^\perp = r+\hat{r}$. □

The following result will be crucial for investigating Conjecture V.2.

Theorem V.4: Let $M \in \mathbb{F}^{2\delta\times 2\delta}$ be as in Lemma V.3. Then

$$\hat{\lambda}_{X,Y} = q^{-k}\boldsymbol{H}(\ell_{(X,Y)M}), \quad \text{for all } (X,Y) \in \hat{\Delta}.$$

Proof: Recall that $\hat{\Delta} = \ker\hat{\Phi} \oplus \hat{\Delta}^*$. For $(X',Y') \in \ker\hat{\Phi}$ and $(X,Y) \in \hat{\Delta}^*$ we have $\hat{\lambda}_{X'+X,Y'+Y} = \hat{\lambda}_{X,Y}$ due to Proposition III.11, part c). Furthermore, $\ell_{(X',Y')M+(X,Y)M} = \ell_{(X,Y)M}$ by Lemma V.3, part b). Hence, it suffices to show the result for $(X,Y) \in \hat{\Delta}^*$. For $(X,Y) = (0,0)$, the result is obviously true by Lemma III.8 and Proposition IV.8. Thus, let $(X,Y) \neq (0,0)$. By Lemma V.3, part e) this yields $(X,Y)M \in (\ker\Phi)^\perp\backslash\Delta^\perp$. Hence, $q^{-k}\boldsymbol{H}\,(\ell_{(X,Y)M})$ needs to be computed according to the last case in Proposition IV.8. In order to do so we need to find a vector $c((X,Y)M)$ satisfying the requirements given there. We will show that $\hat{\varphi}(X,Y)$ is such a vector. First of all, it is clear that $\hat{\varphi}(X,Y) \in C_{\hat{\mathcal{C}}} = (\mathcal{C}_{\text{const}})^\perp$. Moreover, $\hat{\varphi}(X,Y) \notin \hat{\mathcal{C}}_{\text{const}}$ since $(X,Y) \in \hat{\Delta}^*\backslash\{0\}$. Finally, applying (V.3) to $(X',Y') \in ((X,Y)M)^\perp \cap \Delta^*$ shows that $\hat{\varphi}(X,Y) \in [\varphi(((X,Y)M)^\perp \cap \Delta^*)]^\perp$. All this shows that we may choose $c((X,Y)M)$ in Proposition IV.8 as $\hat{\varphi}(X,Y)$. Now that proposition yields

$$\begin{aligned}&q^{-k}\boldsymbol{H}(\ell_{(X,Y)M}) \\ &\quad = \frac{1}{q-1}(\mathrm{we}\,(\langle\hat{\varphi}(X,Y),\hat{\mathcal{C}}_{\text{const}}\rangle) - \mathrm{we}\,(\hat{\mathcal{C}}_{\text{const}}))\end{aligned}$$

and this coincides with $\hat{\lambda}_{X,Y}$ due to Lemma III.8. □

For the sequel let $\mathcal{G}$ be any direct complement of $(\ker\Phi)^\perp$ in $\mathcal{F}$. Due to Lemma V.3(e) we have the following decompositions of $\mathcal{F}$:

$$\begin{array}{ccccccc} \mathcal{F} & = & \overbrace{\widehat{\Delta}^* \;\oplus\; \ker\widehat{\Phi}}^{\widehat{\Delta}} & \oplus & \widehat{\Delta}^- \\ \Big\downarrow f & & \Big\downarrow f_0 \qquad \Big\downarrow f_1 & & \Big\downarrow f_2 \\ \mathcal{F} & = & \underbrace{\mathcal{M} \;\oplus\; \Delta^\perp}_{(\ker\Phi)^\perp} & \oplus & \mathcal{G} \end{array} \qquad \text{(V.4)}$$

where, due to identical dimensions, there exist isomorphisms in each column. For f_0, we choose the isomorphism induced by the matrix M from Lemma V.3, and thus, $\mathcal{M} = \operatorname{im} M$ as before. This picture leads to the following result.

