State Space Realizations and Monomial Equivalence for Convolutional Codes

Heide Gluesing-Luerssen*, Gert Schneider†

February 6, 2007

Abstract: We will study convolutional codes with the help of state space realizations. It will be shown that two such minimal realizations belong to the same code if and only if they are equivalent under the full state feedback group. This result will be used in order to prove that two codes with positive Forney indices are monomially equivalent if and only if they share the same adjacency matrix. The adjacency matrix is an invariant of the code obtained via a minimal state space realization and counts in a detailed way the weights of all possible outputs. It contains full information about the weights of the codewords in the given code.

Keywords: Convolutional codes, minimal realizations, weight adjacency matrix, monomial equivalence

MSC (2000): 94B10, 94B05, 93B15, 93B20

1 Introduction and Preliminaries

In the theory of linear block codes MacWilliams' Equivalence Theorem [14, 15] tells us that two block codes are isometric if and only if they are monomially equivalent. Stated more precisely, codes that are related by a weight-preserving isomorphism differ only by permutation and rescaling of the coordinates. It is of crucial importance that the weight-preserving mapping is linear and not just a bijection. Indeed, it is well known that codes with the same weight enumerator need not be monomially equivalent (unless they are one-dimensional). In other words, the weight enumerator does not form a *complete* invariant under monomial equivalence.

In this paper we will show the somewhat surprising result that for a particular class of convolutional codes (not encompassing block codes) a certain generalized weight enumerator does form a complete invariant under monomial equivalence. Thus, two such convolutional codes are monomially equivalent if and only if the share the same generalized weight enumerator.

This generalized weight enumerator will be an adjacency matrix associated with a weighted state transition graph of the code and counts in a very detailed and systematic way the weights of codeword coefficients. It will be introduced in Section 3, and its properties, as found in [5, 6], will be briefly summarized. All that will indicate that it forms an adequate generalization of the classical weight enumerator for block codes. The adjacency matrix is defined via suitable state space realizations of reduced encoders. In this sense, our approach

*University of Kentucky, Department of Mathematics, 715 Patterson Office Tower, Lexington, KY 40506-0027, USA; heidegl@ms.uky.edu

†University of Groningen, Department of Mathematics, P. O. Box 800, 9700 AV Groningen, The Netherlands; schneider@math.rug.nl

follows a series of papers where convolutional codes have been investigated successfully by system-theoretic methods, see, e. g., [10, 24, 25]. Since for a given code neither the reduced encoders nor the associated realizations are unique, we will first discuss in detail the relationship between any two minimal realizations for a given code. This is accomplished in Section 2 by making use of classical realization theory. It turns out that, in essence, two minimal realizations belong to the same code if and only if they are equivalent under the full state feedback group. In Section 3 the weight adjacency matrix associated with a minimal realization will be introduced. The fact that all minimal realizations of a given code are feedback equivalent in the sense above will allow us to turn this matrix into an invariant of the code. Finally, the converse fact, that is, any two feedback equivalent minimal realizations belong to the same code, will lead us to our main result. It states that two convolutional codes with positive Forney indices are monomially equivalent if and only if they share the same adjacency matrix. This result is not true for the class of codes where at least one Forney index is zero; this, of course, includes block codes.

Before further commenting on our result let us first recall the basic notions of coding theory. Let $\mathbb{F}$ be a finite field. A *block code of length* n over $\mathbb{F}$ is, algebraically, just a subspace of $\mathbb{F}^n$. A *convolutional code of length* n is a submodule $\mathcal{C}$ of $\mathbb{F}[z]^n$ of the form

$$\mathcal{C} = \operatorname{im} G := \{uG \,\big|\, u \in \mathbb{F}[z]^k\}$$

where G is a *basic* matrix in $\mathbb{F}[z]^{k\times n}$, i. e.

$$\operatorname{rk} G(\lambda) = k \text{ for all } \lambda \in \overline{\mathbb{F}}, \tag{1.1}$$

with $\overline{\mathbb{F}}$ being an algebraic closure of $\mathbb{F}$. We call such a matrix G an *encoder*, and the number

$$\deg(\mathcal{C}) := \deg(G) := \max\{\deg(M) \mid M \text{ is a } k\text{-minor of } G\} \tag{1.2}$$

is said to be the *degree* of the encoder G or of the code $\mathcal{C}$. It is clear that for two basic matrices G, $G' \in \mathbb{F}[z]^{k\times n}$ one has $\operatorname{im} G = \operatorname{im} G'$ if and only if $G' = UG$ for some $U \in GL_k(\mathbb{F}[z])$. Here $GL_k(\mathbb{F}[z])$ denotes the group of unimodular $k \times k$-matrices over $\mathbb{F}[z]$, i. e., matrices with determinant in $\mathbb{F}^* := \mathbb{F}\backslash\{0\}$. A matrix $G \in \mathbb{F}[z]^{k\times n}$ is said to be *reduced* if the sum of its row degrees equals $\deg(G)$, where the degree of a polynomial row vector is defined as the maximal degree of its entries. For details and characterizations of reducedness see, e. g., [3, Main Thm.] or [18, Thm. A.2]. It is well known [3, p. 495] that each convolutional code admits a reduced encoder. The row degrees of a reduced encoder are, up to ordering, uniquely determined by the code 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 is, in a natural way, a block code.

Beyond these purely algebraic concepts the most important notion in error-control coding is certainly the weight. Recall that for a vector $v = (v_1, \ldots, v_n) \in \mathbb{F}^n$ the *(Hamming) weight* is defined to be $\operatorname{wt}(v) := \#\{i \mid v_i \neq 0\}$. For a polynomial vector $v = \sum_{j=0}^{N} v^{(j)} z^j$, where $v^{(j)} \in \mathbb{F}^n$, one defines its weight as $\operatorname{wt}(v) = \sum_{j=0}^{N} \operatorname{wt}(v^{(j)})$. The *distance* of a (block or convolutional) code $\mathcal{C}$ is defined as $\operatorname{dist}(\mathcal{C}) = \min\{\operatorname{wt}(v) \mid v \in \mathcal{C},\, v \neq 0\}$.

Finally we need to introduce the notion of monomial equivalence for convolutional codes. Motivated by the idea that monomial equivalence should consist of the most obvious transformations that leave invariant all algebraic and coding-theoretically relevant parameters one arrives at the same notion as for block codes.

Definition 1.1 Let $G, G' \in \mathbb{F}[z]^{k\times n}$ be two basic matrices with rank k. We call G, G' *monomially equivalent* if there exists a permutation matrix $P \in GL_n(\mathbb{F})$ and a diagonal matrix $R \in GL_n(\mathbb{F})$ such that $G' = GPR$. The codes $\mathcal{C} = \operatorname{im} G$ and $\mathcal{C}' = \operatorname{im} G'$ are said to be *monomially equivalent* if $G' = WGPR$ for some $W \in GL_k(\mathbb{F}[z])$ and P, R as above. In other words, there exist monomially equivalent encoder matrices for $\mathcal{C}$ and $\mathcal{C}'$.

