Algebraic aspects of 2D convolutional codes

Ettore Fornasini and Maria Elena Valcher
Dept. of Electronics and Comp. Sci., University of Padova
Via Gradenigo 6/a, 35131 PADOVA, Italy

Abstract

Two-dimensional (2D) codes are introduced as linear shift-invariant spaces of admissible signals on the discrete plane. Convolutional and, in particular, basic codes are characterized both in terms of their internal properties and by means of their input-output representations. The algebraic structure of the class of all encoders that correspond to a given convolutional code is investigated and the possibility of obtaining 2D decoders, free from catastrophic errors, as well as efficient syndrome decoders is considered.
Some aspects of the state space implementation of 2D encoders and decoders via (finite memory) 2D systems are finally discussed.

Keywords: encoders and decoders of 2D sequences, dual codes, behaviours, 2D state models, inverse systems.

1 Introduction

The algebraic theory of 1D convolutional codes was originated by G.D.Forney in a noteworthy paper of 1970 [1]. Employing the same polynomial matrix techniques utilized in researches on multivariable linear systems, Forney laid on firm foundations the notions of equivalence, minimality and duality of convolutional encoders and showed how one could apply the state space realization methods for implementing a code in a transmission chain.

In recent times the extension of the above techniques to polynomial matrices in two variables [2 ÷ 5], guaranteed a fairly good understanding of their algebraic properties and made possible two significant advances in 2D signal modelling and realization, which seem very promising for applications in multidimensional data coding.
The first such development is the behavioural approach, introduced by J.C.Willems and P.Rocha [6÷8] in the description of the admissible 2D systems trajectories. This approach, indeed, allows to investigate the recursive structure of the codes without making any a priori assumption on the direction of the recursion and, consequently, on the specific kind of causality the encoding process refers to. Moreover, once a convolutional code $\mathcal{C}$ has been selected on the basis of some internal requirements (such as the reliability of the transmitted message, the minimal distance between two distinct codewords, etc.), it is possible to provide a complete description of all Laurent polynomial encoders which produce $\mathcal{C}$, and find out among them the most efficient ones.

The second major development is the introduction of 2D finite memory systems [9÷11], which constitute the natural state model for realizing polynomial transfer matrices in two indeterminates and, therefore, for implementing a code using digital hardware.

Seeking to make a contribution to the evolutionary trend described above, this paper outlines an algebraic theory of 2D convolutional codes, which encompasses both a behavioural approach to the internal structure of the codes and a state space procedure for synthesizing 2D encoders and decoders.

In the first part of the paper, 2D convolutional codes are introduced as modules of doubly indexed sequences. Several connections with the submodule of finite codewords are discussed, thus providing different characterizations of the convolutional property and a complete classification of all equivalent encoders.

Next part deals with 2D basic codes and injective encoders. Unlike the 1D case, a 2D convolutional code needs not admit necessarily an injective encoder. So "good" codes constitute only a proper subclass of the convolutional ones, and characterizing such class requires to introduce the notions of extendability and left zero-prime encoders.

Most of the concepts introduced in the previous parts are revisited in the section devoted to the duality notion. The different point of view therein adopted finds a very natural application in the synthesis of 2D syndrome decoders.

In the last section we concentrate on some aspects of the realization problem, considering finite memory 2D systems as candidates for its solution. The quarter plane causality which underlies the state updating of these models requires to introduce some restrictions on the supports of the information signals to be encoded, and to cope with standard polynomials, instead of Laurent polynomials, in representing encoders and decoders. Finally, in order to reduce the computational effort involved in designing the transmission chain, we investigate the possibility of realizing 2D decoders as inverse state models of the corresponding encoders.

Due to the intrinsic complexity of the subject, some results have still a preliminary character and some topics remain rather inexplored. Nevertheless, it's hoped that the main features of the theory have been covered, and some directions for future developments are broadly visible from our exposition. In particular, looking for the future, it would be clearly desirable to relate the properties of a 2D polynomial matrix with the dimension of its minimal state space realizations. This could eventually lead to express the requirement of obtaining an optimal encoder for a given code $\mathcal{C}$ as a constraint on the polynomial structure of the encoder itself. At the present time, however, little is known concerning the structure of 2D minimal realizations and, as a consequence, there's no possibility to single out, among the equivalent encoders of a given 2D code, those which exhibit the most economic realizations.

2 2D convolutional codes and their encoders

A 2D code $\mathcal{C}$ of lenght n over a finite field $\mathbf{F}$ can be viewed as a set of sequences indexed on the discrete plane $\mathbf{Z} \times \mathbf{Z}$ and taking values in $\mathbf{F}^n$. Thus, denoting the sequence space $(\mathbf{F}^n)^{\mathbf{Z}\times\mathbf{Z}}$ as $\mathcal{F}_\infty^n$, it follows that $\mathcal{C}$ is a subset of $\mathcal{F}_\infty^n$.

In 1D coding theory, the natural order of $\mathbf{Z}$ is usually associated with the time ordering and, therefore, with the sequential structure of the data flow. This motivates the habit of considering 1D codewords with left compact support, and to represent them [12] as vectors with components in the field $\mathbf{F}((z))$ of formal power series with left compact support.

When encoding two-dimensional data, there is no natural notion of causality inducing a particular ordering in $\mathbf{Z} \times \mathbf{Z}$ and, consequently, some a priori restrictions on the supports of the sequences in $\mathcal{C}$. So, adopting this point of view, we will, in general, assume that the supports of the elements of the code could extend indefinitely in all the directions of the discrete plane. Special attention, however, will be deserved to the class of codes whose elements have finite supports and to the possibility of characterizing complete codes as the duals of the above class.
In the sequel, it will be convenient to represent the signals of $\mathcal{F}_\infty^n$ and, hence, the codewords of $\mathcal{C}$, via formal power series, by associating any sequence $\{w(i,j)\}$ with the series

$$\sum_{i,j \in \mathbf{Z}} w(i,j) z_1^i z_2^j. \tag{2.1}$$

To avoid cumbersome notations, we will adopt the symbol $\mathbf{w}$ for denoting both the sequence and the associated power series (2.1). The context will always make clear which object we are referring to. Sometimes, mostly when a power series $\mathbf{v}$ is obtained as a Cauchy product, it will be useful to denote the coefficient of $z_1^i z_2^j$ in $\mathbf{v}$ as $(\mathbf{v}, z_1^i z_2^j)$. The main advantage in using formal power series is that many linear operators on $\mathcal{F}_\infty^n$ can be represented by appropriate matrices, with elements in $\mathcal{F}_\pm := \mathbf{F}[z_1, z_2, z_1^{-1}, z_2^{-1}]$, the ring of 2D Laurent polynomials (*L-polynomials*). This way, several fundamental operator properties find an immediate counterpart in terms of the structure of the corresponding matrices and, in particular, of their factors.

Definition *A matrix $G(z_1, z_2) \in \mathcal{F}_\pm^{k \times n}$ is*

- *$\mathcal{F}_\pm-$ unimodular, if $k = n$ and $\det G$ is a unit in $\mathcal{F}_\pm$;*
- *left factor-prime (ℓFP), $if\, for\, every\, factorization\, G = T\, \bar{G}$, with $T \in \mathcal{F}_\pm^{k \times k}$, T is $\mathcal{F}_\pm$ − unimodular;*
- *left zero-prime (ℓZP), $if\, the\, ideal\, generated\, by\, the\, maximal\, order\, minors\, of\, G\, is\, the\, ring\, F_\pm$ itself.*

Introducing a convolutional structure on $\mathcal{C}$ requires to endow the set of its sequences with some closure properties, which constitute the mathematical formalization of very natural constraints of regularity. The most common requirements on $\mathcal{C}$ are linearity and shift invariance:

(a) [Linearity] *If $\mathbf{w}_1$ and $\mathbf{w}_2$ belong to $\mathcal{C}$, then $\alpha\mathbf{w}_1 + \beta\mathbf{w}_2$ belongs to $\mathcal{C}$ for every α and β in $\mathbf{F}$.*

(b) [Shift Invariance] *$\mathbf{w} \in \mathcal{C}$ implies that $\mathbf{v} = z_1^h z_2^k \mathbf{w} \in \mathcal{C}$, $\forall\ h, k \in \mathbf{Z}$, i.e. $\mathcal{C}$ is invariant w.r.t. the shifts in $\mathbf{Z} \times \mathbf{Z}$ along the coordinate axes.*

As the set of formal power series $\mathcal{F}_\infty^n$ is naturally endowed with a module structure w.r.t. $\mathcal{F}_\pm$, codes which satisfy properties (a) and (b) can be characterized as $\mathcal{F}_\pm-$ submodules of $\mathcal{F}_\infty^n$. They will be called *admissible codes*.

Example 1 Every submodule $\mathcal{C}$ of $\mathcal{F}_\pm^n$ is an admissible code. Since $\mathcal{F}_\pm^n$ is an $\mathcal{F}_\pm-$ Noetherian module [13], $\mathcal{C}$ is finitely generated, i.e. there exists a finite set of row vectors $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_h$ in $\mathcal{F}_\pm^n$ such that

$$\mathcal{C} = \{\sum_{i=1}^{h} a_i \mathbf{g}_i,\ a_i \in \mathcal{F}_\pm\} = \{\mathbf{a}G,\ \mathbf{a} \in \mathcal{F}_\pm^h\} =: \mathrm{Im}_\pm G,$$

where G denotes the polynomial matrix $G = \mathrm{col}\{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_h\}$.

Example 2 A sequence $\mathbf{w}$ has past compact support if, for every $(l, m) \in \mathbf{Z} \times \mathbf{Z}$, the corresponding past cone $\{(i, j) : i \leq l, j \leq m\}$ intersects the support of $\mathbf{w}$ in a finite number of points.
The set

$$\mathcal{C} = \{\mathbf{w} \in \mathcal{F}_\infty^n : \mathrm{supp}(\mathbf{w}) \text{ past compact}\}$$

is an admissible code.

Example 3 Let $M_1,\ M_2 \in \mathbf{F}^{\nu\times\nu}$ consitute a pair of commuting invertible matrices and let K be in $\mathbf{F}^{\nu\times n}$. The set

$$\mathcal{C} = \{\mathbf{w} \in \mathcal{F}_\infty^n : \mathbf{w} = \sum_{i,j\in\mathbf{Z}} x M_1^i M_2^j K z_1^i z_2^j,\ x \in \mathbf{F}^\nu\}$$

is an admissible code. Moreover its dimension, as $\mathbf{F}-$ vector space, is finite. It can be shown that all finite dimensional admissible codes have the above structure [14].