Theorem V.5: Consider the diagram (V.4) and let the isomorphism f_0 be induced by the matrix M from Lemma V.3. Fix any isomorphisms f_1 and f_2 in the diagram and let $f := f_0 \oplus f_1 \oplus f_2$ be the associated automorphism on $\mathcal{F}$. Then

$$\hat{\lambda}_{X,Y} = q^{-k}\boldsymbol{H}((\mathcal{H}\Lambda\mathcal{H})_{f(X,Y)}), \quad \text{for all } (X,Y) \in \mathcal{F}. \quad \text{(V.5)}$$

As a consequence

$$\hat{\lambda}_{f^{-1}(-Y,X)} = q^{-k}\boldsymbol{H}((\mathcal{H}\Lambda^t\mathcal{H}^{-1})_{X,Y}), \text{ for all } (X,Y) \in \mathcal{F}.$$

In particular, the entries of $\hat{\Lambda}$ and $q^{-k}\boldsymbol{H}(\mathcal{H}\Lambda^t\mathcal{H}^{-1})$ coincide up to reordering.

Proof: Recall from (IV.2) that $(\mathcal{H}\Lambda\mathcal{H})_{f(X,Y)} = \ell_{f(X,Y)}$. We have to consider three cases.

1) If $(X,Y) \notin \hat{\Delta}$, then $f(X,Y) \notin (\ker\Phi)^\perp$ and $\hat{\lambda}_{X,Y} = 0 = q^{-k}\boldsymbol{H}\,(\ell_{f(X,Y)})$ due to the very definition of $\hat{\Delta}$ and Proposition IV.8.

2) If $(X,Y) \in \ker\hat{\Phi}$ then $\hat{\varphi}(X,Y) \in \hat{\mathcal{C}}_{\text{const}}$ and $f(X,Y) \in \Delta^\perp$. Now Lemma III.8 as well as Proposition IV.8 yield $\hat{\lambda}_{X,Y} = \text{we}(\hat{\mathcal{C}}_{\text{const}}) = q^{-k}\boldsymbol{H}(\ell_{f(X,Y)})$.

3) For the remaining case, we have $(X,Y) \in \hat{\Delta}\backslash\ker\hat{\Phi}$. Writing $(X,Y) = (X_1,Y_1) + (X_2,Y_2)$ where $(X_1,Y_1) \in \hat{\Delta}^*$ and $(X_2,Y_2) \in \ker\hat{\Phi}$, Proposition III.11, part c) yields $\hat{\lambda}_{X,Y} = \hat{\lambda}_{X_1,Y_1}$ while Theorem IV.7 implies $\ell_{f(X,Y)} = \ell_{(X_1,Y_1)M}$. Now the result follows from Theorem V.4.

For the second statement, put $\Gamma := \mathcal{H}\Lambda^t\mathcal{H}^{-1}$. Notice first that Lemma IV.4 and the definition of $\mathcal{P}(-I)$ as given in (III.4) yield $\mathcal{H}\Lambda\mathcal{H} = \mathcal{P}(-I)\Gamma^t$. This implies $(\mathcal{H}\Lambda\mathcal{H})_{X,Y} = \Gamma_{Y,-X}$. Now we obtain from (V.5) $\hat{\lambda}_{f^{-1}(X,Y)} = q^{-k}\boldsymbol{H}\,(\Gamma_{Y,-X})$ and, thus, $\hat{\lambda}_{f^{-1}(-Y,X)} = q^{-k}\boldsymbol{H}\,(\Gamma_{X,Y})$. This concludes the proof. □

It needs to be stressed that the theorem does not prove Conjecture V.2 since we did not show that the automorphism is of the form $f^{-1}(-Y,X) = (XQ,YQ)$ for some suitable $Q \in \text{GL}_\delta(\mathbb{F})$ and all $(X,Y) \in \mathcal{F}$. The difficulty in proving Conjecture V.2 consists precisely in finding isomorphisms f_0, f_1, f_2 for Diagram (V.4) such that $f^{-1}(-Y,X)$ has such a form. This will be accomplished next for the class of convolutional codes for which either r or $\hat{r}$ is equal to δ.

We begin with the case where $\hat{r} = \delta$. Notice that this is equivalent to saying that all nonzero Forney indices of $\hat{\mathcal{C}}$ have value one. Therefore, in this case

$$\hat{G} = \begin{pmatrix} \hat{E}_1 \\ \hat{E}_2 \end{pmatrix} + D\begin{pmatrix} \hat{C} \\ 0 \end{pmatrix}$$

where

$$\hat{E} = \begin{pmatrix} \hat{E}_1 \\ \hat{E}_2 \end{pmatrix} = \begin{pmatrix} \hat{B}^t\hat{E} \\ \hat{E}_2 \end{pmatrix} \text{ and rank } \begin{pmatrix} \hat{C} \\ \hat{E}_2 \end{pmatrix} = n-k \quad \text{(V.6)}$$

where the last part follows from minimality of the encoder $\hat{G}$. Furthermore, $\operatorname{im}\hat{E}_2 = \hat{\mathcal{C}}_{\text{const}}$. We will need the following technical lemma.