Notice that we require that, just like for block codes, the rescaling factors (the diagonal elements of R) are constant rather than polynomials. It is obvious that monomially equivalent codes have the same dimension, Forney indices, degree, and distance. They also share the same column distances, extended row distances, and active burst distances. All these are parameters relevant for the error-correcting quality of the code, see [11, Ch.3], [12], and [8]. Furthermore, the mapping $uG \longmapsto uGPR$ is weight-preserving and $\mathbb{F}[z]$-linear and thus monomially equivalent codes are isometric. The isometry is even degree-preserving. Below we will address the issue of isometries for convolutional codes in a bit more detail. It should be observed that, in general, testing whether two codes $\mathcal{C} = \operatorname{im} G$ and $\mathcal{C}' = \operatorname{im} G'$ of the same size are monomially equivalent can be quite a formidable task. Indeed, one has to check whether there exists a unimodular matrix W, a permutation P, and a diagonal matrix R such that $G' = WGPR$. For codes with positive Forney indices our main result provides an alternative test: the coincidence of the adjacency matrices. It depends on the algebraic parameters of the codes which way is more efficient.

We want to close the introduction with briefly addressing the issue of isometry for codes. Recall that two block codes are said to be *isometric* if there exists a weight-preserving isomorphism between them. MacWilliams' Equivalence Theorem (see, e. g., [9, Thm. 7.9.4]) tells us that isometry in this sense coincides with monomial equivalence. This theorem became the cornerstone of the notion of equivalence for block codes and allows us to classify these codes. Since the discovery of the importance of linear block codes over $\mathbb{Z}_4$ for nonlinear codes, the Equivalence Theorem has enjoyed various generalizations to block codes over certain finite rings, see for instance the articles [2, 7, 27, 29].

For convolutional codes an intrinsic coding-theoretic classification has not yet been established. In other words, it is not yet clear as to when two such codes should be identified. It is easily seen that the usual notion of isometry (that is, weight-preserving isomorphism) is too weak in order to yield a reasonable concept of code equivalence. In fact, the block code $\mathcal{C} = \operatorname{im}(1, 1)$ is isometric to the (proper) convolutional code $\mathcal{C}' = \operatorname{im}(z, 1)$. But these codes should certainly not be called equivalent as they have completely different algebraic and coding-theoretic properties. But even if we require, in addition, that the codes share the same degree, isometry will in general change the coding-theoretic properties. For instance, the codes

$$\mathcal{C} := \operatorname{im}(1, 1 + z) \text{ and } \mathcal{C}' := \operatorname{im}(z, 1 + z)$$

are isometric with a degree-preserving isomorphism. Hence they have the same degree and distance. But, again, the codes have different error-correcting capabilities, as can be seen from their column distances. These distances are refined parameters relevant for error-control via sequential decoding, see [11, pp. 110]. In this case, the codes have different zeroth column distances, which simply reflects the fact that the lowest nonzero coefficient of each codeword in the first code always has weight 2 while for the second code this is 1. As a consequence, the codes should not be identified.

It should be noticed that in both examples the codes $\mathcal{C}$ and $\mathcal{C}'$ are not monomially equivalent in the sense of Definition 1.1. However, they are equivalent under rescaling the first entry by the factor z.

We believe that our main theorem will be helpful in order to establish an appropriate notion of equivalence for convolutional codes. It should be clear that a reasonable notion should involve those isometries that leave all error-correcting properties of the code invariant. Since the adjacency matrix uniquely determines many (if not all) of the parameters characterizing these properties [5, Sec. 3] we believe that it is reasonable to require that this matrix be invariant under code equivalence. In this sense our main theorem can be regarded as a generalization of MacWilliams' Equivalence Theorem to convolutional codes with positive Forney indices. However, the result does not tell us how an intrinsic notion of code equivalence should look like, and we have to leave this open for future research.

2 State Space Descriptions of Reduced Encoders

In this section we will study state space descriptions of convolutional codes and discuss their non-uniqueness. This section can be regarded as a recollection of certain results from classical linear systems theory applied to the particular situation of coding theory.

Let us fix a code $\mathcal{C} = \operatorname{im} G$ and concentrate on the encoding process

$$G : \mathbb{F}[z]^k \longrightarrow \mathcal{C}, \quad u \longmapsto v := uG \tag{2.1}$$

for various choices of the (reduced) encoder $G \in \mathbb{F}[z]^{k\times n}$. Obviously, the encoding (2.1) can be interpreted as a dynamical input-output system and thus can be described as a state space system in the system theoretic sense. In this section we will describe all possible state space descriptions of a given code $\mathcal{C}$ with minimal state space dimension and investigate their relation to each other. The main issue will be the non-uniqueness of the encoder matrix G. As we will see later on the considerations of this section can directly be deduced from classical realization theory. However, the polynomial rather than proper rational setting and the fact that not the encoder but rather the code is the object under consideration lead to certain differences, and we consider it worth presenting the results and the differences in some detail. In addition, we wish to show how, due to the specific form of our transfer matrices, all assertions can easily be obtained by some matrix algebra. Notice that the results of this section are true for arbitrary fields $\mathbb{F}$ and do not require the finiteness of $\mathbb{F}$.

In order to use standard notation of systems theory it will be most convenient to associate with a given polynomial matrix $G \in \mathbb{F}[z]^{k\times n}$ the proper rational transfer matrix

$$T_G(z) := G(z^{-1}) \in \mathbb{F}(z)^{k\times n}. \tag{2.2}$$

Notice that the transfer function T_G is polynomial in z^{-1}, or, in other words, T_G does not have any poles in $\overline{\mathbb{F}}\backslash\{0\}$. Recall that the *McMillan degree* $\delta_{\mathrm{M}}(T)$ of a proper rational matrix $T \in \mathbb{F}(z)^{k\times n}$ can be defined as $\delta_{\mathrm{M}}(T) := \deg(\det Q)$ where

$$T = Q^{-1}P \text{ is a coprime factorization with matrices } Q \in \mathbb{F}[z]^{k\times k},\ P \in \mathbb{F}[z]^{k\times n}. \tag{2.3}$$

Coprimeness of the factorization $Q^{-1}P$ simply means that the matrix $[Q, P]$ is basic. It is well known that such a factorization always exists (e. g., the Smith-McMillan form), and the McMillan degree does not depend on the choice of the coprime factorization.

Proposition 2.1 *Let $G \in \mathbb{F}[z]^{k\times n}$ be a polynomial matrix and let T_G be as in* (2.2)*. Then $\delta_M(T_G) \geq \deg(G)$. Moreover, if G is reduced then $\delta_M(T_G) = \deg(G)$.*

PROOF: Let $\nu_1, \ldots, \nu_k$ be the row degrees of G and put $\alpha(G) := \max\{\deg(M) \mid M \text{ minor of } G \text{ of any size}\}$. Then obviously $\sum_{i=1}^{k} \nu_i \geq \alpha(G) \geq \deg(G)$, and we have equality at both steps if and only if G is reduced. From [23, Thm. 2(i)] it is known that $\delta_{\mathrm{M}}(T_G) = \alpha(G)$. This yields the desired results. □

The last statement of Proposition 2.1 is not an if-and-only-if statement. This can easily be verified using the matrix G at the end of Remark 2.5 below.