When testing whether a sequence $\mathbf{w}$ belongs to a code $\mathcal{C}$ which includes codewords with infinite support, the possibility of resorting to a finite set of autoregressive equations, applied at every point of $\mathbf{Z}\times\mathbf{Z}$, constitutes a very favourable situation. Actually, in this case, we can recognize a codeword by using only a finite set of samples $w(i, j)$ at each step of the testing procedure. Such a possibility clearly corresponds [7] to the existence of an L-polynomial matrix $H^T(z_1, z_2)$, such that

$$\mathcal{C} = \ker H^T(z_1, z_2) = \{\mathbf{w} \in \mathcal{F}_\infty^n : \mathbf{w}H^T = \mathbf{0}\}, \tag{2.2}$$

and it can be restated as a closure property of the code, as follows:

(c) [Completeness] *Let $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \mathcal{S}_3 \ldots$ be a sequence of finite windows invading $\mathbf{Z} \times \mathbf{Z}$, so that every point $(i, j) \in \mathbf{Z} \times \mathbf{Z}$ eventually belongs to all windows of the sequence, and let $\mathbf{w} \in \mathcal{F}_\infty^n$. If for every nonnegative integer m there exists $\mathbf{v}_m \in \mathcal{C}$ such that*

$$\mathbf{v}_m|\mathcal{S}_m = \mathbf{w}|\mathcal{S}_m, \tag{2.3}$$

then $\mathbf{w} \in \mathcal{C}$.

An equivalent way of stating the above property is the following. Introduce in $\mathbf{Z} \times \mathbf{Z}$ a distance function $d(\cdot, \cdot)$, by assuming

$$d\Big((i,j),(h,k)\Big) = |i-h| + |j-k|, \quad \forall\ (i,j),(h,k) \in \mathbf{Z} \times \mathbf{Z},$$

and define distance $\Delta(\cdot, \cdot)$ between two sequences $\mathbf{v}$ and $\mathbf{w}$ in $\mathcal{F}_\infty^n$ as follows

$$\Delta(\mathbf{v}, \mathbf{w}) = \begin{cases} 0, & \text{if } \mathbf{v} = \mathbf{w}; \\ 2^{-\min\{d((i,j),(0,0)):\ v(i,j) \neq w(i,j)\}}, & \text{otherwise.} \end{cases} \tag{2.4}$$

Then $\mathcal{F}_\infty^n$ becomes a metric space, and property (c) is exactly the completeness of $\mathcal{C}$ in the topology induced by the distance function Δ. This means that, if a sequence $\mathbf{v}_1, \mathbf{v}_2, \ldots$ in $\mathcal{C}$ converges to $\mathbf{w}$, then $\mathbf{w} \in \mathcal{C}$.

Proposition 2.1 *Let $\mathcal{C} \subseteq \mathcal{F}_\infty^n$ be an admissible code. Then $\mathcal{C}$ is a* complete code*, i.e. it satisfies condition (c), if and only if (2.2) holds.*

PROOF The proposition above has been proved by Paula Rocha in [7]. For an alternative proof see [15] ■

As an immediate corollary, we have that properties (a) ÷ (c) are equivalent to the possibility of representing $\mathcal{C}$ as the kernel of an L-polynomial matrix.

Remark I The codes considered in Examples 1 and 2 are not complete. On the other hand the code of Example 3 is complete, in fact it can be proved that it is the kernel of a polynomial matrix [15].

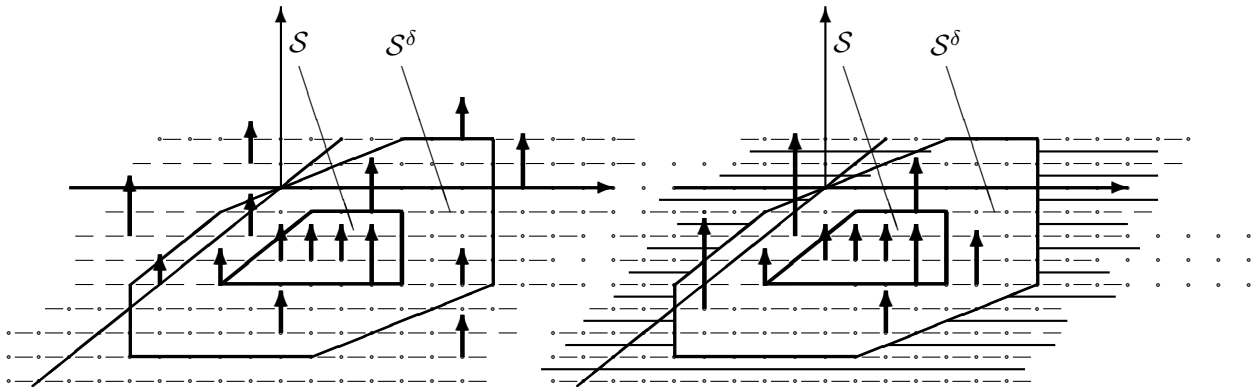
Given a finite window $\mathcal{S}$ and a set of samples, obtained by restricting to $\mathcal{S}$ a (possibly infinite) codeword $\mathbf{w}$, it's interesting to investigate whether the data set, $\mathbf{w} \mid \mathcal{S}$, can be completed into an appropriate finite codeword $\mathbf{v}$, whose support does not "exceed too much" $\mathcal{S}$. If so, the values a codeword $\mathbf{w}$ assumes on the window s, constrain only the samples $w(i,j)$ in a finite neighbourhood of it or, equivalently, do not provide any information on $\mathbf{w}$ at points which are far enough from $\mathcal{S}$. Therefore, if no additional information on $\mathbf{w}$ is available, we can always assume that the partial data at our disposal come from a finite codeword. The above property can be stated as follows:

(d) [Controllability] *There exists a positive integer δ such that, for every finite set $\mathcal{S} \subset \mathbf{Z} \times \mathbf{Z}$ and every $\mathbf{v}_1 \in \mathcal{C}$, there is a codeword $\mathbf{v}_2 \in \mathcal{C}$, such that*

$$\mathbf{v}_1 \mid \mathcal{S} = \mathbf{v}_2 \mid \mathcal{S}$$

and

$$\operatorname{supp}(\mathbf{v}_2) \subseteq \mathcal{S}^\delta := \{(i,j) \in \mathbf{Z} \times \mathbf{Z} : d((i,j), \mathcal{S}) < \delta\}.$$



codeword $\mathbf{v}_1$ codeword $\mathbf{v}_2$

FIG 1

A fundamental objective of coding theory is the investigation of the intrinsic structure of the codes, without taking into account the way codewords are generated, and the analysis of those features, good codes must be endowed with. These should make them as efficient as possible with respect to design requirements, such as the distance among the codewords and the noise sensitivity.
Under the hypothesis that a 2D code $\mathcal{C}$ is complete, there are several equivalent formulations of the controllability property, which concern the internal structure of the codewords set. Some of them refer to the *submodule of the finite codewords*,

$$\mathcal{C}_f := \{\mathbf{w} \in \mathcal{C} : \mathrm{supp}(\mathbf{w})\mathrm{finite}\} = \mathcal{C} \cap \mathcal{F}_{\pm}^n,$$

others to the possibility of obtaining the code by a sort of "patching" of appropriate finite codewords.

($\mathbf{d}_1$) There exists a positive integer ρ such that, given two disjoints subsets of $\mathbf{Z}\times\mathbf{Z}$, $\mathcal{S}_1$ and $\mathcal{S}_2$ whose distance d is greater than ρ, and two codewords $\mathbf{w}_1$ and $\mathbf{w}_2$ in $\mathcal{C}$, there is a codeword $\mathbf{v} \in \mathcal{C}$ such that

$$\mathbf{w}_1 \mid \mathcal{S}_1 = \mathbf{v} \mid \mathcal{S}_2 \qquad \text{and} \qquad \mathbf{w}_2 \mid \mathcal{S}_2 = \mathbf{v} \mid \mathcal{S}_2.$$

This result can be rephrased as the possibility of "concatenating" two portions of distinct codewords into a single codeword, provided their supports are far enough [6÷8].

($\mathbf{d}_2$) There exists a finite set $\mathcal{I}$ of finite support codewords with the property that $\mathbf{w} \in \mathcal{F}_\infty^n$ belongs to $\mathcal{C}$ if and only $\mathbf{w}$ is represented as a locally finite sum of some, possibly shifted, elements of $\mathcal{I}$. So every codeword of $\mathcal{C}$ is obtained by resorting to an appropriate "covering" of the discrete plane with codewords and shifted codewords of $\mathcal{I}$.

($\mathbf{d}_3$) The code $\mathcal{C}$ can be completely reconstructed from $\mathcal{C}_f$, the $\mathcal{F}_\pm-$module of its finite codewords, by means of a limit operation, namely, $\mathbf{w}$ belongs to $\mathcal{C}$ if and only

if there is a sequence $\mathbf{w}_1$, $\mathbf{w}_2$,.. in $\mathcal{C}_f$, converging to $\mathbf{w}$ in the sense of the pointwise topology. So $\mathcal{C}$ can be viewed as the closure of the module $\mathcal{C}_f$, that is as the smallest (complete) code containing $\mathcal{C}_f$.

On the other hand, a code is naturally understood as the result of an encoding process applied to the information signals. Therefore many concepts in coding theory are connected with the existence of an input-output transformation, whose image is the code itself. In this perspective, if the information signals are sequences in $\mathbf{Z} \times \mathbf{Z}$, with values in $\mathbf{F}^k$ for some integer k, and the code $\mathcal{C}$ is linear and shift invariant, it is natural to associate the transformation with a $k \times n$ L-polynomial matrix $G(z_1, z_2)$ and represent the code as

$$\mathcal{C} = \mathrm{Im}G := \{\mathbf{w} : \mathbf{w} = \mathbf{u}G, \mathbf{u} \in \mathcal{F}_\infty^k\}. \tag{2.5}$$

As we shall see, property (2.5) is equivalent to the "internal" properties (a) $\div$ (d). Consequently, the convolutional nature of $\mathcal{C}$, i.e. the possibility of generating all codewords of $\mathcal{C}$ by convolving the input sequences with the matrix G of the impulse response, has an exact counterpart in terms of the internal structure of the code, which can be characterized without any reference to the encoding process. We call *convolutional* a complete code satisfying condition (d), or equivalently a code described as in (2.5).

Proposition 2.2, below, formalizes the main statements concerning the controllability property.

Proposition 2.2 [Equivalent characterizations of Convolutional Codes] *Let $\mathcal{C} \subseteq \mathcal{F}_\infty^n$. The following are equivalent:*
(1) $\mathcal{C} = \mathrm{Im}G$, $G \in \mathcal{F}_\pm^{k\times n}$;
(2) $\mathcal{C} = \ker H^T$, $H^T \in \mathcal{F}_\pm^{n\times p}$, H^T right factor prime;
(3d) $\mathcal{C}$ is complete and satisfies property (d);
(3d_i) $\mathcal{C}$ is complete and satisfies property (d_i), $i = 1, 2, 3$.