Lemma V.6: Let $\hat{r} = \delta$. Then the controller canonical forms satisfy

1) $\hat{B}^t\hat{E}C^t \in \text{GL}_\delta(\mathbb{F})$,
2) $\hat{C}E^tB + \hat{C}C^tA = -\hat{B}^t\hat{E}C^t$.

Proof:

1) Since $\hat{r} = \delta$, we have $\hat{B}^t = (I_\delta, 0)$ and thus, $\hat{B}^t\hat{E}$ consists of the first δ rows of $\hat{E}$. As a consequence, $\hat{B}^t\hat{E}$ has full row rank δ. Suppose now that $\operatorname{rank}\hat{B}^t\hat{E}C^t < \delta$. Since $\hat{E}E^t = 0$ this implies $\operatorname{rank}\hat{B}^t\hat{E}(C^t,E^t) < \delta$. Hence, there exists a nonzero vector $a \in \mathbb{F}^\delta$ such that $a\hat{B}^t\hat{E}(C^t,E^t) = 0$. In other words, $a\hat{B}^t\hat{E} \in (\mathcal{C}_C)^\perp = \hat{\mathcal{C}}_{\text{const}}$, where the last identity follow from Proposition II.8. Now the full row rank of $\hat{B}^t\hat{E}$ shows that $\operatorname{im}\hat{B}^t\hat{E} \cap \hat{\mathcal{C}}_{\text{const}} \neq \{0\}$, a contradiction to Remark II.6, part 2) applied to the code $\hat{\mathcal{C}}$. Hence, $\operatorname{rank}\hat{B}^t\hat{E}C^t = \delta$.

2) By duality, we have $\hat{G}G^t = 0$. From Remark II.2 we know that the controller canonical forms determine the corresponding encoders via

$$\begin{aligned} G(D) &= B(D^{-1}I - A)^{-1}C + E \\ &= B\sum_{l\geq 1} D^lA^{l-1}C + E \end{aligned}$$

and

$$\hat{G}(D) = \hat{B}\sum_{l\geq 1} D^l\hat{A}^{l-1}\hat{C} + \hat{E}$$

where due to nilpotency the sums are finite. Since $\hat{r} = \delta$, all Forney indices of $\hat{\mathcal{C}}$ are at most one and therefore $\hat{A} = 0$. Hence, $\hat{G} = D\hat{B}\hat{C} + \hat{E}$. Now we compute

$$\begin{aligned} 0 &= \hat{G}G^t \\ &= D(\hat{E}C^tB^t + \hat{B}\hat{C}E^t) \\ &\quad + \sum_{l\geq 2} D^l(\hat{E}C^t(A^t)^{l-1}B^t + \hat{B}\hat{C}C^t(A^t)^{l-2}B^t). \end{aligned}$$

Hence, the coefficients of $D^l, l \geq 1$ are zero. Left-multiplying the coefficient of D by $\hat{B}^t$, right-multiplying it by B, and using $\hat{B}^t\hat{B} = I$ implies the identity

$$\hat{B}^t\hat{E}C^tB^tB + \hat{C}E^tB = 0. \quad \text{(V.7)}$$

Furthermore, if we multiply the coefficient of D^l by $\hat{B}^t$ from the left and by BA^{l-1} from the right we obtain

$$\begin{aligned} 0 &= (\hat{B}^t\hat{E}C^t(A^t)^{l-1}B^t + \hat{C}C^t(A^t)^{l-2}B^t)BA^{l-1} \\ &= (\hat{B}^t\hat{E}C^tA^t + \hat{C}C^t)((A^t)^{l-2}B^tBA^{l-1}) \\ &= (\hat{B}^t\hat{E}C^tA^t + \hat{C}C^t)((A^t)^{l-2}(I - A^tA)A^{l-1}) \end{aligned}$$

where the last identity is due to Remark II.4. Addition of these equations yields