Let us now turn to state space realizations of encoders and recall some well known results from realization theory as to be found, e. g., in [13, Ch. 6]. First of all, each proper rational matrix $T \in \mathbb{F}(z)^{k\times n}$ has a *realization* $(A, B, C, D) \in \mathbb{F}^{\delta\times\delta+k\times\delta+\delta\times n+k\times n}$. In our setting where transfer matrices act on the right, see (2.1), this means that $T(z) = B(zI - A)^{-1}C + D$. Furthermore, $\delta \geq \delta_{\mathrm{M}}(T)$, and $\delta = \delta_{\mathrm{M}}(T)$ if and only if (A, B, C, D) is *controllable* and *observable*, that is, $\mathrm{rk}\,(\lambda I - A^{\mathsf{T}}, B^{\mathsf{T}}) = \delta = \mathrm{rk}\,(\lambda I - A, C)$ for all $\lambda \in \overline{\mathbb{F}}$. Controllable and observable realizations do always exist. They are unique up to *similarity*, that is, given any two such realizations (A, B, C, D) and $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ of T then there exists a matrix $S \in GL_\delta(\mathbb{F})$ such that $(\bar{A}, \bar{B}, \bar{C}, \bar{D}) = (SAS^{-1}, BS^{-1}, SC, D)$.

Assume now that $T = T_G$ for some matrix $G \in \mathbb{F}[z]^{k\times n}$ with full row rank. One can show straightforwardly that any realization (A, B, C, D) of T leads to the equivalence

$$v = uG \Longleftrightarrow \left\{ \begin{array}{rcl} x_{t+1} & = & x_t A + u_t B \\ v_t & = & x_t C + u_t D \end{array} \quad \text{for all } t \geq 0 \right\} \text{ where } x_0 = 0 \tag{2.4}$$

for any $u = \sum_{t\geq 0} u_t z^t \in \mathbb{F}[z]^k$ and $v = \sum_{t\geq 0} v_t z^t \in \mathbb{F}[z]^n$, see also [5, Thm. 2.3]. Due to this interpretation we simply call the quadruple (A, B, C, D) a *(state space) system* over $\mathbb{F}$. This gives rise to the following definition.

Definition 2.2 Let $(A, B, C, D) \in \mathbb{F}^{\delta\times\delta+k\times\delta+\delta\times n+k\times n}$ be a system over $\mathbb{F}$.

(1) Let $G \in \mathbb{F}[z]^{k\times n}$ be a polynomial matrix with full row rank. Then (A, B, C, D) is said to be a *realization of order* δ of G if

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

As usual, the system is called *canonical* if it is controllable and observable.

(2) We call (A, B, C, D) a *realization of the code* $\mathcal{C} \subseteq \mathbb{F}[z]^n$ if there exists an encoder $G \in \mathbb{F}[z]^{k\times n}$ of $\mathcal{C}$ such that (A, B, C, D) is a realization of G. If G is reduced and (A, B, C, D) is a canonical realization of G, then it is said to be a *canonical minimal realization* of $\mathcal{C}$.

Since a realization of G is, by definition, a realization of the proper matrix T_G in the system theoretic sense, it follows from the discussion above that each polynomial matrix G has a realization, and the order of any realization is at least $\delta_{\mathrm{M}}(T_G)$. Each such G also has a canonical realization, and a given realization is canonical if and only if its order equals $\delta_{\mathrm{M}}(T_G)$. Moreover, each code has a canonical minimal realization; it has order $\deg(\mathcal{C})$. However, not each realization with that order is canonical minimal, see Remark 2.5 and Example 2.7 below. Let us also note that in the special case where $\deg(\mathcal{C}) = 0$, i. e., $\mathcal{C}$ is a block code, the matrices A, B, C of a canonical minimal realization do not exist and $D = G$, where G is a constant encoder of $\mathcal{C}$.

We will single out a particularly simple realization of a given encoder. It is a particular instance of the well-known controller form in systems theory, see, e. g., [1, p. 285].

Proposition 2.3 *Let $G \in \mathbb{F}[z]^{k\times n}$ be a polynomial matrix with rank k and row degrees $\nu_1, \ldots, \nu_k$. Put $\delta := \sum_{i=1}^k \nu_i$. Let G have rows $g_i = \sum_{\ell=0}^{\nu_i} g_{i,\ell} z^\ell$, $i = 1, \ldots, k$, where $g_{i,\ell} \in \mathbb{F}^n$. For $i = 1, \ldots, k$ define the matrices*

$$A_i = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{F}^{\nu_i\times\nu_i},\ B_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{F}^{\nu_i},\ C_i = \begin{pmatrix} g_{i,1} \\ \vdots \\ g_{i,\nu_i} \end{pmatrix} \in \mathbb{F}^{\nu_i\times n}.$$

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

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix},\ B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix},\ C = \begin{pmatrix} C_1 \\ \vdots \\ C_k \end{pmatrix},\ D = \begin{pmatrix} g_{1,0} \\ \vdots \\ g_{k,0} \end{pmatrix} = G(0).$$

In the case where $\nu_i = 0$ the ith block is missing and in B a zero row occurs. The following is true.

(i) The controller form (A, B, C, D) forms a controllable realization of the matrix G.

(ii) G is reduced if and only if $\left(\begin{smallmatrix} -A & C \\ -B & D \end{smallmatrix}\right) = \delta + k$.

(iii) If G is reduced, then the controller form is a canonical realization of G.

Proof: Part (i) is proved in [5, Prop. 2.1] and part (ii) can be checked directly[1]. Part (iii) is a consequence of (i) and (ii) since observability means that $\operatorname{rk}(\lambda I - A, C) = \delta$ for all $\lambda \in \overline{\mathbb{F}}$. Due to nilpotency of A this is equivalent to $\operatorname{rk}(-A, C) = \delta$, and that follows from (ii). □

It is well known, and can also straightforwardly be shown, that if the polynomial matrix G is reduced the controller form is the shift realization in the sense of Fuhrmann [4, Thm. 10-1] associated with the coprime factorization

$$T_G = \operatorname{diag}(z^{\nu_1}, \ldots, z^{\nu_k})^{-1}\Big(\sum_{l=0}^{\nu_i} g_{ij}^{(\nu_i - l)} z^l\Big). \tag{2.5}$$

Later on we will need some more detailed properties of canonical minimal realizations. In the next theorem we present a slightly more comprehensive picture than necessary in order to sketch the interplay between realizations and polynomial matrices. Only the very last result will be needed later on. After a direct proof of the theorem we will place it into the context of classical systems theory in Remark 2.5 below.

Theorem 2.4 *Let $(A, B, C, D) \in \mathbb{F}^{\delta\times\delta+k\times\delta+\delta\times n+k\times n}$ be a canonical system and put $G := B(z^{-1}I - A)^{-1}C + D \in \mathbb{F}(z)^{k\times n}$. Then G is a polynomial matrix if and only if A is nilpotent. If A is nilpotent one also has the following.*

(a) G is basic if and only if $\operatorname{rk} D = k$ and $\operatorname{rk}\left(\begin{smallmatrix} \lambda I - A & C \\ -B & D \end{smallmatrix}\right) = \delta + k$ for all $\lambda \in \overline{\mathbb{F}}\backslash\{0\}$.

(b) If G is a reduced polynomial matrix then $\operatorname{rk}\left(\begin{smallmatrix} -A & C \\ -B & D \end{smallmatrix}\right) = \delta + k$.

Summarizing, if G is a basic and reduced polynomial matrix then

$$A \text{ nilpotent}, \quad \operatorname{rk} D = k, \quad \operatorname{rk}\begin{pmatrix} \lambda I - A & C \\ -B & D \end{pmatrix} = \delta + k \text{ for all } \lambda \in \overline{\mathbb{F}}. \tag{2.6}$$

[1] The equivalence given in [5, (2.2)] is false in general. It is true, however, if all row degrees of G are positive.