Proof The equivalences (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3d_1) have been proved by P.Rocha in [7]. An independent proof, based on the notion of duality, will be provided in section 4. For the remaining equivalences we proceed by showing that (3d_1) $\Leftrightarrow$ (3d) and (3d_3) $\Leftrightarrow$ (1) $\Leftrightarrow$ (3d_2).

(**3d**$_1$) $\Rightarrow$ (**3d**) Take $\delta := \rho$ and apply (3d_1) to $\mathbf{w}_1 := \mathbf{v}_1$, $\mathbf{w}_2 = \mathbf{0}$, $\mathcal{S}_1 := \mathcal{S}$ (finite) and $\mathcal{S}_2 := C\mathcal{S}^\delta$, the complementary set of $\mathcal{S}^\delta$. The finite codeword $\mathbf{v}$ obtained in (3d_1) is the codeword $\mathbf{v}_2$ we are looking for.

(**3d**) $\Rightarrow$ (**3d**$_1$) The space $\mathcal{F}_\infty^n$ is sequentially compact, i.e. every sequence of elements in $\mathcal{F}_\infty^n$ contains a subsequence which converges to an element in $\mathcal{F}_\infty^n$. As a consequence of this property, one can show [15] that (3d) extends to the infinite subsets of $\mathbf{Z} \times \mathbf{Z}$. Set $\rho := \delta$. There exist two codewords $\mathbf{v}_1$ and $\mathbf{v}_2$ such that $\mathbf{v}_i|\mathcal{S}_i = \mathbf{w}_i|\mathcal{S}_i$, $i = 1, 2$, and $\mathrm{supp}(\mathbf{v}_i) \subseteq \mathcal{S}_i^\delta$. The signal $\mathbf{v} := \mathbf{v}_1 + \mathbf{v}_2$ belongs to $\mathcal{C}$ and satisfies $\mathbf{v}|\mathcal{S}_i = \mathbf{v}_i \mid \mathcal{S}_i = \mathbf{w}_i|\mathcal{S}_i$, $i = 1, 2$, as required.

(**1**) $\Rightarrow$ (**3d**$_2$) Consider the set $\mathcal{I} := \{[\,\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_k\,]G,\ \alpha_\nu \in \mathbf{F}\}$, whose elements are the codewords corresponding to the "atomic" input signals $[\,\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_k\,] \in \mathbf{F}^k$.

As $\mathbf{F}$ is a finite field, $\mathcal{I}$ is finite too. Every codeword in $\mathcal{C}$ can be written as

$$\mathbf{w} = \mathbf{u}G = \sum_{i,j\in\mathbf{Z}} z_1^i z_2^j \Big([\, u_1(i,j) u_2(i,j) \quad \dots \quad u_k(i,j)\,] G \Big). \tag{2.6}$$

Since all codewords $\mathbf{w}_{ij}(z_1,z_2) := [\, u_1(i,j) \quad u_2(i,j) \quad \dots \quad u_k(i,j)\,]G$ are elements of $\mathcal{I}$, (2.6) represents $\mathbf{w}$ as a locally finite sum of elements and shifted elements of $\mathcal{I}$.

$(\mathbf{3d}_2) \Rightarrow (\mathbf{1})$ Let $\mathcal{I} := \{\mathbf{c}_1(z_1,z_2), \mathbf{c}_2(z_1,z_2), \dots, \mathbf{c}_p(z_1,z_2)\}$, with $\mathbf{c}_i(z_1,z_2) \in \mathcal{F}_\pm^n$, $i = 1,2,...,p$. By assumption, the codewords in $\mathcal{C}$ are the elements in $\mathcal{F}_\infty^n$ which can be expressed as

$$\begin{aligned} \sum_{i,j\in\mathbf{Z}} & \quad \sum_{t=1}^{p} \delta_t(i,j) z_1^i z_2^j \; \mathbf{c}_t(z_1,z_2) \\ = & \sum_{i,j\in\mathbf{Z}} [\, \delta_1(i,j) z_1^i z_2^j \delta_2(i,j) z_1^i z_2^j \dots \; \delta_p(i,j) z_1^i z_2^j \,] \begin{bmatrix} \mathbf{c}_1(z_1,z_1) \\ \mathbf{c}_2(z_1,z_1) \\ \dots \\ \mathbf{c}_p(z_1,z_1) \end{bmatrix} \end{aligned} \tag{2.7}$$

where $\delta_t(\cdot,\cdot)$ takes values in $\{0,1\}$.
Letting $G(z_1,z_2) = \mathrm{col}\{\mathbf{c}_1(z_1,z_1), \mathbf{c}_2(z_1,z_1), \dots, \mathbf{c}_p(z_1,z_1)\}$, we have that the $\mathcal{F}_\pm-$module generated by the rows of G is included in $\mathcal{C}$. We aim to prove that $\mathbf{u}G$ is in $\mathcal{C}$ for every $\mathbf{u} \in \mathcal{F}_\pm^p$. Actually, given any sequence of finite sets $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \dots$ invading $\mathbf{Z}\times\mathbf{Z}$, the sequence of input signals $\mathbf{u}_\nu$ defined by

$$u_\nu(i,j) = \begin{cases} u(i,j) & \text{if } (i,j) \in \mathcal{S}_\nu; \\ 0 & \text{otherwise,} \end{cases}$$

converges to $\mathbf{u}$. By the continuity of G [15], $\mathbf{u}_\nu G$ converges to $\mathbf{u}G$. Since $\mathcal{C}$ is complete, $\mathbf{u}_\nu G \in \mathcal{C},\ \forall\nu \Rightarrow\ \mathbf{u}G \in \mathcal{C}$. Consequently, $\mathrm{Im}G \subseteq \mathcal{C}$.
On the other hand, by (2.7) every codeword can be expressed as the $G-$image of a series in $\mathcal{F}_\infty^p$ with coefficients in $\{0,1\}$. Therefore $\mathcal{C} = \mathrm{Im}G$.

$(\mathbf{1}) \Rightarrow (\mathbf{3d}_3)$ Let $\mathbf{w}$ be in $\mathcal{C} = \mathrm{Im}G$ and $\mathbf{w} = \mathbf{u}G$. Consider an L-polynomial sequence $\{\mathbf{u}_\nu\}$ converging to $\mathbf{u}$. Because of the continuity of G, the sequence of L-polynomial codewords $\{\mathbf{w}_\nu\} := \{\mathbf{u}_\nu G\}$ converges to $\mathbf{w} = \mathbf{u}G$.

$(\mathbf{3d}_3) \Rightarrow (\mathbf{1})$ The finite codewords of $\mathcal{C}$ constitute an $\mathcal{F}_\pm-$module $\mathcal{C}_f$, which is finitely generated as a submodule of $\mathcal{F}_\pm^n$. Let $\mathbf{g}_\nu \in \mathcal{F}_\pm^n$, $\nu = 1,2,...k$, constitute a set of generators for $\mathcal{C}_f$ and $G := \mathrm{col}\{\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_k\}$. Clearly $\mathrm{Im}G \subseteq \mathcal{C}$.
To prove the reverse inclusion, consider any codeword $\mathbf{w}$ in $\mathcal{C}$ and a sequence of finite codewords, $\mathbf{w}_\nu \in \mathcal{C}_f$, $\nu = 1,2,...$, converging to $\mathbf{w}$. Since all finite codewords in $\mathcal{C}$ are linear combinations in $\mathcal{F}_\pm$ of the rows of G, there is a sequence $\{\mathbf{u}_\nu\}$, $\mathbf{u}_\nu \in \mathcal{F}_\pm^k$, such that $\{\mathbf{u}_\nu G\} = \{\mathbf{w}_\nu\}$. By the sequential compactness of $\mathcal{F}_\infty^k$ [15], we can extract from $\{\mathbf{u}_\nu\}$ a subsequence $\{\mathbf{u}_{\nu_j}\}$ that converges to some $\mathbf{u}$ in $\mathcal{F}_\infty^k$. So, by the continuity of the operator G, we have

$$\mathbf{w} = \lim_{j\to\infty} (\mathbf{u}_{\nu_j} G) = (\lim_{j\to\infty} \mathbf{u}_{\nu_j}) G = \mathbf{u}G \qquad \blacksquare$$

It can be easily realized that, while a $k \times n$ polynomial matrix $G(z_1, z_2)$ uniquely identifies a convolutional code $\mathrm{Im}\, G = \{\mathbf{w} = \mathbf{u}G : \mathbf{u} \in \mathcal{F}_\infty^k\}$, the converse does not hold, as the same code $\mathcal{C}$ can be described as the image of different L-polynomial matrices. Two matrices $G_1(z_1, z_2)$ and $G_2(z_1, z_2)$, with elements in $\mathcal{F}_\pm$ and the same number of columns, are *equivalent encoders* if $\mathrm{Im}\ G_1 = \mathrm{Im}\ G_2$. Since each convolutional code biuniquely corresponds to a class of equivalent encoders, the natural problems arise to investigate what conditions guarantee that two matrices belong to the same class and to find out in every equivalence class the most efficient encoders.

To answer these questions we need some preliminary results, concerning the relationships between the $\mathcal{F}_\pm-$ module

$$\mathrm{Im}_\pm G := \{\mathbf{u}G : \mathbf{u} \in \mathcal{F}_\pm^k\}$$

and the $\mathcal{F}_\pm-$ submodule $\mathcal{C}_f$ of the finite codewords of $\mathcal{C} = \mathrm{Im}\ G$, that will be also denoted as $(\mathrm{Im}\ G)_f$.

Lemma 2.3 [15] *Let $\bar{G}(z_1, z_2)$ be a $k \times n$ $\ell FPL-$ polynomialmatrix.Then*
i) the $\mathcal{F}_\pm-$ module $Im_\pm\bar{G}$ is free;
ii) if $T(z_1, z_2)$ is a $k' \times k$ L-polynomial matrix, of rank k over $\mathbf{F}(z_1, z_2)$, then $(\mathrm{Im}\ T\bar{G})_f$ coincides with $(\mathrm{Im}\ \bar{G})_f$. ■

Lemma 2.4 *Let $G(z_1, z_2)$ be a $k \times n$ L-polynomial matrix, with full row rank over $\mathbf{F}(z_1, z_2)$. The following properties are equivalent*
i) $G(z_1, z_2)$ is ℓFP;
ii) the module $(\mathrm{Im}\ G)_f$ of the finite codewords in $\mathrm{Im}(G)$ coincides with $\mathrm{Im}_\pm G$, i.e. every finite codeword $\mathbf{w}$ of $\mathrm{Im}G$ is the image of a finite input sequence;
iii) $\ker G$, the $\mathbf{F}-$vector space of all information sequences in $\mathcal{F}_\infty^k$ which produce the zero codeword, is finite dimensional.