$$\begin{aligned} 0 &= (\hat{B}^t\hat{E}C^tA^t + \hat{C}C^t)\sum_{l\geq 2}((A^t)^{l-2}A^{l-1} - (A^t)^{l-1}A^l) \\ &= \hat{B}^t\hat{E}C^tA^tA + \hat{C}C^tA \end{aligned}$$

where the last identity follows from the nilpotency of A. Now we conclude with Remark II.4 and the aid of (V.7)

$$\begin{aligned} \hat{C}C^tA &= -\hat{B}^t\hat{E}C^tA^tA \\ &= -\hat{B}^t\hat{E}C^t + \hat{B}^t\hat{E}C^tB^tB \\ &= -\hat{B}^t\hat{E}C^t - \hat{C}E^tB \end{aligned}$$

which is what we wanted. □

Now we can present an isomorphism f_1 for Diagram (V.4).

Lemma V.7: Let $\hat{r} = \delta$. Define

$$M_1 := \begin{pmatrix} -\hat{C}C^t & \hat{C}C^tA \\ 0 & 0 \end{pmatrix} \in \mathbb{F}^{2\delta \times 2\delta}.$$

Then

a) $\operatorname{im} M_1 \subseteq \Delta^{\perp}$,
b) $\ker \hat{\Phi} \cap \ker M_1 = \{0\}$,
c) $\operatorname{rank} M_1 = \delta - r$.

As a consequence, the matrix M_1 induces an isomorphism $f_1 : \ker \hat{\Phi} \longrightarrow \Delta^{\perp}$.

Proof:

a) Using Lemma III.7 it suffices to show that $(\hat{C}C^t)_{ij} = 0$ for all $1 \le i \le \delta$ and $j \in \mathcal{J}$. From (V.6), we see that all rows in $\hat{C}$ are leading coefficient rows in $\hat{G}$. By definition of the controller canonical form, the jth rows of C, where $j \in \mathcal{J}$, are leading coefficient rows of the encoder G. Now $\hat{G}G^t = 0$ implies the desired result. As a consequence, we also have $\operatorname{rank} M_1 \le \delta - r$.

b) From Lemma V.6 we know that $\hat{B}^t\hat{E}C^t \in \mathrm{GL}_\delta(\mathbb{F})$. Let now $(X, Y) \in \ker \hat{\Phi} \cap \ker M_1$. Then $X\hat{C}C^t = 0$ and, due to Lemma V.3, part b) we also have $(X, Y)M = 0$. As a consequence, $X\hat{C}C^t + Y\hat{B}^t\hat{E}C^t = Y\hat{B}^t\hat{E}C^t = 0$ and from the above we conclude $Y = 0$. Now $(X, 0) \in \ker \hat{\Phi}$ yields $\hat{\varphi}(X, 0) = X\hat{C} \in \hat{\mathcal{C}}_{\text{const}}$. Using the full row rank of the rightmost matrix in (V.6) we conclude that $X = 0$.

c) follows from items a) and b), since $\dim \ker \hat{\Phi} = \delta - r$. □

Now we are able to prove our main result. The crucial step will be a suitable choice for the space $\hat{\Delta}^*$. Recall that, so far, $\hat{\Delta}^*$ was just any direct complement of $\ker \hat{\Phi}$ in $\hat{\Delta}$. As we will see below, $\ker M_1$ is such a direct complement.

Theorem V.8: Let $\hat{r} = \delta$, that is, each Forney index of the code $\hat{\mathcal{C}}$ is at most 1. Then $Q := -\hat{B}^t\hat{E}C^t \in \mathrm{GL}_\delta(\mathbb{F})$ and

$$\hat{\lambda}_{X,Y} = q^{-k}\boldsymbol{H}((\mathcal{H}\Lambda^t\mathcal{H}^{-1})_{XQ,YQ}) \text{ for all } (X, Y) \in \mathcal{F}.$$

As a consequence

$$\hat{\Lambda} = q^{-k}\boldsymbol{H}(\mathcal{P}(Q)\mathcal{H}\Lambda^t\mathcal{H}^{-1}\mathcal{P}(Q)^{-1}) \tag{V.8}$$

where the MacWilliams transformation $\boldsymbol{H}$ has to be applied entry-wise.