PROOF: The if-part of the first statement follows immediately from $G = D + \sum_{i=1}^{\infty} BA^{i-1}Cz^i$. The other direction is a well-known result in systems theory, too. Indeed, if $T_G = Q^{-1}P$ is a coprime polynomial factorization then $Q^{-1}P = B(zI - A)^{-1}C + D$ implies $\det(Q) = \det(zI - A)$ up to some nonzero constant, see [13, Thms. 8.3-2 and 8.2-3]. Since G being polynomial yields $\det(Q) = \alpha z^l$ for some $\alpha \in \mathbb{F}^*$, $l \in \mathbb{N}$ we arrive at the nilpotency of A.

(a) By nilpotency of A the matrix $\lambda I - A$ is regular for $\lambda \neq 0$. Thus

$$\operatorname{rk} \begin{pmatrix} \lambda I - A & C \\ -B & D \end{pmatrix} = \operatorname{rk} \begin{pmatrix} \lambda I - A & C \\ 0 & D + B(\lambda I - A)^{-1}C \end{pmatrix} = \operatorname{rk} \begin{pmatrix} \lambda I - A & C \\ 0 & G(\lambda^{-1}) \end{pmatrix}$$

along with $G(0) = D$ completes the proof of (a).

(b) Let G be reduced and consider the controller form (A, B, C, D) of G. Then the required rank condition is satisfied by Proposition 2.3(ii). By part (iii) of that proposition the controller form is canonical. Now (b) follows for arbitrary canonical realizations of G by using the facts that each such realization is similar to the controller form and that the rank of $\left(\begin{smallmatrix} \lambda I - A & C \\ -B & D \end{smallmatrix}\right)$ is invariant under similarity. □

Remark 2.5 Using the transformation $T = T_G$ part (a) of the last theorem is a particular instance of the well-known system theoretic fact that the transmission zeros (i. e., the zeros of the transfer matrix) coincide with the invariant zeros of a canonical realization (i. e., the zeros of the rightmost matrix in (2.6)), see for instance [13, p. 578]. Part (b) reflects the fact that row reduced matrices have no zeros at infinity, see [13, 6.5.-19, p. 468]. Indeed, by definition G has a zero at infinity if T_G has a zero at zero, meaning that $\operatorname{rk} P(0) < k$ for a coprime factorization $Q^{-1}P = T_G$. But if G is reduced then the factorization in (2.5) shows that T_G has no zeros at zero. The converse of Theorem 2.4(b), and thus the converse of this last statement, is not true. This can be seen from the system $(A, B, C, D) = \left(0, \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right), (0, 0, 1), \left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{smallmatrix}\right)\right)$ over $\mathbb{F}_3$. It is canonical and satisfies (2.6), but $G := B(z^{-1}I - A)^{-1}C + D = \left(\begin{smallmatrix} 0 & 1 & 1+2z \\ 1 & 0 & z \end{smallmatrix}\right)$ is not reduced. As a consequence, (A, B, C, D) is not a canonical minimal realization of the code $\mathcal{C} = \operatorname{im} G$. It can also easily be checked that the matrix T_G has no zeros at zero.

Let us now turn to different canonical minimal realizations of a given code (in the sense of Definition 2.2(2)) and present the main result of this section. As we will see in the proof it is an application of a classical result from Wolovich in exact model matching for linear systems. Our specific situation where transfer matrices are polynomial and basic in z^{-1} makes the following formulation possible.

Theorem 2.6 *Let $G, \bar{G} \in \mathbb{F}[z]^{k \times n}$ be basic and reduced and let $\deg(G) = \deg(\bar{G}) = \delta$. Let (A, B, C, D) and $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ be associated canonical realizations, respectively. Then the following are equivalent.*

(i) $G = W\bar{G}$ for some $W \in GL_k(\mathbb{F}[z])$.

(ii) The systems (A, B, C, D) and $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ are equivalent under the full state feedback group, that is, there exist matrices $T \in GL_\delta(\mathbb{F})$, $U \in GL_k(\mathbb{F})$, $M \in \mathbb{F}^{\delta \times k}$ such that

$$\bar{A} = T^{-1}(A - MB)T,\ \bar{B} = UBT,\ \bar{C} = T^{-1}(C - MD),\ \bar{D} = UD. \tag{2.7}$$

PROOF: (ii) ⇒ (i): Define the $k \times k$-matrix $V := I + B(z^{-1}I - A)^{-1}M$. From systems theory it is well known [1, p. 346, Eq. (2.43)] that

$$B(z^{-1}I - A)^{-1}C + D = VU^{-1}\big(UB(z^{-1}I - A + MB)^{-1}(C - MD) + UD\big), \tag{2.8}$$

thus $G = VU^{-1}\bar{G}$. Due to nilpotency of A the matrix V is polynomial. But then $W := VU^{-1}$ is even unimodular since G and $\bar{G}$ are both basic. This yields (i).
(i) $\Rightarrow$ (ii): Also this implication can directly be deduced from linear systems theory. Indeed, the identity $G = W\bar{G}$ implies $T_G = T_W T_{\bar{G}}$ for the associated transfer matrices. This can be read as exact model matching of the system G from the system $\bar{G}$. In the paper [28] it has been characterized (at least for systems where all Forney indices are positive) as to when exact model matching can be realized via the full state feedback group. It is lengthy but straightforward to check that the sufficient and necessary conditions given in [28, Thm., p. 519] are satisfied in our situation. This shows the implication (i) $\Rightarrow$ (ii).
However, in our very specific situation the proof simplifies considerably and we think it is worth giving a direct argument. In order to do so first notice that equivalence under the full state feedback group is indeed an equivalence relation. Since the controller form of a reduced matrix is canonical and all canonical realizations are similar, we may assume without loss of generality that both (A, B, C, D) and $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ are in controller form. Assumption (i) implies that G and $\bar{G}$ have the same row degrees. Since reordering of the rows of G retains the specific requirements of the controller form we may further assume that G and $\bar{G}$ both have row degrees $\nu_1 \geq \ldots \geq \nu_k$. Then $A = \bar{A}$ and $B = \bar{B}$ since they are both fully determined by the row degrees. Due to reducedness of G and $\bar{G}$ the ith row of W has degree at most ν_i for $i = 1, \ldots, k$, see [3, Main Thm. (4)]. We will show now that

$$W = \big(I + B(z^{-1}I - A)^{-1}M\big)U^{-1} \text{ for some } M \in \mathbb{F}^{\delta\times k},\ U \in GL_k(\mathbb{F}). \tag{2.9}$$

We certainly have to put $U := W(0)^{-1}$ and need to find M such that $B(z^{-1}I - A)^{-1}M = WU - I$. The latter matrix is of the form $WU - I = \Big(\sum_{j=1}^{\nu_i} a_{ij}z^j\Big)_{i=1,\ldots,k}$ for suitable $a_{ij} \in \mathbb{F}^k$. Using that $B(z^{-1}I - A)^{-1} = \operatorname{diag}\big(\begin{pmatrix} z & z^2 & \cdots & z^{\nu_i}\end{pmatrix}\big)_{i=1,\ldots,k} \in \mathbb{F}[z]^{k\times\delta}$ one sees that the matrix $M = (M_1, \ldots, M_k)^{\mathsf{T}}$ where $M_i = ({a_{i1}}^{\mathsf{T}}, \ldots, {a_{i\nu_i}}^{\mathsf{T}})$, satisfies (2.9). Notice that if $\nu_i = 0$ the result is true as well since in that case the ith block of M is missing and a zero row appears in $WU - I$ and $B(z^{-1}I - A)^{-1}$. Now we have the identity $G = VU^{-1}\bar{G}$ where, again, $V = I + B(z^{-1}I - A)^{-1}M$. Using (2.8) this reads as

$$UB(z^{-1}I - A + MB)^{-1}(C - MD) + UD = B(z^{-1}I - A)^{-1}\bar{C} + \bar{D} = \bar{G}(z). \tag{2.10}$$

Hence $(A - MB, UB, C - MD, UD)$ is a realization of $\bar{G}$ of order $\deg(\bar{G})$ and therefore canonical. As a consequence, (2.10) implies that the realizations $(A-MB, UB, C-MD, UD)$ and $(A, B, \bar{C}, \bar{D})$ are similar, and this yields (ii). □

The result just proven tells us that two canonical minimal realizations of a given code are equivalent under the full state feedback group. One should bear in mind, however, that the action of the full state feedback group does in general not preserve the property of being canonical minimal. This is being illustrated by the following example.