PROOF **i) ⇒ ii)** Clearly $\mathrm{Im}_\pm G$ is included in $(\mathrm{Im}G)_f$, since every linear combination in $\mathcal{F}_\pm$ of the rows of G is a finite codeword of $\mathcal{C}$.
We aim to prove the inverse inclusion. The $\ell FP condition implies [5] the existence of two matrices \mathrm{X}(\mathrm{z}_1, z_2)$ and $Y(z_1, z_2)$, with elements in $\mathcal{F}_\pm$, and two polynomials $h(z_1) \in \mathbf{F}[z_1, z_1^{-1}]$, $k(z_2) \in \mathbf{F}[z_2, z_2^{-1}]$ such that (A2) holds. We therefore have

$$\mathbf{w}X = \mathbf{u}GX = \mathbf{u}h(z_1), \quad \mathbf{w}Y = \mathbf{u}GY = \mathbf{u}k(z_2)$$

and, by multiplying the first equation by $k(z_2)$ and the second one by $h(z_1)$, we get $\mathbf{w}Xk(z_2) = \mathbf{w}Yh(z_1)$. Since $h(z_1)$ and $k(z_2)$ are coprime, $h(z_1)$ is a common factor of all the components of $\mathbf{w}X$, that is there exists an L-polynomial vector $\mathbf{p}(z_1, z_2) \in \mathcal{F}_\pm^k$ such that $\mathbf{w}X = \mathbf{p}h(z_1)$. Thus $h(z_1)\mathbf{w} = h(z_1)(\mathbf{u}G) = h(z_1)(\mathbf{p}G)$ and, consequently,

$$h(z_1)(\mathbf{w} - \mathbf{p}G) = 0 \qquad (2.8).$$

Since all entries in (2.8) are in $\mathcal{F}_\pm$, it follows that $\mathbf{w} = \mathbf{p}G$, and therefore $(\mathrm{Im}G)_f$ is included in $\mathrm{Im}_\pm G$.

ii) $\Rightarrow$ i) Assume that G is not $\ell FP. As a consequence of Corollary A.2 in the Appendix, G can be rewritten as G = T $\bar{G}$ where $\bar{G} \in \mathcal{F}_{\pm}^{k\times n}$ is ℓFP and T $\in \mathcal{F}_{\pm}^{k\times k}$ is a full rank nonunimodular matrix. Thus there exists a vector $\mathbf{p}(z_1, z_2) \in \mathcal{F}_{\pm}^{k}$ such that equation

$$\mathbf{u}T = \mathbf{p} \tag{2.9}$$

has no solution in $\mathcal{F}_{\pm}^{k}$. However, as T is a full rank square matrix, (2.9) admits a unique solution in $\mathbf{F}(z_1, z_2)$ given by

$$\mathbf{u} = \mathbf{p}T^{-1} = \mathbf{p}\frac{\mathrm{adj}T}{\det T}.$$

The entries of $\mathbf{u}$ can be viewed as series in $\mathcal{F}_{\infty}$, and hence $\mathbf{u}$ is an infinite input sequence in $\mathcal{F}_{\infty}^{k}$. We aim to show that $\mathbf{w} = \mathbf{u}G$ is a finite codeword that does not belong to $\mathrm{Im}_{\pm}G$. Actually, $\mathbf{w} = \mathbf{u}G = (\mathbf{u}T)\bar{G} = \mathbf{p}\bar{G}$ is finite.
On the other hand, assume that there is a finite input sequence $\mathbf{v}$ such that $\mathbf{w} = \mathbf{v}G$. Then we have $\mathbf{w} = \mathbf{v}G = (\mathbf{v}T)\bar{G}$, which implies $(\mathbf{p} - \mathbf{v}T)\bar{G} = 0$. Since $\mathrm{Im}_{\pm}\bar{G}$ is a free module, we have that $\mathbf{v}T = \mathbf{p}$, and equation (2.9) has an L-polynomial solution, a contradiction.

i) $\Leftrightarrow$ iii) $\ker G := \{\mathbf{u} \in \mathcal{F}_{\infty}^{k} : \mathbf{u}G = 0, \}$ can be viewed as an autoregressive description of a complete behaviour. It has been proved [7,14] that a necessary and sufficient condition for a behaviour being finite dimensional is that G is ℓFP ■

We are now in a position for introducing the basic results about the equivalence of two encoders.

Proposition 2.5 [Equivalent Encoders] *Let $G_1(z_1, z_2)$ and $G_2(z_1, z_2)$ be two matrices with elements in $\mathcal{F}_{\pm}$ and dimensions $k_1 \times n$ and $k_2 \times n$, respectively.*
G_1 and G_2 are equivalent encoders if and only if
i) under the assumption that both $G_1(z_1, z_2)$ and $G_2(z_1, z_2)$ are ℓFP, we have $k_1 = k_2$ and $G_2(z_1, z_2) = U(z_1, z_2)G_1(z_1, z_2)$, with $U(z_1, z_2)$ unimodular; ii) under the assumption that $G_1(z_1, z_2)$ is ℓFP, there is a $k_2 \times k_1$ full column rank L-polynomial matrix, $P_1(z_1, z_2)$, such that $G_2 = P_1G_1$;

iii) in the general case, there exist two full column rank L-polynomial matrices $P_1(z_1, z_2)$ and $P_2(z_1, z_2)$, of suitable dimensions, such that

$$P_1G_1 = P_2G_2. \tag{2.10}$$

Proof Given a full column rank matrix $P \in \mathcal{F}_{\pm}^{h\times k}$, the map $P : \mathcal{F}_{\infty}^{h} \to \mathcal{F}_{\infty}^{k} : \mathbf{u} \mapsto \mathbf{u}P$ is onto. Consequently the convolutional codes ImG and ImPG coincide for any $G \in \mathcal{F}_{\pm}^{k\times n}$. Thus (2.10) in iii), and in particular $G_2 = UG_1$ and $G_2 = P_1G_1$ in i) and ii), imply that G_1 and G_2 are equivalent encoders.

Viceversa, assume first that G_1 and G_2 are ℓFP equivalent encoders. Then, by property (d_3) and Lemma 2.4 we have

$$\mathrm{Im}_{\pm}G_1 = (\mathrm{Im}G_1)_f = (\mathrm{Im}G_2)_f = \mathrm{Im}_{\pm}G_2 \tag{2.11}$$

As each row of G_1 (of G_2) is an $\mathcal{F}_{\pm}$- linear combination of the rows of G_2 (of G_1), there exist L-polynomial matrices P_1 and P_2 such that $P_1G_1 = G_2$ and $P_2G_2 = G_1$. We will have then $G_1 = P_1P_2G_1$ and $G_2 = P_2P_1G_2$. The $\ell FP property yields P_1P_2 = I_{k_1}$, $P_2P_1 = I_{k_2}$, showing that $\mathrm{k}_1 = k_2$ and both P_1 and P_2 are unimodular.

Assume next that G_1 and G_2 are equivalent encoders, and only G_1 is ℓFP. By Corollary A.2, G_2 can be factorized as $G_2 = T\bar{G}_2$, where $\bar{G}_2$ is ℓFP and T full column rank. Thus $\mathrm{Im}\mathrm{G}_2 = \mathrm{Im}\bar{G}_2$ and, consequently, G_1 and $\bar{G}_2$ are ℓFP equivalent encoders. It follows that $\bar{G}_2 = UG_1$, for a suitable unimodular matrix U, and, letting $P_1 = TU$, one gets $G_2 = P_1G_1$, as required.

Finally, (case iii)) suppose that G_1 and G_2 are equivalent encoders, and neither G_1 nor G_2 are ℓFP. Clearly we have $\mathrm{G}_i = \tilde{T}_i\bar{G}_i$, $i = 1, 2$, with $\tilde{T}_i$ full column rank and $\bar{G}_i$ ℓFP matrices. Moreover $\bar{G}_1 = U\bar{G}_2$ for some $\mathcal{F}_{\pm}$-unimodular matrix U. So, letting $T_1 = \tilde{T}_1U$ and $T_2 = \tilde{T}_2$ we get

$$G_1 = T_1\bar{G}_2, \quad G_2 = T_2\bar{G}_2$$

Consider any pair of L-polynomial matrices X_1 and X_2 with the property that $A_i := [\, T_i \quad X_i \,]$, $i = 1, 2$, is a nonsingular $k_i \times k_i$ L-polynomial matrix, and assume $k_1 \geq k_2$. Then we have

$$A_1^{-1}G_1 = \begin{bmatrix} \bar{G}_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} A_2^{-1} \\ 0 \end{bmatrix}$$

Introduce the following left matrix fraction description (MFD)

$$\begin{bmatrix} A_2^{-1} \\ 0 \end{bmatrix} = L^{-1}N,$$

with N full column rank, and rewrite A_1L^{-1} as a left MFD, $A_1L^{-1} = Q^{-1}B$. As $G_1 = A_1L^{-1}NG_2 = Q^{-1}BNG_2$, we end up with $QG_1 = BNG_2$, which corresponds to (2.10), upon assuming $P_1 = Q$ and $P_2 = BN$ ■

Remark II It's worthwhile to underline that every convolutional code $\mathcal{C}$ can be represented as the image of a ℓFP matrix. Actually, given any encoder $\mathrm{G}(\mathrm{z}_1, z_2)$ of $\mathcal{C}$, by Corollary A.2 we can extract a greatest left factor, obtaining

$$G(z_1, z_2) = T(z_1, z_2)\bar{G}(z_1, z_2),$$

with T full column rank and $\bar{G}$ ℓFP. By the above proposition, $\bar{G}(z_1, z_2)$ is a ℓFP encoder of C.

3 Injectivity and decoding

The purpose of an encoding scheme is to associate every input sequence with a specific codeword, which preserves the information message, but is less sensitive to the action of noise. So, in order to make possible the retrieval of the original message at the decoding stage, it's quite obvious that every codeword has to be the image of a unique information sequence, which amounts to assume that the map from the input space $\mathcal{F}_{\infty}^{k}$ to the codewords space $\mathcal{C}$ is injective.

As proved in the previous section, a convolutional code can be expressed as the image or the kernel of appropriate Laurent polynomial matrices. The following proposition shows that the injectivity requirement reduces to a zero primeness condition on the above matrices. This involves some relevant consequences on both the internal properties of the code $\mathcal{C}$ and the classes of encoders and decoders of $\mathcal{C}$.

Proposition 3.1 [Injective Encoders] *Let $\mathcal{C}$ be a convolutional code of length n and rank k. The following are equivalent:*
i) $\mathcal{C}$ *admits an injective encoder;*
ii) $\mathcal{C} = \text{Im } G(z_1, z_2), \quad G \in \mathcal{F}_{\pm}^{k\times n}$ ℓZP;
iii) $\mathcal{C} = \ker H^T(z_1, z_2)$, $H^T \in \mathcal{F}_{\pm}^{n\times(n-k)}$ rZP.