Proof: The invertibility of $Q = \hat{B}^t\hat{E}C^t$ has been shown in Lemma V.6, part 1). Choose the matrices M and M_1 as in Lemmas V.3 and V.7. By Lemma V.6, part 2) we have

$$M + M_1 = \begin{pmatrix} 0 & Q \\ -Q & 0 \end{pmatrix} \in \mathrm{GL}_{2\delta}(\mathbb{F}). \tag{V.9}$$

Notice that due to $\hat{r} = \delta$ we have $\mathcal{F} = \hat{\Delta}$. In particular, the last column of Diagram V.4 is trivial. Lemma V.7 shows that $\hat{\Delta}^* := \ker M_1$ is a direct complement of $\ker \hat{\Phi}$ in $\mathcal{F}$. Now define the automorphism $f : \mathcal{F} \longrightarrow \mathcal{F}$ as $f(X, Y) = (X, Y)(M + M_1)$. Using Lemmas V.3 and V.7, in particular $\hat{\Delta}^* = \ker M_1$ and $\ker \hat{\Phi} \subseteq \ker M$, we see that f is of the form $f = f_0 \oplus f_1$ as required in Diagram (V.4). Here f_0 and f_1 are induced by the matrices M and M_1, respectively. It is easy to see that $f^{-1}(-Y, X) = (XQ^{-1}, YQ^{-1})$ and, therefore ,Theorem V.5 yields $\hat{\lambda}_{XQ^{-1},YQ^{-1}} = q^{-k}\boldsymbol{H}\left((\mathcal{H}\Lambda^t\mathcal{H}^{-1})_{X,Y}\right)$. Using (III.2) this implies the desired result. □

Example V.9: Note that the code $\mathcal{C}$ from Example II.9 satisfies $\hat{r} = \delta$. Hence we can apply Theorem V.8 to this code. The automorphism Q can be calculated as

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

and, using the lexicographic ordering (III.1), the permutation matrix is given by

$$\mathcal{P}(Q) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Using this as well as the adjacency matrices given in Example III.2 and the MacWilliams matrix $\mathcal{H}$ in Example IV.5, one can check straightforwardly identity (V.6).

We can easily transfer our result to convolutional codes with $\delta = r$.

Theorem V.10: Let $\delta = r$, that is, each Forney index of the code $\mathcal{C}$ is at most 1. Then $P := -\hat{C}E^tB \in \mathrm{GL}_\delta(\mathbb{F})$ and for all $(X, Y) \in \mathcal{F}$ we have $\hat{\lambda}_{X,Y} = q^{-k}\boldsymbol{H}((\mathcal{H}\Lambda^t\mathcal{H}^{-1})_{XP,YP})$. In other words

$$\hat{\Lambda} = q^{-k}\boldsymbol{H}(\mathcal{P}(P)\mathcal{H}\Lambda^t\mathcal{H}^{-1}\mathcal{P}(P)^{-1}).$$

Proof: We can apply Theorem V.8 to the code $\hat{\mathcal{C}}$. Hence, the matrix $Q := -B^tE\hat{C}^t$ is regular and, since $\dim \hat{\mathcal{C}} = n - k$, we obtain

$$\Lambda = q^{-n+k}\boldsymbol{H}(\mathcal{P}(Q)\mathcal{H}\hat{\Lambda}^t\mathcal{H}^{-1}\mathcal{P}(Q)^{-1}).$$

Applying $\boldsymbol{H}^2(f) = q^n f$ and the $\mathbb{C}$-linearity of $\boldsymbol{H}$ we arrive at $q^{-k}\boldsymbol{H}(\mathcal{H}^{-1}\mathcal{P}(Q)^{-1}\Lambda\mathcal{P}(Q)\mathcal{H}) = \hat{\Lambda}^t$. Transposing this equation and remembering that $\mathcal{H}^t = \mathcal{H}$ while $\mathcal{P}(Q)^t = \mathcal{P}(Q)^{-1}$, one gets

$$\hat{\Lambda} = q^{-k}\boldsymbol{H}(\mathcal{H}\mathcal{P}(Q)^{-1}\Lambda^t\mathcal{P}(Q)\mathcal{H}^{-1}).$$

Now Lemma IV.4 yields

$$\mathcal{H}\mathcal{P}(Q)^{-1} = \mathcal{H}\mathcal{P}(Q^{-1}) = \mathcal{P}(Q^t)\mathcal{H} = \mathcal{P}(P)\mathcal{H}.$$

This concludes the proof. □

Incorporating the permutation matrix into the MacWilliams matrix, see Lemma IV.4, we can state Theorems V.8 and V.10 in terms of P-MacWilliams matrices only.