Example 2.7 Let $\mathbb{F} = \mathbb{F}_2$. Then $(A, B, C, D) := \Big(\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1&1\\0&0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0&1\\0&1&0\end{smallmatrix}\right)\Big)$ is canonical. Moreover, $G = B(z^{-1}I - A)^{-1}C + D = \left(\begin{smallmatrix}1&z&1+z\\0&1&z\end{smallmatrix}\right)$ is basic and reduced. Using the feedback $M = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)$ and $T = U = I_2$ the system $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ in (2.7) leads to a nilpotent matrix $\bar{A}$ and a non-reduced encoder matrix $\bar{G} = \bar{B}(z^{-1}I - \bar{A})^{-1}\bar{C} + \bar{D} = \left(\begin{smallmatrix}1&z&1+z\\z&1+z^2&z^2\end{smallmatrix}\right)$. Hence the realization (A, B, C, D) of the code $\mathcal{C}$ is canonical minimal while $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ is not.

The last example and Proposition 2.1 suggest that the requirement of reducedness for encoders seems too strong for this type of considerations. Indeed, the results of this section

become somewhat more elegant if we replace reducedness by semi-reducedness where we call a matrix $G \in \mathbb{F}[z]^{k\times n}$ *semi-reduced* if $\delta_{\mathrm{M}}(T_G) = \deg(G)$. It is straightforward to show that the results remain true even for semi-reduced encoders and Proposition 2.1 as well as Theorem 2.4(b) become if-and-only-if statements. We omit the details.

3 The Weight Adjacency Matrix and Monomial Equivalence

In this section we will return to the particular situation of convolutional codes as dynamical systems over finite fields. Thus from now on let

$$\mathbb{F} = \mathbb{F}_q \text{ be a finite field with } q \text{ elements.} \tag{3.1}$$

Recall the weight of constant and polynomial vectors of length n from the introduction. We will need the *weight enumerator* of sets $S \subseteq \mathbb{F}^n$ given as

$$\mathrm{we}(S) := \sum_{i=0}^{n} \lambda_i W^i \in \mathbb{Z}[W], \text{ where } \lambda_i := \#\{v \in S \mid \mathrm{wt}(v) = i\}. \tag{3.2}$$

The weight enumerator $\mathrm{we}(\mathcal{C})$ of a block code $\mathcal{C} \subseteq \mathbb{F}^n$ has been investigated intensively in the block coding literature. For instance, the famous MacWilliams Identity Theorem [15] tells us how to completely derive $\mathrm{we}(\mathcal{C}^\perp)$ from $\mathrm{we}(\mathcal{C})$, where $\mathcal{C}^\perp$ is the dual of $\mathcal{C}$ with respect to the standard inner product on $\mathbb{F}^n$.

In order to introduce an appropriate generalization of the weight enumerator for convolutional codes we need a state space realization. Consider the system in (2.4). Due to (3.1) the system has q^δ different state vectors x_t where δ is the order of the realization (A, B, C, D). We consider now for each pair of states $(X, Y) \in \mathbb{F}^{2\delta}$ all (finitely many) state transitions from $x_t = X$ to $x_{t+1} = Y$ via suitable input $u_t = u$ and count the weights of all corresponding outputs $v = XC + uD$. This leads to the following definition, see also [19, Sec. 2] and [5, Def. 3.4].

Definition 3.1 Let $G \in \mathbb{F}[z]^{k\times n}$ be a basic and reduced matrix such that $\deg(G) = \delta$ and let (A, B, C, D) be a canonical realization of G. We call $\mathbb{F}^\delta$ the *state space* of the realization. The *weight adjacency matrix associated with* (A, B, C, D) is defined to be the matrix $\Lambda \in \mathbb{Z}[W]^{q^\delta \times q^\delta}$ that is indexed by the state pairs $(X, Y) \in \mathbb{F}^{2\delta}$ and has the entries

$$\Lambda_{X,Y} := \mathrm{we}\{XC + uD \mid u \in \mathbb{F}^k : Y = XA + uB\} \in \mathbb{Z}[W] \text{ for } (X, Y) \in \mathbb{F}^{2\delta}. \tag{3.3}$$

Recall that in the case where $\delta = 0$ the matrices A, B, C do not exist while $D = G$. As a consequence, $\Lambda = \Lambda_{0,0} = \mathrm{we}(\mathcal{C})$ is the ordinary weight enumerator of the block code $\mathcal{C} = \{uG \mid u \in \mathbb{F}^k\}$.

The weight adjacency matrix is the adjacency matrix of the weighted state-transition diagram as considered in [11, Sec. 3.10] and [19, Sec. 2]. Its properties have been studied in detail in the papers [5] and [6]. Among other things it has been discussed in detail in [5, Sec. 3] that the weight adjacency matrix contains full information about the extended row distances and the active burst distances of the convolutional code $\mathcal{C} = \mathrm{im}\, G$. These parameters are closely related to the error-correcting performance of $\mathcal{C}$ and are studied intensively in the more engineering-oriented literature, see, e. g., [12, 8]. In a slightly different form the weight adjacency matrix appears also in other papers on convolutional coding theory, see, e. g., [11, Sec. 3.10]. It is mainly used to compute the path weight enumerator (cf. [11, pp. 154])

counting the number of atomic codewords (fundamental paths in the state diagram) of given weight and length. As has been shown in [26] (see also [11, Thm. 4.2]) the path weight enumerator yields an upper bound for the burst error probability of the convolutional code used on a binary symmetric channel with maximum-likelihood decoding. In the paper [6], alternative formulas for the entries of the weight adjacency matrix are given. They are used in order to formulate a conjecture for a MacWilliams Identity for convolutional codes and their duals which then is proven in special cases. All this makes sense only because the weight adjacency matrix can indeed nicely be turned into an invariant of the code. This will be shown in the discussion leading to Definition 3.5 below.

Remark 3.2 In the literature on convolutional codes also the notion of *extended* path weight enumerator has been introduced. It is obtained by not only counting codeword weights, but also keeping track of the weights of the associated message words (see [11, pp. 156] or [17, pp. 215]). It leads to bounds of certain error probabilities concerning original vs. decoded message. A lot of effort has been made in order to efficiently compute the extended path weight enumerator for a given code, see [20, 21, 22]. For our considerations however, it needs to be pointed out that the extended path weight enumerator as well as the associated extended weight adjacency matrix do not lead to an invariant of the code but rather depend on the encoder matrix. Since we are studying code properties, and not encoder properties, we will not pursue this approach.