PROOF **i) $\Leftrightarrow$ ii)** If G is $\ell ZP, by Proposition A.3 there exists an n \times k$ matrix, $K(z_1, z_2)$, with elements in $\mathcal{F}_{\pm}$, such that $GK = I_k$. So $\mathbf{u}G = 0$ implies $0 = (\mathbf{u}G)K = \mathbf{u}(GK) = \mathbf{u}$, which means that G defines an injective input-output map.

Conversely, we aim to prove that, if G is not $\ell ZP, it is not an injective encoder. If rank$G ¡ k, the result is trivial, so we confine ourselves to the case rank$G = k.

Consider first the case when G is not $\ell FP. By Corollary A.2, there exist two L-polynomial matrices, \bar{G}(z_1, z_2)$, $k \times n$ and $\ell FP, and$T(z_1, z_2), $k \times k$ with $\det T \neq 0$ and not a unit in $\mathcal{F}_{\pm}$, such that $G = T\bar{G}$.

- If $\det T$ is not a unit in $\mathbf{F}(z_1)[z_2, z_2^{-1}]$, in the (renormalized) Hermite form [2,5] of T w.r.t $\mathbf{F}[z_1, z_1^{-1}]$ $[z_2, z_2^{-1}]$, we have $S(z_1, z_2) = L(z_1, z_2)T(z_1, z_2)$, where $S \in \mathcal{F}_{\pm}^{k\times k}$ is upper triangular, and $L \in \mathcal{F}_{\pm}^{k\times k}$ has determinant in $\mathbf{F}[z_1, z_1^{-1}]$.
As $\det S = \det T \det L$, the assumption on $\det T$ implies that at least one diagonal element in S is a nonunit polynomial in $\mathbf{F}(z_1)[z_2, z_2^{-1}]$. Let $S_{ii}(z_1, z_2)$ be the first element with this property, and consider $v_i(z_1, z_2)$, a series in $\mathcal{F}_{\infty}$ such that

$$v_i S_{ii} = 0 \quad \text{and} \quad v_i \ \det L \neq 0. \tag{3.1}$$

Then there exists a vector $\mathbf{v} \in \mathcal{F}_{\infty}^k$, with the first $i-1$ entries identically zero, such that $\mathbf{v}S = 0$. On the other hand $\mathbf{v}L$ is not zero, otherwise $0 = \mathbf{v}L\text{adj}L = \mathbf{v}\det L$ would imply $v_i \det L = 0$, which contradicts (3.1). So $\mathbf{v}L(z_1, z_2)$ is a nonzero element in $\ker T$, and hence in $\ker G$.

- If $\det T$ is a unit in $\mathbf{F}(z_1)[z_2, z_2^{-1}]$, it cannot be also a unit in $\mathbf{F}(z_2)[z_1, z_1^{-1}]$, otherwise $\det T$ would be a unit in $\mathcal{F}_{\pm}$. So we can resort to the Hermite form of T w.r.t. $\mathbf{F}[z_2, z_2^{-1}][z_1, z_1^{-1}]$ to prove that ker$G$ is nontrivial.

When G is $\ell FP (but not \ell ZP), ker$G$ is a finite dimensional vector space$[14], and$G$ is not an injective encoder.$

ii) $\Leftrightarrow$ iii) By Proposition 2.2 and Remark II, the equivalence of ii) and iii) holds for factor prime matrices. Since for every $\mathbf{u} \in \mathcal{F}_{\infty}^k$, $(\mathbf{u}G)H^T = \mathbf{u}(GH^T) = 0$, it follows that $GH^T = 0$, and therefore, by Proposition A.4, G is $\ell ZP if and only if$H^T$ is rZP ■

Given a convolutional code $\mathcal{C} = \text{Im } G$, of length n and rank k, it is natural to wonder whether it admits injective encoders. Clearly injective (i.e. ℓZP)$encoders, if any, have to be looked for among t$ is a $k \times n$ $\ell FP encoder of$C$, $anyother \ell FP encoder is given by \tilde{G} = U\bar{G}$, U unimodular.
Since the premultiplication by $\mathcal{F}_{\pm}-$unimodular matrices preserves the $\ell ZP property, the existence of a \ell ZP en$

Unlike the 1D case, left factor-primeness does not imply left zero-primeness, and examples can be given of 2D convolutional codes devoid of injective encoders.

Example 3 Let $\mathbf{F} = \mathrm{GF}(2)$. It's easy to check that the following L-polynomial matrix

$$G_1(z_1, z_2) = \begin{bmatrix} z_1^{-1} + 1 & 0 & z_1^2 \\ z_2^{-1} & z_2 + 1 & 0 \end{bmatrix}$$

is $\ell ZP, as the only input sequence producing the zero codeword is \mathbf{u} = 0$.
On the contrary

$$G_2(z_1, z_2) = \begin{bmatrix} z_1^2 + 1 & 0 & z_1 \\ z_2 + 1 & z_2^2 + z_1 & 0 \end{bmatrix}$$

is $\ell FP but not \ell ZP, since all maximal order minors have a common zero in (1,1). Therefore \mathrm{G}_2$ is not an injective encoder. It can be easily realized that $\mathbf{u} = [\ 0 \quad \sum_{i,j} z_1^i z_2^j\ ]$ is the unique nonzero input sequence in $\ker G_2$.

According to the above discussion, we can single out among 2D convolutional codes those which admit a $\ell ZP encoder. They will be called basic, in analogy with the 1D case [1], and will be characteriz$

A complete code $\mathcal{C}$, and, a fortiori, a convolutional one, is a submodule of $\mathcal{F}_\infty^n$ whose elements satisfy a finite set of autoregressive equations, the *parity checks* of the code. By associating each equation with an L-polynomial column vector $\mathbf{h}_i^T(z_1, z_2)$, we have that $\mathbf{w} \in \mathcal{C}$ if and only if $\mathbf{w}\mathbf{h}_i^T = 0$, $i = 1, 2, ..., p$. So, by juxtaposing the columns $\mathbf{h}_i^T$ into a matrix $H^T = \mathrm{col}\{\mathbf{h}_1^T, \mathbf{h}_2^T, ..., \mathbf{h}_p^T\}$, we get the usual kernel representation $\mathcal{C} = \ker H^T = \{\mathbf{w} \in \mathcal{F}_\infty^n : \mathbf{w}H^T = 0\}$.

Definition *A sequence $\mathbf{v} \in \mathcal{F}_\infty^n$ satisfies the parity checks of the code in $(r, s) \in \mathbf{Z} \times \mathbf{Z}$ if*

$$\left(\mathbf{v}H^T, z_1^\mu z_2^\nu\right) = 0, \qquad \forall (\mu, \nu) \in (r, s) + \mathrm{supp}(H^T), \tag{3.2}$$

where $(r, s) + \mathrm{supp}(H^T) := \{(r + i, s + j) : (i, j) \in \mathrm{supp}(H^T)\}$.
More generally, if $\mathcal{T}$ is an arbitrary subset of $\mathbf{Z} \times \mathbf{Z}$, $\mathbf{v}$ satisfies the parity checks of the code on $\mathcal{T}$, if it satisfies them in every point $(r, s) \in \mathcal{T}$, i.e.

$$\left(\mathbf{v}H^T, z_1^\mu z_2^\nu\right) = 0, \qquad \forall (\mu, \nu) \in \mathcal{T} + \mathrm{supp}(H^T) \tag{3.3}$$

where $\mathcal{T} + \mathrm{supp}(H^T) := \bigcup_{(r,s)\in\mathcal{T}} \Big((r, s) + \mathrm{supp}(H^T)\Big)$.

Letting $H^T(z_1, z_2) = \sum_{ij} H_{ij}^T z_1^i z_2^j$, condition (3.2) reduces to the following system of linear equations

$$\sum_{(i,j)\ \in\ \mathrm{supp}(H^T)} v(\mu - i, \nu - j) H_{ij}^T = 0, \qquad \forall (\mu, \nu) \in (r, s) + \mathrm{supp}(H^T), \tag{3.4}$$

and hence to the system of all difference equations which regard the sample $v(r, s)$. Analogously, $\mathbf{v}$ meets condition (3.3) if all difference equations involving the samples $v(r, s)$, with (r, s) in $\mathcal{T}$, are satisfied. In FIG.2, each dashed polygon intersecting $\mathcal{T}$ represents the coordinates $(\mu - i, \nu - j)$ of the samples which appear in a system like (3.4).

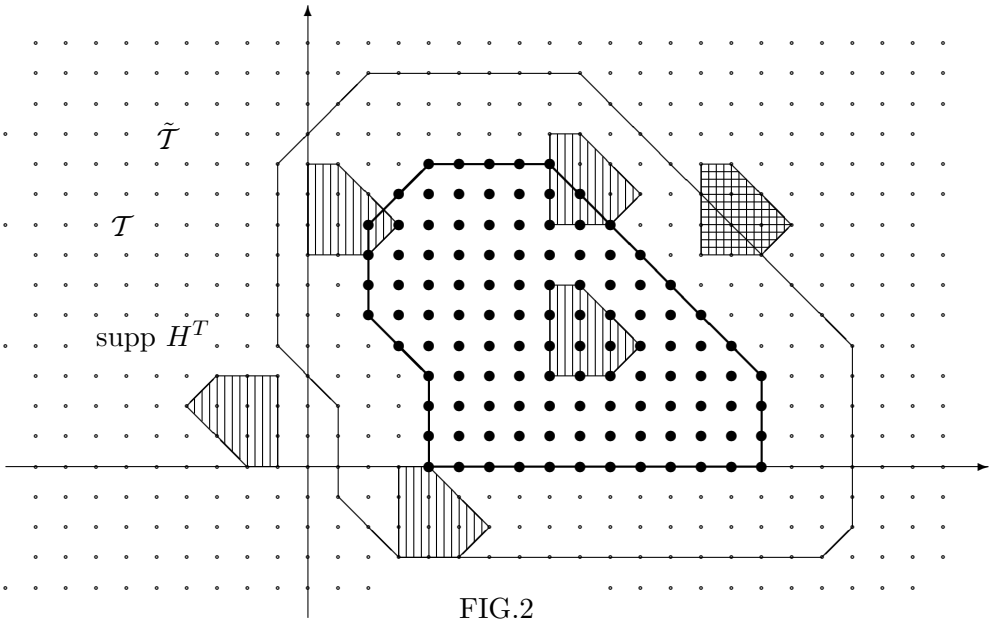


FIG.2

Clearly, when verifying whether $\mathbf{v}$ satisfies the parity checks of the code on $\mathcal{T}$, we take into account not only the samples on $\mathcal{T}$, but also those which belong to an appropriate set $\tilde{\mathcal{T}} \supseteq \mathcal{T}$. The remaining tests we have to perform, when deciding whether $\mathbf{v}$ is a codeword, are represented by systems of difference equations which involve only the samples of $\mathbf{v}$ on $C\mathcal{T}$. Some of them, however, utilize again the samples on $\tilde{\mathcal{T}} \setminus \mathcal{T}$, as suggested by the chequered polygon in FIG.2.
So, it could happen that the data on $\tilde{\mathcal{T}}$ allow to satisfy the parity checks on $\mathcal{T}$, yet none selection of the data on $C\tilde{\mathcal{T}}$ makes possible the fulfillment of the parity checks on $C\tilde{\mathcal{T}}$. Otherwise stated, the specific assignment of the values of $\mathbf{v}$ on $\tilde{\mathcal{T}}$ compromises any possibility of extending the data on $\tilde{\mathcal{T}}$ into a legal codeword.
In these situations, the natural question arises whether such an extension could be made possible by changing only the data which are "close" to the border of $\tilde{\mathcal{T}}$. More precisely, we wonder whether there is a positive integer δ, such that any sequence $\mathbf{v}$, satisfying the parity checks of the code on $\mathcal{T}^\delta$, can be modified into a codeword $\mathbf{w}$, which coincides with $\mathbf{v}$ on the window $\mathcal{T}$.
A positive answer is very important from the *syndrome decoder* point of view. Actually, when the parity checks of the code are verified in $\mathcal{T}^\delta$, we can assume the restriction $\mathbf{v}|\mathcal{T}$ as correct and, whenever the parity checks fail in some point $(r, s) \notin \mathcal{T}^\delta$, we have to modify only the values of $\mathbf{v}$ on $C\mathcal{T}$.