Corollary V.11:

1) If $\hat{r} = \delta$, then the matrix $Q = -\hat{B}^t\hat{E}C^t$ is in $\mathrm{GL}_\delta(\mathbb{F})$ and $\hat{\Lambda} = q^{-k}\boldsymbol{H}(\mathcal{H}(Q)\Lambda^t\mathcal{H}(Q)^{-1})$.
2) If $r = \delta$, then the matrix $P = -\hat{C}E^tB$ is in $\mathrm{GL}_\delta(\mathbb{F})$ and $\hat{\Lambda} = q^{-k}\boldsymbol{H}(\mathcal{H}(P)\Lambda^t\mathcal{H}(P)^{-1})$.

Using the notion of the generalized adjacency matrix as defined in Remark III.5, part b) we obtain the following consequence, formulated independently of any chosen representation.

Theorem V.12: Let $r = \delta$ or $\hat{r} = \delta$. Then the generalized adjacency matrix of $\mathcal{C}$ uniquely determines the generalized adjacency matrix of the dual code $\hat{\mathcal{C}}$. More precisely, let $[\Lambda]$ and $[\hat{\Lambda}]$ be the generalized adjacency matrices of $\mathcal{C}$ and $\hat{\mathcal{C}}$, respectively. Then, in a suggestive notation

$$[\hat{\Lambda}] = q^{-k}\boldsymbol{H}(\mathcal{H}[\Lambda]^t\mathcal{H}^{-1}).$$

We close this section with an example supporting Conjecture V.2 that is not covered by the cases in Theorem V.8 or Theorem V.10.

Example V.13: Let $q = 3$, thus $\mathbb{F} = \mathbb{F}_3$, and define

$$G = \begin{pmatrix} 1+D^2 & 2+D & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Put $\mathcal{C} = \operatorname{im} G$. Then G is a minimal basic matrix and thus $\mathcal{C}$ is a $(3,2,2)$ code. The dual code is given by $\hat{\mathcal{C}} = \operatorname{im} \hat{G}$ where $\hat{G} = (2+D\ \ 2+2D^2\ \ 2+D$. Notice that $r = \hat{r} = 1 \neq \delta$. Using the controller canonical forms one can straightforwardly compute the adjacency matrices $\Lambda, \hat{\Lambda} \in \mathbb{C}[W]^{9\times 9}$. Then, via a systematic search, one finds that identity (V.2) is satisfied if one chooses the regular matrix $P = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{F}_3)$. In the same way one can establish plenty of examples.

VI. Unit Constraint-Length Codes

In the last section, we want to have a closer look at codes with degree $\delta = 1$, also called unit constraint-length codes. Notice that in this case $r = \hat{r} = \delta = 1$. The situation now becomes particularly simple since, first, $\mathrm{GL}_\delta(\mathbb{F}) = \mathbb{F}^*$ and, second, the adjacency matrices Λ and $\hat{\Lambda}$ do not depend on the choice of the encoder matrices G and $\hat{G}$. The latter is a consequence of (III.2) along with Remark III.5, parts a) and b). Notice also that in Diagram (V.4), the second and third columns are trivial. Using once more Remark III.5, part a), we finally see that the statements of both Theorems V.8 and V.10 reduce to the nice short formula

$$\hat{\Lambda} = q^{-k}\boldsymbol{H}(\mathcal{H}\Lambda^t\mathcal{H}^{-1}). \tag{VI.1}$$

In the paper [10], the so-called weight enumerator state diagram has been studied for codes with degree one. They are defined as the state diagram of the encoder where each directed edge is labeled by the weight enumerator of a certain affine code. A type of MacWilliams identity has been derived for these objects [10, Theorem 4]. It consists of a separate transformation formula for each of these labels. After some notational adjustment one can show that the weight enumerator state diagram is in essence identical to the adjacency matrix of the code. Furthermore, if stated in our notation, the MacWilliams identity in [10, Theorem 4] reads as

$$\hat{\lambda}_{X,Y} = \begin{cases} q^{-k-1}\boldsymbol{H}\left(\lambda_{0,0} + (q-1)(\lambda_{0,1} + \sum_{Y\in\mathbb{F}}\lambda_{1,Y})\right), & \text{if } (X,Y) = (0,0) \\ q^{-k-1}\boldsymbol{H}\left(\lambda_{0,0} + q\lambda_{X,Y} - \lambda_{0,1} - \sum_{Y\in\mathbb{F}}\lambda_{1,Y}\right), & \text{else.} \end{cases} \tag{VI.2}$$