Example 3.3 Let

$$G = \begin{pmatrix} z & 1+z^2 & 1+z & z+z^2 \\ 1 & 0 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2[z]^{2\times 4}.$$

Then G is basic and reduced and the controller form is given by

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

In order to explicitly display the weight adjacency matrix we need to fix an ordering on the state space. Let us choose the lexicographic ordering, hence $X_1 = (0,0)$, $X_2 = (0,1)$, $X_3 = (1,0)$, $X_4 = (1,1)$. Going through all possible combinations of states X and inputs u one obtains the weight adjacency matrix

$$\Lambda = \left(\Lambda_{X_i,X_j}\right)_{i,j=1,\ldots,4} = \begin{pmatrix} 1+W^3 & 0 & W^2+W^3 & 0 \\ W^2+W^3 & 0 & W+W^2 & 0 \\ 0 & 1+W^3 & 0 & W^2+W^3 \\ 0 & W^2+W^3 & 0 & W+W^2 \end{pmatrix},$$

where the entry at position (i,j) is Λ_{X_i,X_j} as defined in (3.3). For instance, the entry Λ_{X_3,X_2} is obtained as follows. The state equation $X_2 = X_3A + uB$ together with $X_2 = (0,1)$ and $X_3 = (1,0)$ yields the two options $u_1 = (0,0)$ and $u_2 = (0,1)$ for the input. This then leads to the two outputs $v_1 = (1,0,1,1)$ and $v_2 = (0,0,0,0)$, and thus to the weight enumerator $\Lambda_{X_3,X_2} = W^3 + 1$.

The weight adjacency matrix does not form an invariant of a code but rather depends on the choice of both the reduced encoder and the canonical realization. This dependence, however, can nicely be described.

Theorem 3.4 *Let $\mathcal{C} \subseteq \mathbb{F}[z]^n$ be a code of degree δ, and let (A, B, C, D) and $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ both be canonical minimal realizations of $\mathcal{C}$. Furthermore, let Λ and $\bar{\Lambda}$ be the associated weight adjacency matrices, respectively. Then there exists a state space isomorphism $T \in GL_\delta(\mathbb{F})$ such that*

$$\Lambda_{X,Y} = \bar{\Lambda}_{XT,YT} \text{ for all } (X, Y) \in \mathbb{F}^{2\delta}. \tag{3.4}$$

In particular, $\bar{\Lambda} = P\Lambda P^{-1}$ for some permutation matrix $P \in GL_{q^\delta}(\mathbb{Z})$.

The result appeared first in [5, Remark 3.6, Theorem 4.1]. Using Theorem 2.6 we can give an alternative, very short proof for this theorem. Indeed, by Theorem 2.6 the two realizations are equivalent under the full feedback group, thus we may assume (2.7). But then one can straightforwardly check that for any $(X, Y, u, v) \in \mathbb{F}^{2\delta+k+n}$

$$Y = XA + uB,\ v = XC + uD$$

is equivalent to

$$YT = XT\bar{A} + (uU^{-1} + XMU^{-1})\bar{B},\ v = XT\bar{C} + (uU^{-1} + XMU^{-1})\bar{D}.$$

Since for any given X the mapping $u \longmapsto uU^{-1} + XMU^{-1}$ is bijective on $\mathbb{F}^k$, Equation (3.4) is immediate from the definition of the weight adjacency matrix.

The result above shows that we obtain an invariant of the code after factoring out the effect of the state space isomorphisms. This brings us to the following weight counting invariant for convolutional codes. It has been introduced first in [5, p. 314].

Definition 3.5 Let $\mathcal{C} \subseteq \mathbb{F}[z]^n$ be a code of degree δ and let Λ be the weight adjacency matrix associated with a canonical minimal realization of $\mathcal{C}$. We call

$$\bar{\Lambda}(\mathcal{C}) := \{\Lambda' \mid \exists\, T \in GL_\delta(\mathbb{F}) :\ \Lambda'_{X,Y} = \Lambda_{XT,YT} \text{ for all } (X, Y) \in \mathbb{F}^{2\delta}\}$$

the *adjacency matrix* of $\mathcal{C}$.

Let us now study the adjacency matrix with respect to transformations of monomial equivalence as introduced in Defintion 1.1. The following result is easy to see; it appeared first in [5, Thm. 4.4].

Proposition 3.6 *Let $\mathcal{C}$ and $\mathcal{C}'$ be monomially equivalent codes. Then $\bar{\Lambda}(\mathcal{C}) = \bar{\Lambda}(\mathcal{C}')$.*

The main result of this paper states that under a certain condition on the Forney indices the converse of the statement above is true as well. That is, the adjacency matrix even forms a *complete* invariant for monomial equivalence. Indeed, we have the following result.

Theorem 3.7 *Let $\mathcal{C}, \mathcal{C}' \subseteq \mathbb{F}[z]^n$ be two codes and assume that all Forney indices of $\mathcal{C}$ are positive. Then $\mathcal{C}$ and $\mathcal{C}'$ are monomially equivalent if and only if $\bar{\Lambda}(\mathcal{C}) = \bar{\Lambda}(\mathcal{C}')$.*

Notice that we require that $\mathcal{C}$ and $\mathcal{C}'$ are defined over the same field $\mathbb{F}$ and have the same length n. Just like in block coding theory we consider this a reasonable assumption for this kind of considerations. In the proof we will see that if $\bar{\Lambda}(\mathcal{C}) = \bar{\Lambda}(\mathcal{C}')$ the codes $\mathcal{C}$ and $\mathcal{C}'$ have the same dimension and Forney indices. Thus the assumption above on the Forney indices is true for $\mathcal{C}'$ as well.

Remembering that the adjacency matrix can be regarded as a generalization of the weight enumerator of block codes (see the paragraph right after Definition 3.1) this result comes

somewhat surprisingly. Indeed, there exist block codes sharing the same weight enumerator but are not monomially equivalent; see also Example 3.8(a) at the end of this section. This shows that the positivity of the Forney indices is certainly a necessary condition for the above result to be true.

PROOF: The only-if part is in Proposition 3.6. Thus let us assume that $\bar{\Lambda}(\mathcal{C}) = \bar{\Lambda}(\mathcal{C}')$. The outline of the proof is as follows. We will consider canonical minimal realizations of the two codes and show that the identity $\bar{\Lambda}(\mathcal{C}) = \bar{\Lambda}(\mathcal{C}')$ will imply that these realizations are equivalent under the full state feedback group followed by reordering and rescaling of the output coordinates. With the aid of Theorem 2.6 we then can conclude that the two associated encoder matrices satisfy an identity of the form $G' = WGPR$ for some unimodular matrix W and permutation and rescaling matrices P, R. This tells us that the codes are monomially equivalent. We proceed in several steps.