Generally, neither the completeness assumption nor the more restrictive hypothesis that $\mathcal{C}$ is a convolutional code, imply that the code $\mathcal{C}$ exhibits the aforementioned features. As we shall see, these constitute the exact counterpart, from an internal point of view, of the condition for the existence of an injective encoder (stated in Proposition 3.1), and provide an equivalent definition of 2D basic codes.
The formal definition of these properties will be assumed as a further constraint on the structure of $\mathcal{C}$.

(e) [Extension Property] *Let $\mathcal{C} = \ker H^T$. There exists a positive integer δ such that, for every finite subset $\mathcal{S} \subset \mathbf{Z} \times \mathbf{Z}$ and every $\mathbf{v} \in \mathcal{F}_\infty^n$, which satisfies on $\mathcal{S}^\delta$ the parity checks of the code, there is a codeword $\mathbf{w} \in \mathcal{C}$ such that*

$$\mathbf{w}|\mathcal{S} = \mathbf{v}|\mathcal{S}. \tag{3.5}$$

Lemma 3.2 *Let $\mathcal{C} = \ker H^T(z_1, z_2)$ be a code satisfying the extension property. Then property (e) holds for all (nonnecessarily finite) subsets of $\mathbf{Z} \times \mathbf{Z}$.*

PROOF Assume that the sequence $\mathbf{v} \in \mathcal{F}_\infty^n$ satisfies the parity checks of the code on an infinite set $\mathcal{S}^\delta$, and let $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \mathcal{S}_3 \subset ...$ be a sequence of finite sets in $\mathbf{Z} \times \mathbf{Z}$ invading $\mathcal{S}$. Since $\mathbf{v}$ satisfies the parity checks on $\mathcal{S}_i^\delta$, $i = 1, 2, 3, ...$, there exists a sequence of codewords $\mathbf{w}_i$, $i = 1, 2, 3, ...$ such that $\mathbf{w}_i|\mathcal{S}_i = \mathbf{v}|\mathcal{S}_i$.
As $\mathcal{F}_\infty^n$ is sequentially compact, there is a subsequence $\{\mathbf{w}_{\nu_j}\}$ of $\{\mathbf{w}_i\}$ converging to $\mathbf{w} \in \mathcal{F}_\infty^n$. The proposition is proved by observing that

- as $\mathbf{w}_{\nu_j} \in \mathcal{C}$ for every j, by the completeness of $\mathcal{C}$, $\mathbf{w} \in \mathcal{C}$;
- since $\mathbf{w}|\mathcal{S}_{\nu_j} = \mathbf{w}_{\nu_j}|\mathcal{S}_{\nu_j} = \mathbf{v}|\mathcal{S}_{\nu_j}$ for every ν_j, it follows that $\mathbf{w}|\mathcal{S} = \mathbf{v}|\mathcal{S}$ ■

Proposition 3.3 *The extension property implies controllability.*

PROOF Let $\mathbf{w} \in \mathcal{C}$, with $\mathcal{C}$ a code satisfying property (e), and consider a finite set $\mathcal{S}$. Define $\mathbf{v} \in \mathcal{F}_\infty^n$ as follows

$$v(h,k) = \begin{cases} 0, & \text{for every } (h,k) \in \mathcal{S}^\delta; \\ w(h,k), & \text{otherwise} \end{cases}$$

where δ is like in (e).
Since $\mathbf{v}$ satisfies the parity checks in $\mathcal{T} = \mathcal{S} \cup C(\mathcal{S}^{2\delta})$, by the previous lemma there is a codeword $\bar{\mathbf{v}} \in \mathcal{C}$ such that $\bar{\mathbf{v}}|\mathcal{T} = \mathbf{v}|\mathcal{T}$. Clearly $(\mathbf{w} - \bar{\mathbf{v}})$ is in $\mathcal{C}$, and $(\mathbf{w} - \bar{\mathbf{v}})|\mathcal{S} = \mathbf{w}|\mathcal{S}$. Moreover $(\mathbf{w} - \bar{\mathbf{v}})|C(\mathcal{S}^{2\delta}) = 0$, implies that $\text{supp}(\mathbf{w} - \bar{\mathbf{v}}) \subseteq \mathcal{S}^{2\delta}$, so $\mathcal{C}$ satisfies property (d) ■

Remark III As a consequence of the above proof, given any sequence $\mathbf{v} \in \mathcal{F}_\pm^n$ which satisfies the parity checks of the code in $\mathcal{S}^\delta$, there is a codeword $\mathbf{w}$ which coincides with $\mathbf{v}$ in $\mathcal{S}$ and whose support does not exceed $\mathcal{S}^{2\delta}$.

Extension property and controllability are very close each other. To further highlight the strict connection between the two notions, we will show that, for a complete code, property (e) is equivalent to property (e_1), (which represents the natural counterpart of (d_1), and hence is called "strong controllability" in [7]).

(e_1) *Let $\mathcal{C} = \ker H^T$. There exists a positive integer ρ such that for every pair of subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathbf{Z} \times \mathbf{Z}$, with $d(\mathcal{S}_1, \mathcal{S}_2) > 2\rho$, and for every pair of sequences $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathcal{F}_\infty^n$, which satisfy the parity checks of the code on $\mathcal{S}_1^\rho$ and $\mathcal{S}_2^\rho$ respectively, a codeword $\mathbf{w} \in \mathcal{C}$ exists, such that*

$$\mathbf{w}|\mathcal{S}_1 = \mathbf{v}|\mathcal{S}_1 \quad \textit{and} \quad \mathbf{w}|\mathcal{S}_2 = \mathbf{v}|\mathcal{S}_2. \tag{3.6}$$

The proof of the equivalence between (e) and (e_1) is included in the following summarizing proposition, which provides also the connections between basic codes, introduced from an "external" point of view in Proposition 3.1, and codes endowed with the extension property.

Proposition 3.4 [Equivalent characterizations of Basic Codes] *Let $\mathcal{C} \subseteq \mathcal{F}_\infty^n$. The following are equivalent:*
(1) $\mathcal{C} = \mathrm{Im}G$, $G \in \mathcal{F}_\pm^{k\times n}$ ℓZP;
(2) $\mathcal{C} = \ker H^T$, $H^T \in \mathcal{F}_\pm^{n\times(n-k)}$ rZP;
(3e) $\mathcal{C}$ satisfies property (e);
(3e$_1$) $\mathcal{C}$ satisfies property (e$_1$).

PROOF **(1) $\Leftrightarrow$ (2)** See Proposition 3.1.

(2) $\Rightarrow$ (3e) Assume that $H^T \in \mathcal{F}_\pm^{n\times(n-k)}$ is rZP and $W \in \mathcal{F}_\pm^{(n-k)\times n}$ is a polynomial left inverse of H^T.
Let $\delta_1 = \max\{|i| + |j| : \ (H^T, z_1^i z_2^j) \neq 0\}$, $\delta_2 = \max\{|i| + |j| : \ (W, z_1^i z_2^j) \neq 0\}$, and $\delta = \delta_1 + \delta_2$. If $\mathbf{v} \in \mathcal{F}_\infty^n$ satisfies the parity checks of the code on a finite set $\mathcal{S}^\delta \subseteq \mathbf{Z}\times\mathbf{Z}$, the series $\mathbf{a} = \mathbf{v}H^T$ satisfies $(\mathbf{a}, z_1^\mu, z_2^\nu) = 0$ for every $(\mu,\nu) \in \mathcal{S}^\delta + \ \mathrm{supp}(H^T)$. Since supp $(H^T) \subseteq \mathcal{S}^{\delta_1}$, we have

$$\mathcal{S}^{\delta_2} \subseteq \mathcal{S}^\delta + \mathrm{supp}(H^T) \subseteq C(\mathrm{supp}(\mathbf{a})).$$

Introduce next the series $\mathbf{x} = \mathbf{a}W$. As $\mathrm{supp}(\mathbf{x}) \subseteq \mathrm{supp}(\mathbf{a}) + \mathrm{supp}(W)$ and $\mathrm{supp}(\mathbf{a}) \subseteq C\mathcal{S}^{\delta_2}$, it follows that

$$\mathrm{supp}(\mathbf{x}) \subseteq C\mathcal{S}^{\delta_2} + \mathrm{supp}(W) \subseteq (C\mathcal{S}^{\delta_2})^{\delta_2} = C\mathcal{S}$$

and therefore $\mathbf{x}|\mathcal{S} = 0$.
Finally, let $\mathbf{w} := \mathbf{v} - \mathbf{x}$. As a consequence of $\mathbf{x}H^T = \mathbf{a}WH^T = \mathbf{a} = \mathbf{v}H^T$, we have $\mathbf{w}H^T = (\mathbf{v} - \mathbf{x})H^T = 0$, which implies $\mathbf{w} \in \mathcal{C}$. Moreover $\mathbf{w} \mid \mathcal{S} = (\mathbf{v} - \mathbf{x})|\mathcal{S} = \mathbf{v}|\mathcal{S}$.