In the sequel, we will briefly sketch that this result coincides with identity (VI.1). In order to do so, use again $\ell_{X,Y}$ as introduced in (IV.2). Then (VI.1) turns into

$$\hat{\lambda}_{X,Y} = q^{-k}\boldsymbol{H}(\ell_{-Y,X}), \quad \text{for all } (X,Y) \in \mathcal{F}. \tag{VI.3}$$

Now we are in a position to derive (VI.2). Consider first the case $(X,Y) = (0,0)$. Recalling Theorem IV.7, Proposition III.13, and Remark III.5, part a) we find

$$\begin{aligned} q\ell_{0,0} = \mathrm{we}(\mathcal{C}_C) &= \sum_{(X,Y)\in\mathcal{F}} \lambda_{X,Y} \\ &= \lambda_{0,0} + \sum_{Y\in\mathbb{F}^*}\lambda_{0,Y} + \sum_{X\in\mathbb{F}^*}\sum_{Y\in\mathbb{F}}\lambda_{X,Y} \\ &= \lambda_{0,0} + (q-1)(\lambda_{0,1} + \sum_{Y\in\mathbb{F}}\lambda_{1,Y}). \end{aligned}$$

Using (VI.3), this establishes the first case of (VI.2). For the second case, let $(X,Y) \in \mathcal{F}\backslash(0,0)$. Since $\Delta^\perp = \{0\}$ and $(\ker\Phi)^\perp = \mathcal{F}$ one observes that in Theorem IV.7 the third case has to be applied. Along with Proposition III.13 and Remark III.5, part a) this yields

$$\begin{aligned} &q\ell_{-Y,X} \\ &= \frac{1}{q-1}\left(q\sum_{(Z_1,Z_2)\in(-Y,X)^\perp}\lambda_{Z_1,Z_2} - \mathrm{we}(\mathcal{C}_C)\right) \\ &= \frac{1}{q-1}\left(q\sum_{\alpha\in\mathbb{F}}\lambda_{\alpha X,\alpha Y} - \mathrm{we}(\mathcal{C}_C)\right) \\ &= \frac{1}{q-1}\left(q(q-1)\lambda_{X,Y} + q\lambda_{0,0} - \sum_{(Z_1,Z_2)\in\mathcal{F}}\lambda_{Z_1,Z_2}\right) \\ &= q\lambda_{X,Y} + \lambda_{0,0} - \frac{1}{q-1}\sum_{(Z_1,Z_2)\in\mathcal{F}\backslash\{(0,0)\}}\lambda_{Z_1,Z_2} \\ &= q\lambda_{X,Y} + \lambda_{0,0} - \lambda_{0,1} - \sum_{Y\in\mathbb{F}}\lambda_{1,Y}. \end{aligned}$$

Combining this with (VI.3) leads to the second case of (VI.2).

VII. Conclusion

In this paper, we studied the adjacency matrices for convolutional codes. We introduced a transformation consisting of conjugation with the MacWilliams matrix followed by entrywise application of the MacWilliams Identity for block codes. We proved that the resulting matrix coincides up to reordering of the entries with the adjacency matrix of the dual code, and we presented the reordering mapping explicitly. This result can be regarded as a weak MacWilliams Identity for convolutional codes. However, we strongly believe that the reordering of the entries can even be expressed in terms of an isomorphism on the state space, and indeed, we proved this statement for a particular class of convolutional codes. The general case has to remain open for future research.

REFERENCES

[1] F. J. MacWilliams, "Combinatorial Problems of Elementary Abelian Groups," Ph.D. dissertation, Harvard Univ., Cambridge, MA, 1962.

[2] F. J. MacWilliams, "A theorem on the distribution of weights in a systematic code," *Bell Syst. Tech. J.*, vol. 42, pp. 79–94, 1963.

[3] F. J. MacWilliams and N. J. A, "Sloane," in *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[4] W. C. Huffman and V. Pless, *Fundamentals of Error-Correcting Codes*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[5] J. A. Wood, "Duality for modules over finite rings and applications to coding theory," *Amer. J. Math.*, vol. 121, pp. 555–575, 1999.