1) We first study the algebraic parameters of the codes and fix suitable realizations. Since the adjacency matrices have the same size, the two codes have the same degree, say δ. Let G, G' be any basic and reduced encoder matrices of $\mathcal{C}$ and $\mathcal{C}'$ and (A, B, C, D) and (A', B', C', D') be the corresponding controller forms, respectively. Then the two systems have order δ and, according to Proposition 2.3, they form canonical minimal realizations of the codes $\mathcal{C}$ and $\mathcal{C}'$. Let Λ and Λ' be the associated weight adjacency matrices. By assumption there exist some $T \in GL_\delta(\mathbb{F})$ such that

$$\Lambda'_{X,Y} = \Lambda_{XT,YT} \text{ for all } (X, Y) \in \mathbb{F}^{2\delta}. \tag{3.5}$$

In [5, Thm. 5.1] is has been proven that codes satisfying (3.5) have the same dimension and the same Forney indices. Thus let $k := \dim(\mathcal{C}) = \dim(\mathcal{C}')$. Using Theorem 3.4 we may assume that both codes have their Forney indices in the same ordering. Let us denote them by $\nu_1 \geq \ldots \geq \nu_k \geq 1$. Notice that $\delta = \sum_{i=1}^{k} \nu_i$. Now the controller form implies $A' = A$ and $B' = B$.

2) Next we will show that

$$A = T(A - MB)T^{-1} \text{ and } B = UBT^{-1} \text{ for some matrices } M \in \mathbb{F}^{\delta \times k},\ U \in GL_k(\mathbb{F}). \tag{3.6}$$

By definition of the weight adjacency matrix we have for any $(X, Y) \in \mathbb{F}^{2\delta}$

$$Y - XA \in \operatorname{im} B \iff \Lambda'_{X,Y} \neq 0 \iff \Lambda_{XT,YT} \neq 0 \iff YT - XTA \in \operatorname{im} B.$$

Putting $\tilde{A} = TAT^{-1}$, $\tilde{B} = BT^{-1}$, we thus get

$$Y - XA \in \operatorname{im} B \iff Y - X\tilde{A} \in \operatorname{im} \tilde{B}.$$

Using $X = 0$ this implies $\operatorname{im} \tilde{B} = \operatorname{im} B$ and hence $BT^{-1} = \tilde{U}B$ for some $\tilde{U} \in GL_k(\mathbb{F})$. On the other hand, for each $X \in \mathbb{F}^\delta$ there exists $u \in \mathbb{F}^k$ and $Y \in \mathbb{F}^\delta$ such that $Y - XA = uB$, hence there exists $\tilde{u} \in \mathbb{F}^k$ such that $Y - X\tilde{A} = \tilde{u}B$. This implies $X(\tilde{A} - A) = (u - \tilde{u})B$. Using for X all standard basis vectors we obtain the identity $\tilde{A} = A + \tilde{M}B$ for some matrix $\tilde{M} \in \mathbb{F}^{\delta \times k}$. Hence we arrive at $A = T^{-1}(A + \tilde{M}B)T$ and $B = \tilde{U}BT$. This in turn yields (3.6).

3) In this step we will prove that (A, B, C', D') and (A, B, C, D) are related via the full state feedback group followed by reordering and rescaling of the output coordinates, see (3.8) below. In order to do so we will compare the entries of the weight adjacency matrices. Consider the canonical minimal realization $(\bar{A}, \bar{B}, \bar{C}, \bar{D}) = (TAT^{-1}, BT^{-1}, TC, D)$ of the code $\mathcal{C}$. It is easy to see [5, Rem. 3.6] that the associated weight adjacency matrix $\bar{\Lambda}$ satisfies $\bar{\Lambda}_{X,Y} = \Lambda_{XT,YT}$ for all $(X, Y) \in \mathbb{F}^{2\delta}$ and hence Equation (3.5) implies

$$\bar{\Lambda} = \Lambda'.$$

Now we can study the entries of these weight adjacency matrices. Since all Forney indices are positive, the matrix B has full rank k (see the controller form). As a consequence, for each pair of states $(X, Y) \in \mathbb{F}^{2\delta}$ the set $\{XC' + uD' \mid u \in \mathbb{F}^k : Y = XA + uB\}$ has at most one element. Recalling the definition of the weight adjacency matrix in (3.3) one obtains that the nonzero entries are given by

$$\Lambda'_{X,XA+uB} = \bar{\Lambda}_{X,XA+uB} \text{ for all } (X, u) \in \mathbb{F}^\delta \times \mathbb{F}^k, \tag{3.7}$$

and these entries have the value $\Lambda'_{X,XA+uB} = W^\alpha$ where $\alpha = \mathrm{wt}(XC' + uD')$. On the other hand notice that, due to (3.6), for any $(X, u) \in \mathbb{F}^\delta \times \mathbb{F}^k$ we have

$$XA + uB = X(TAT^{-1} - TMBT^{-1}) + uUBT^{-1} = X\bar{A} + \bar{u}\bar{B} \text{ where } \bar{u} = uU - XTM.$$

Thus (3.3) yields $\bar{\Lambda}_{X,XA+uB} = \bar{\Lambda}_{X,X\bar{A}+\bar{u}\bar{B}} = W^\beta$ where $\beta = \mathrm{wt}(X\bar{C} + \bar{u}\bar{D})$. As a consequence, (3.7) implies

$$\mathrm{wt}\left((X, u)\begin{pmatrix} C' \\ D' \end{pmatrix}\right) = \mathrm{wt}\big(X\bar{C} + (uU - XTM)\bar{D}\big) = \mathrm{wt}\left((X, u)\begin{pmatrix} \bar{C} - TM\bar{D} \\ U\bar{D} \end{pmatrix}\right)$$

for all $(X, u) \in \mathbb{F}^\delta \times \mathbb{F}^k$. Now [5, Lemma 5.4], which is basically MacWilliams' Equivalence Theorem for block codes, yields the existence of a permutation matrix $P \in GL_n(\mathbb{F})$ and a nonsingular diagonal matrix $R \in GL_n(\mathbb{F})$ such that

$$\begin{pmatrix} C' \\ D' \end{pmatrix} = \begin{pmatrix} \bar{C} - TM\bar{D} \\ U\bar{D} \end{pmatrix} PR.$$

With the help of (3.6) we see that the realization (A, B, C', D') of $\mathcal{C}'$ is of the form

$$\left.\begin{aligned} (A, B, C', D') &= (T(A - MB)T^{-1}, UBT^{-1}, (\bar{C} - TM\bar{D})PR, U\bar{D}PR) \\ &= (T(A - MB)T^{-1}, UBT^{-1}, T(C - MD)PR, UDPR). \end{aligned}\right\} \tag{3.8}$$

4) Now we can apply Theorem 2.6 and obtain for the associated encoder matrices

$$G' = WGPR \text{ for some } W \in GL_k(\mathbb{F}[z]).$$

Thus $\mathcal{C} = \mathrm{im}\, G$ and $\mathcal{C}' = \mathrm{im}\, G'$ are monomially equivalent. This completes the proof. □

Let us briefly mention an immediate consequence of the theorem above. Indeed, for codes $\mathcal{C}$, $\mathcal{C}' \subseteq \mathbb{F}[z]^n$ with the assumptions as in the theorem we have the implication

$$\bar{\Lambda}(\mathcal{C}) = \bar{\Lambda}(\mathcal{C}') \Longrightarrow \bar{\Lambda}(\mathcal{C}^\perp) = \bar{\Lambda}(\mathcal{C}'^\perp)$$

where $\mathcal{C}^\perp := \{w \in \mathbb{F}[z]^n \mid wv^\mathsf{T} = 0 \text{ for all } v \in \mathcal{C}\}$ is the dual code of $\mathcal{C}$. This follows directly from Theorem 3.7 and Proposition 3.6 along with the fact that if $\mathcal{C}$ and $\mathcal{C}'$ are monomially equivalent then so are their dual codes. As a consequence, the adjacency matrix of a code with solely positive Forney indices completely determines the adjacency matrix of its dual. In the paper [6] it is shown that this is true for a much bigger class of codes and a concrete formula is given for computing $\bar{\Lambda}(\mathcal{C}^\perp)$ from $\bar{\Lambda}(\mathcal{C})$. It forms a generalization of the famous MacWilliams' Identity for the weight enumerators of block codes, see [16, p. 146, Thm. 13].