(3e) $\Rightarrow$ (2) Let $\mathcal{C}$ satisfy the extension property. Then $\mathcal{C}$ is a convolutional code, and can be described as $\mathcal{C} = \ker H^T$, with H^T rFP. To prove that H^T is rZP, by Proposition A5 it's sufficient to show that the equation

$$\mathbf{x}H^T = \mathbf{a} \tag{3.7}$$

admits an L-polynomial solution for all vectors $\mathbf{a}$ in $\mathcal{F}_\pm^{n-k}$.
As H^T has full column rank over $\mathbf{F}(z_1, z_2)$, equation (3.7) admits a solution $\mathbf{v} \in \mathcal{F}_\infty^n$. To complete the proof we will show that there is a codeword $\mathbf{w}$, differing from $\mathbf{v}$ on a finite set $\mathcal{T}$. Actually, in this case we have

$$(\mathbf{v} - \mathbf{w})H^T = \mathbf{a}, \qquad \mathrm{supp}(\mathbf{v} - \mathbf{w}) \subseteq \mathcal{T}$$

and, consequently, $(\mathbf{v} - \mathbf{w})$ is an L-polynomial solution of (3.7).
By assumption, $\mathbf{v}$ satisfies the parity checks of $\mathcal{C}$ on the set

$$\mathcal{H} = \{(r,s) : ((r,s) + \ \mathrm{supp}(H^T)) \cap \ \mathrm{supp}(\mathbf{a}) = \emptyset\}$$

whose complement is a finite set.

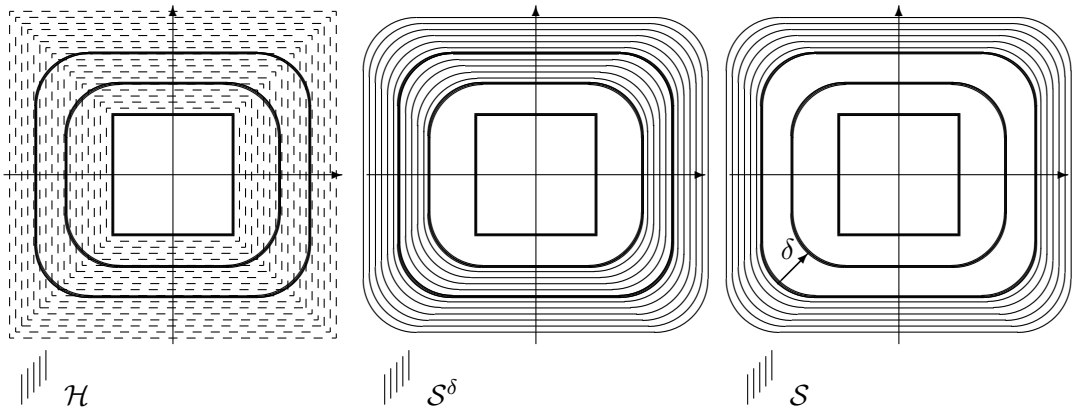


FIG.3

Clearly,(see FIG.3), a set $\mathcal{S}$ exists such that $C\mathcal{S}$ is finite and $\mathcal{S}^\delta \subseteq \mathcal{H}$. As $\mathbf{v}$ satisfies the parity checks of the code in $\mathcal{S}^\delta$, by $(3e)$ there is a codeword $\mathbf{w}$ such that $\mathbf{w}|\mathcal{S} = \mathbf{v}|\mathcal{S}$. Therefore $\mathbf{v}$ and $\mathbf{w}$ differ only on a (finite) subset of $C\mathcal{S}$.

$(\mathbf{3e}) \Leftrightarrow (\mathbf{3e}_1)$ The proof can be performed along the same lines followed in showing the equivalence $(\mathbf{3d}) \Leftrightarrow (\mathbf{3d}_1)$ in Proposition 2.2 ■

Once an encoder commits an information message to the corresponding codeword, the encoded message $\mathbf{w}$ is transmitted over a noisy channel. Generally, the received sequence $\mathbf{r}$ not only differs from the original message $\mathbf{w}$, but also does not belong to the code $\mathcal{C}$. So we need to project $\mathbf{r}$ on the codewords space, in order to find out the best approximation of $\mathbf{r}$ in $\mathcal{C}$, namely a codeword $\hat{\mathbf{r}}$ whose distance from $\mathbf{r}$ is minimal. Often $\hat{\mathbf{r}}$ differs from the transmitted codeword $\mathbf{w}$ by a nonzero reconstruction error

$$\mathbf{e} = \hat{\mathbf{r}} - \mathbf{w} \tag{3.8}$$

which is a codeword too. Having no possibility of finding out $\mathbf{e}$, all we can do is to reconstruct $\hat{\mathbf{u}}$, the input sequence corresponding to $\hat{\mathbf{r}}$, and assume it as an approximation of the correct information sequence. This step is performed by a decoder, namely a right inverse of the encoder matrix $G(z_1, z_2)$, which produces the sequence $\hat{\mathbf{u}}$, when receiving $\hat{\mathbf{r}} = \hat{\mathbf{u}}G$ as its input.

It can be easily realized from (3.8) that if there exist finite codewords generated by infinite information sequences, then a finite error $\mathbf{e}$ in the reconstruction of the codeword $\mathbf{w}$ could produce an infinite error when decoding $\hat{\mathbf{r}}$ instead of $\mathbf{w}$. Such *catastrophic errors*, however, are avoided when, to preserve the injectivity property, we confine ourselves to the class of basic encoders. Indeed, if G is a left zero prime L-polynomial matrix, it admits at least one L-polynomial right inverse $G^{-1}(z_1, z_2)$, and therefore every finite codeword $\mathbf{e}$ in $\mathcal{C}$ is generated by one (and only one) finite input sequence $\mathbf{u_e} = \mathbf{e}G^{-1}$. An interesting consequence of the above reasoning is that, when achieving the injectivity of a convolutional encoder, one also guarantees the existence of a polynomial decoder, thus ruling out the possibility of catastrophic errors.

Remark IV As G defines an injective input-output map, a decoder represented by a rational right inverse $G^{-1}(z_1, z_2)$ of G would associate to any finite codeword $\mathbf{e}$ in $\mathcal{C}$, the same input sequence $\mathbf{u}_e$ as the polynomial decoder. In this case, however, expression $\mathbf{e}G^{-1}$ would be meaningless when $\mathbf{e}$ is not polynomial, and we should restrict our attention to codewords whose supports do not extend to the whole discrete plane.

4 Dual codes and syndrome decoders

The structure of 2D codes, as discussed in section 2, can be clarified further through the duality relation between finite and complete codes.
When referring to a *finite code* $\mathcal{C}$ of length n, we mean (see Example I of section 2) a submodule of $\mathcal{F}_{\pm}^n$ defined as $\mathcal{C} = \mathrm{Im}_{\pm}G = \{\mathbf{w} = \mathbf{u}G : \mathbf{u} \in \mathcal{F}_{\pm}^h\}$, where G denotes an arbitrary matrix in $\mathcal{F}_{\pm}^{h\times n}$. On the other hand, a complete (and, in particular, a convolutional) code $\mathcal{D}$ of length n is defined as $\mathcal{D} = \ker H^T = \{\mathbf{w} \in \mathcal{F}_{\infty}^n : \mathbf{w}H^T = 0\}$, with H^T an arbitrary matrix in $\mathcal{F}_{\pm}^{n\times q}$.
The two are dual concepts, that play together in the encoding and decoding processes. In most cases, it is quite reasonable to assume that 2D information signals are finite support, and, therefore, finite codes are easily regarded as the result of an encoding operation. Even if complete and convolutional codes can be introduced by simply extending this point of view to infinite information signals, in algebraic terms, however, it's very convenient to give them a different interpretation.
A complete code is more naturally viewed as a family of $\mathbf{F}-$ valued linear functions on the space of the finite sequences, via the canonical algebraic duality [16] between $\mathcal{F}_{\pm}^n$ and the space of linear functionals $L(\mathcal{F}_{\pm}^n)$. So our philosophy will be to characterize a finite code $\mathcal{C}$ as the set of codewords which are in the kernel of a suitable space of linear functionals, and viceversa, a (complete) dual code $\mathcal{D}$ as the set of linear parity checks necessary to decide whether a finite sequence is a legal codeword.
The duality properties find an obvious application in the syndrome decoders synthesis. Indeed, a complete characterization of the syndrome decoders of $\mathcal{C}$ can be achieved by resorting to a systematical analysis of the class of its dual codes.

Introduce in $\mathcal{F}_{\pm}^m \times \mathcal{F}_{\infty}^m$ the following non degenerate bilinear form

$$\langle \cdot, \cdot \rangle_m : \mathcal{F}_{\pm}^m \times \mathcal{F}_{\infty}^m \to \mathbf{F}$$

defined by : $\langle \mathbf{u}, \mathbf{v} \rangle_m = (\mathbf{u}\mathbf{v}^T, 1) = \sum_{i,j\in\mathbf{Z}} u(i,j)v^T(-i,-j)$.
Two vectors $\mathbf{u} \in \mathcal{F}_{\pm}^m$ and $\mathbf{v} \in \mathcal{F}_{\infty}^m$ are called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle_m = 0$. Given any submodule $\mathcal{M}$ of $\mathcal{F}_{\pm}^m$, its *orthogonal complement* $\mathcal{M}^{\perp}$, is constituted by all the vectors of $\mathcal{F}_{\infty}^m$ which are orthogonal to $\mathcal{M}$. Similarly, every submodule $\mathcal{N}$ of $\mathcal{F}_{\infty}^m$ identifies an orthogonal complement $\mathcal{N}^{\perp}$ in $\mathcal{F}_{\pm}^m$.
The space $\mathcal{F}_{\infty}^m$ can be viewed as $L(\mathcal{F}_{\pm}^m)$, the algebraic dual of $\mathcal{F}_{\pm}^m$. In fact, we can associate with every $\mathbf{v} \in \mathcal{F}_{\infty}^m$ the linear functional on $\mathcal{F}_{\pm}^m$ defined by

$$f_v(\cdot) = \langle \cdot, \mathbf{v} \rangle_m \tag{4.1}$$

and, conversely, every linear functional on $\mathcal{F}_\pm^m$ can be represented as in (4.1) for an appropriate choice of $\mathbf{v} \in \mathcal{F}_\infty^m$. The identification of $\mathcal{F}_\infty^m$ with $L(\mathcal{F}_\pm^m)$ makes it possible to use some results, not valid for arbitrary pairs of dual spaces [16].