[6] H. Q. Dinh and S. R. López-Permouth, "On the equivalence of codes over finite rings," *Appl. Algebra Eng. Comm. Comput.*, vol. 15, pp. 37–50, 2004.

[7] J. B. Shearer and R. J. McEliece, "There is no MacWilliams identity for convolutional codes," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 6, pp. 775–776, Dec. 1977.

[8] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, no. 5, pt. 1, pp. 751–772, Oct. 1971.

[9] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding*. Piscataway, NJ: IEEE Press, 1999.

[10] K. A. S. Abdel-Ghaffar, "On unit constrained-length convolutional codes," *IEEE Trans. Inf. Theory*, vol. 38, no. 1, pp. 200–206, Jan. 1992.

[11] R. J. McEliece, , M. Darnell and B. Honory, Eds., "How to Compute Weight Enumerators for Convolutional Codes," in *Communications and Coding (P. G. Farrell 60th Birthday Celebration)*. New York: Wiley, 1998, pp. 121–141.

[12] J. L. Massey and M. K. Sain, "Codes, automata, and continuous systems: Explicit interconnections," *IEEE Trans. Autom. Contr.*, vol. AC-12, no. 6, pp. 644–650, Dec. 1967.

[13] J. Rosenthal, "Connections between linear systems and convolutional codes," in *Codes, Systems, and Graphical Models*, B. Marcus and J. Rosenthal, Eds. New York: Springer, 2001, pp. 39–66.

[14] J. Rosenthal and E. V. York, "BCH convolutional codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1833–1844, Sep. 1999.

[15] F. Fagnani and S. Zampieri, "System-theoretic properties of convolutional codes over rings," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2256–2274, Sep. 2001.

[16] R. Hutchinson, J. Rosenthal, and R. Smarandache, "Convolutional codes with maximum distance profile," *Syst. Contr. Lett.*, vol. 54, pp. 53–63, 2005.

[17] H. Gluesing-Luerssen, "On the weight distribution of convolutional codes," *Linear Algebra and its Applications*, vol. 408, pp. 298–326, 2005.

[18] G. D. Forney, Jr., "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Contr.*, vol. 13, pp. 493–520, 1975.

[19] R. J. McEliece, "The algebraic theory of convolutional codes," in *Handbook of Coding Theory*, V. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, 1998, vol. 1, pp. 1065–1138.

[20] G. D. Forney, Jr., "Structural analysis of convolutional codes via dual codes," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 512–518, Jul. 1973.

[21] G. D. Forney, Jr., "Algebraic structure of convolutional codes, and algebraic system theory," in *Mathematical System Theory, The Influence of R. E. Kalman*, A. Antoulas, Ed. Berlin, Germany: Springer, 1991, pp. 527–557.

[22] R. Johannesson, P. Ståhl, and E. Wittenmark, "A note on type II convolutional codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1510–1514, Jul. 2000.

[23] G. D. Forney, Jr., "Codes on graphs: Normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.

[24] G. D. Forney, Jr. and M. D. Trott, "The dynamics of group codes: Dual abelian group codes and systems," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2935–2965, Dec. 2004.

[25] T. Mittelholzer, "Convolutional codes over groups: A pragmatic approach," in *Proc. 33rd Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, 1995, pp. 380–381.

[26] H.-A. Loeliger and T. Mittelholzer, "Convolutional codes over groups," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1660–1686, Nov. 1996.

[27] H.-A. Loeliger, "The convoluted history of minimal encoders," in *Proc. 35th Allerton Annu. Conf. Communications, Control, and Computing*, Monticello, IL, 1997, pp. 687–690.

[28] G. D. Forney, Jr., M. Grassl, and S. Guha, "Convolutional and tail-biting quantum error-correcting codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 865–880, Mar. 2007.

[29] J. Justesen, E. Paaske, and M. Ballan, "Quasicyclic unit memory convolutional codes," *IEEE Trans. Inf. Theory*, vol. 36, no. 3, pp. 540–547, May 1990.

[30] S. Höst, R. Johannesson, and V. V. Zyablov, "Woven convolutional codes I: Encoder properties," *IEEE Trans. Inf. Theory*, vol. 48, no. 1, pp. 149–161, Jan. 2002.

[31] R. Lidl and H. Niederreiter, *Finite Fields*. Cambridge, U.K.: Cambridge Univ. Press, 1997.