We close the paper with presenting some examples showing that the theorem above is not true if some of the Forney indices are zero.

Example 3.8

(a) Recall that for a block code $\mathcal{C} = \operatorname{im} G$, thus $G \in \mathbb{F}^{k\times n}$, the adjacency matrix is the ordinary weight enumerator. In this case it is well known that block codes with the same weight enumerator are, in general, not monomially equivalent. The following example is taken from [9, Exa. 1.6.1]. The matrices

$$G_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \; G_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{3\times 6}$$

generate codes with the same weight enumerator $1 + 3W^2 + 3W^4 + W^6$, but are not monomially equivalent. This can be seen as follows. Since there is no non-trivial rescaling over the field $\mathbb{F}_2$ monomial equivalence of the two codes is the same as $G_2 = UG_1P$ for some $U \in GL_3(\mathbb{F}_2)$ and a permutation matrix P. As a consequence, $G_2G_2^{\mathsf{T}} = UG_1G_1^{\mathsf{T}}U^{\mathsf{T}}$. But this is a contradiction since $G_1G_1^{\mathsf{T}} = 0 \neq G_2G_2^{\mathsf{T}}$.

(b) From the previous data one can also construct an example with positive degree. Using the rows of the matrices above in a suitable way one obtains

$$G = \begin{pmatrix} 1 & 1 & z & z & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad \bar{G} = \begin{pmatrix} z+1 & 1 & z & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2[z]^{2\times 6}.$$

Both matrices are basic and reduced. The weight adjacency matrices of the associated controller forms are both given by

$$\Lambda = \begin{pmatrix} 1+W^6 & W^2+W^4 \\ W^2+W^4 & W^2+W^4 \end{pmatrix}.$$

But the codes $\mathcal{C} = \operatorname{im} G$ and $\bar{\mathcal{C}} = \operatorname{im} \bar{G}$ are not monomially equivalent. This can be seen by computing UG for all $U \in GL_2(\mathbb{F}_2[z])$ such that UG is reduced with indices 1 and 0 again. The only options are

$$U \in \left\{ I_2, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1+z \\ 0 & 1 \end{pmatrix} \right\}$$

and it is seen by inspection that in none of these cases UG has, up to ordering, the same columns as $\bar{G}$ (again,over $\mathbb{F}_2$ we can disregard rescaling matrices).

Conclusion

In this note we have shown that codes with all Forney indices being positive are monomially equivalent if and only if they share the same adjacency matrix. Hence this matrix forms a complete invariant under monomial equivalence for this class of codes. The result is not true for codes with at least one Forney index being zero (unless they are one-dimensional block codes). The adjacency matrix contains in a detailed way information about the error-correcting quality of the code in question. It has to remain open for future research if there is a way to generalize this result to arbitrary convolutional codes. In this context the investigation of isometries with further coding-theoretically meaningful properties should play a role. Once a concept has been established the question whether the adjacency matrix forms an invariant under such isometries needs to be addressed.

Acknowledgments: We would like to thank the reviewers for their careful reading of the first version and their valuable comments that helped to improve the readability of the paper.

References

[1] P. J. Antsaklis and A. N. Michel. *Linear Systems*. McGraw-Hill, New York, 1997.

[2] H. Q. Dinh and S. R. López-Permouth. On the equivalence of codes over rings and modules. *Finite Fields & their Appl.*, 10:615–625, 2004.

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

[4] P. A. Fuhrmann. *Linear Systems and Operators in Hilbert space*. McGraw-Hill, New York, 1981.

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

[6] H. Gluesing-Luerssen and G. Schneider. On the MacWilliams identity for convolutional codes. Preprint 2006. Submitted. Available at http://arxiv.org/pdf/cs.IT/0603013.

[7] M. Greferath and S. E. Schmidt. Finite ring combinatorics and MacWilliams' Equivalence Theorem. *J. Combin. Theory Ser. A*, 92:17–28, 2000.

[8] S. Höst, R. Johannesson, and V. V. Zyablov. Woven convolutional codes I: Encoder properties. *IEEE Trans. Inform. Theory*, IT-48:149–161, 2002.

[9] W. C. Huffman and V. Pless. *Fundamentals of Error-Correcting Codes*. Cambridge University Press, Cambridge, 2003.

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

[11] R. Johannesson and K. S. Zigangirov. *Fundamentals of Convolutional Coding*. IEEE Press, New York, 1999.

[12] J. Justesen, E. Paaske, and M. Ballan. Quasi-cyclic unit memory convolutional codes. *IEEE Trans. Inform. Theory*, IT-36:540–547, 1990.

[13] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

[14] F. J. MacWilliams. *Combinatorial problems of elementary abelian groups*. PhD thesis, Harvard University, 1962.

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

[16] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, 1977.

[17] R. J. McEliece. *The Theory of Information and Coding*. Addison-Wesley Publishing Company, 1977.

[18] R. J. McEliece. The algebraic theory of convolutional codes. In V. Pless and W. Huffman, editors, *Handbook of Coding Theory, Vol. 1*, pages 1065–1138. Elsevier, Amsterdam, 1998.

[19] R. J. McEliece. How to compute weight enumerators for convolutional codes. In M. Darnell and B. Honory, editors, *Communications and Coding (P. G. Farrell 60th birthday celebration)*, pages 121–141. Wiley, New York, 1998.

[20] M. Motani and C. Heegard. Computing weight distributions of convolutional codes via shift register synthesis. In *Applied Algebra, Algorithms and Error-Correcting Codes; 13th Intern. Symp. AAECC-13 (Honolulu/USA). Lecture Notes in Computer Science LN 1719*, pages 314–323. Springer, 1999.

[21] I. Onyszchuk. Finding the complete path and weight enumerators of convolutional codes. JPL TDA Progress Report 42-100, 1990.

[22] C. Pimentel. On the computation of weight enumerators for convolutional codes. *IEEE Trans. Commun.*, 51:313–317, 2003.

[23] A. C. Pugh. The McMillan degree of a polynomial system matrix. *Int. J. Contr.*, 24:129–135, 1976.

[24] J. Rosenthal. Connections between linear systems and convolutional codes. In B. Marcus and J. Rosenthal, editors, *Codes, Systems, and Graphical Models*, pages 39–66. Springer, Berlin, 2001.

[25] J. Rosenthal and E. V. York. BCH convolutional codes. *IEEE Trans. Inform. Theory*, IT-45:1833–1844, 1999.

[26] A. J. Viterbi. Convolutional codes and their performance in communication systems. *IEEE Trans. Commun. Technol.*, COM-19:751–772, 1971.

[27] H. N. Ward and J. A. Wood. Characters and the equivalence of codes. *J. Combin. Theory Ser. A*, 73:348–352, 1996.

[28] W. A. Wolovich. The use of state feedback for exact model matching. *SIAM J. Contr. & Opt.*, 10:512–523, 1972.

[29] J. A. Wood. Weight functions and the extension theorem for linear codes over finite rings. *Contemp. Math.*, 225:231–243, 1999.