Let $\mathcal{C}$ be a finite code, described as the image of the map

$$G : \mathcal{F}_\pm^k \to \mathcal{F}_\pm^n : \mathbf{u} \mapsto \mathbf{u}G,$$

and consider the map

$$G^T : \mathcal{F}_\infty^n \to \mathcal{F}_\infty^k : \mathbf{v} \mapsto \mathbf{v}G^T.$$

G and G^T are dual mappings, since $\langle \mathbf{u}G, \mathbf{v}\rangle_n = (\mathbf{u}G\mathbf{v}^T, 1) = (\mathbf{u}(\mathbf{v}G^T)^T, 1) = \langle \mathbf{u}, \mathbf{v}G^T\rangle_k$. By resorting to the well known relations

$$(\mathbf{A}) \qquad (\mathrm{Im}_\pm G)^\perp = \ker G^T \qquad (\ker G^T)^\perp = \mathrm{Im}_\pm G,$$

we induce a bijective correspondence between finite codes of length n, represented as images of appropriate L-polynomial matrices, and complete codes of the same length, described as kernels of L-polynomial matrices. This correspondence associates a finite code $\mathrm{Im}_\pm G$, with its dual, namely the $\mathcal{F}_\pm-$ module $\ker\, G^T \subseteq \mathcal{F}_\infty^n$ of all the parity checks of the code. Viceversa, the dual of a complete code $\ker\, G^T$ is the module $\mathrm{Im}_\pm G \subseteq \mathcal{F}_\pm^k$ of its parity checks.
As a straightforward consequence of (A), one gets

$$(\mathbf{B}) \qquad (\mathrm{Im}_\pm G)^{\perp\perp} = \mathrm{Im}_\pm G \qquad (\ker G^T)^{\perp\perp} = \ker G^T,$$

which means that every code can be exactly reconstructed from the space of its parity checks.

The duality between complete and finite codes can be better understood by analyzing the correspondence between convolutional codes and a particular subclass of finite codes. As we have seen, every complete code can be described as the kernel, in $\mathcal{F}_\infty^n$, of an L-polynomial matrix, while only a convolutional code, i.e. the kernel of a rFP matrix, can be represented as the image of an L-polynomial matrix (see [7] and Proposition 2.2). Lemma 4.1, below, shows that for finite codes a dual situation holds. Actually they are always the images of L-polynomial matrices, but only the images of $\ell FP matrices can be expressed as kernels.$

Lemma 4.1 *Let $\mathcal{C}$ be a submodule of $\mathcal{F}_\pm^n$. $\mathcal{C}$ is the kernel of a L-polynomial matrix if and only if there exists a $\ell FP matrix$,$\bar{G}(z_1, z_2)$, such that $\mathcal{C} = \mathrm{Im}_\pm \bar{G}$.*

Proof Let $\mathcal{C} = \mathrm{Im}_\pm \bar{G}$ with $\bar{G} \in \mathcal{F}_\pm^{k\times n}$ $\ell FP, and consider a full column rank matrix$,$\mathrm{H}^T \in \mathcal{F}_\pm^{n\times(n-k)}$, such that $\bar{G}H^T = 0$.
Clearly, if $\mathbf{w} \in \mathcal{C}$, then $\mathbf{w} = \mathbf{u}\bar{G}$ for some $\mathbf{u} \in \mathcal{F}_\pm^k$, and $\mathbf{w}H^T = (\mathbf{u}\bar{G})H^T = \mathbf{u}(GH^T) = 0$, so $\mathbf{w} \in \ker_\pm H^T$. On the other hand, if $\mathbf{w} \in \mathcal{F}_\pm^n$ is in $\ker_\pm H^T$, it belongs to the subspace of $\mathbf{F}(z_1, z_2)^n$ orthogonal to the columns of H^T, and spanned by the rows of $\bar{G}$. Then there exists $\mathbf{f} \in \mathbf{F}(z_1, z_2)^k$ such that $\mathbf{f}\bar{G} = \mathbf{w}$. As $\bar{G}$ is $\ell FP, by Lemma A.1 \mathbf{f} can be chosen in \mathrm{F}_\pm^k$.
So $\mathbf{w}$ belongs to $\mathrm{Im}_\pm \bar{G}$.

Viceversa, let $\mathcal{C} = \ker_{\pm} H^T$ and consider any $\ell FP matrix \bar{G}(z_1, z_2) \in \mathcal{F}_{\pm}^{(n-p)\times p}$ such that $\bar{G}\bar{H}^T = 0$. Using the same arguments as in the first part of the proof, one shows that $\mathcal{C} = \mathrm{Im}_{\pm}\bar{G}$ ■

By completing the duality relations (A) with

$$(\mathbf{C}) \qquad (\mathrm{Im}G^T)^{\perp} = \ker_{\pm} G \qquad (\ker_{\pm} G)^{\perp} = \mathrm{Im}G^T,$$

it's immediate to prove proposition below

Proposition 4.2 *Let $G : \mathcal{F}_{\pm}^k \to \mathcal{F}_{\pm}^n$ and $G^T : \mathcal{F}_{\infty}^n \to \mathcal{F}_{\infty}^k$ be dual mappings. The following are equivalent:*
(1) the finite code $\mathrm{Im}_{\pm}G$ *can be represented as the kernel of an L-polynomial matrix;*
(2) the complete code $\ker G^T$ *is convolutional, i.e. it can be described as the image of an L-polynomial matrix* ■

Remark V Properties (A) and (C) together with Lemma 4.1 allow to obtain an alternative proof of the equivalence between (1) and (2) in Proposition 2.2.
Actually, if $\mathcal{D} = \mathrm{Im}G^T$ is a convolutional code, as a consequence of (C) $\mathcal{D}$ is the dual of the finite code $\mathcal{C} := \ker_{\pm} G$. Then, by Lemma 4.1, there exists a $\ell FP matrix \mathrm{H}(\mathrm{z}_1, z_2)$ such that $\mathcal{C} = \mathrm{Im}_{\pm}H$ and hence, by (A), $\mathcal{D} = (\mathrm{Im}_{\pm}H)^{\perp} = \ker H^T$. So $\mathcal{D}$ is the kernel of a rFP L-polynomial matrix.
Conversely, if $\mathcal{D} = \ker H^T$ is a complete code and $H^T(z_1, z_2)$ is rFP, by (A) $\mathcal{D}$ is the dual of the finite code $\mathcal{C} := \mathrm{Im}_{\pm}H$, with $H\ \ell FP. By Lemma 4.1, there exists a matrix \mathrm{G}(\mathrm{z}_1, z_2)$ such that $\mathcal{C} = \ker_{\pm} G$, and therefore, by (C), $\mathcal{D} = \mathrm{Im}G^T$ is a convolutional code.

It's quite clear that every code described as the kernel of an L-polynomial matrix $H^T(z_1, z_2)$, admits H^T as a syndrome decoder, since a sequence $\mathbf{v}$ belongs to the code if and only if $\mathbf{v}H^T = 0$. Hence every complete code admits a syndrome decoder, while, among finite codes, only those which are the image of a $\ell FP matrix have this property. When a (finite or in fin$ *polynomial matrix, the following proposition provides an algorithm to find out a syndrome decoder.*

Proposition 4.3 *(i) If $\mathcal{C} = \mathrm{Im}_{\pm}\bar{G}$ is a finite code with $\bar{G}(z_1, z_2) \in \mathcal{F}_{\pm}^{k\times n}$ ℓFP, every L−polynomial matrix $\bar{H}^T(z_1, z_2)$ of rank $n-k$, satisfying $\bar{G}H^T = 0$, is a syndrome decoder of $\mathcal{C}$;*

(ii) If $\mathcal{C} = \mathrm{Im}G$ is a convolutional code of rank k, then every rFP L-polynomial matrix $\bar{H}^T(z_1, z_2)$ of rank $n-k$ satisfying $G\bar{H}^T = 0$ is a syndrome decoder of $\mathcal{C}$.

Proof (**i**) It's obvious that, if $\mathbf{w}$ belongs to $\mathcal{C}$, then $\mathbf{w} = \mathbf{u}\bar{G}$, $\mathbf{u} \in \mathcal{F}_{\pm}^k$, satisfies $\mathbf{w}H^T = \mathbf{u}\bar{G}H^T = 0$.
Conversely, every $\mathbf{w} \in \mathcal{F}_{\pm}^n$ satisfying $\mathbf{w}H^T = 0$, belongs to the subspace of $\mathbf{F}(z_1, z_2)^n$ orthogonal to the columns of $H^T(z_1, z_2)$, which is spanned by the rows of $\bar{G}$. As $\bar{G}$ is $\ell FP, by Lemma A.1 \mathbf{w} is a linear combination over \mathrm{F}_{\pm}$ of the rows of $\bar{G}$, and therefore $\mathbf{w} \in \mathcal{C}$.

(**ii**) After factorizing G into $G = L\bar{G}$, with $\bar{G}\ \ell FP, the convolutional code \mathrm{C} can be equivalently represented as \mathrm{C}$ = Im $\bar{G}$ and condition $G\bar{H}^T = 0$ is equivalent to $\bar{G}\bar{H}^T = 0$. So we are reduced to

prove

$$\text{Im } \bar{G} = \ker \bar{H}^T. \tag{4.2}$$

Since $\bar{H}^T$ is rFP, $\ker \bar{H}^T$ is convolutional. As convolutional codes can be uniquely reconstructed from the submodule of the finite codewords (see (d_3)), it will be sufficient to show that (4.2) holds when restricted to $\mathcal{F}_{\pm}^n$, namely $(\text{Im}\bar{G})_f = \text{Im}_{\pm}\bar{G} = \ker_{\pm}\bar{H}^T$. But that's just what has been shown in part (i) ■

Note that the number of the parity checks we have to apply to a sequence $\mathbf{v}$ does not exceed the number of the columns of $H^T(z_1, z_2)$. Implementing a parity check, however, generally involves an infinite number of steps, unless $\mathcal{C}$ is a finite code, described as $\ker_{\pm}H^T$. In this case only a finite number of steps is required to decide whether $\mathbf{v}$ belongs to $\mathcal{C}$, if an upper bound on the diameter of its support is a priori known.

We conclude this section by focusing our attention on the problem of obtaining syndrome decoders for a finite code $\mathcal{C}$, which cannot be represented as the kernel of an L-polynomial matrix. As a general result, we already know that, if $\mathcal{C} = \text{Im}_{\pm}G$ is a finite code, the dual code $\mathcal{D} = \ker G^T$ allows to identify $\mathcal{C}$ as $\mathcal{D}^{\perp}$. The differences from case (i) in Proposition 4.3 come from the fact that a representation of $\mathcal{D}$ as the image of an L-polynomial matrix is no more available.
As the module of the parity checks cannot be generated by the columns of an L-polynomial matrix, the best we can do is to extract from

$$\mathcal{D} = \ker G^T = \{\mathbf{v} \in \mathcal{F}_{\infty}^n : \mathbf{v}G^T = 0\}$$

the submodule

$$\mathcal{D}_f := \ker_{\pm} G^T = \{\mathbf{v} \in \mathcal{F}_{\pm}^n : \mathbf{v}G^T = 0\}$$

and to represent it as the module generated by the rows of an $n \times p$ $\ell FP matrix \bar{H}(z_1, z_2)$, i.e. $\mathcal{D}_f = \text{Im}_{\pm}\bar{H}$.
Clearly, any codeword $\mathbf{w} \in \mathcal{C}$ satisfies $\mathbf{w}\bar{H}^T = 0$. Letting $G = L\bar{G}$, with $\bar{G}$ $\ell FP, we have that the \text{F}_{\pm}$-module of the finite sequences in $\ker \bar{H}^T$ is given by $\bar{\mathcal{C}} := \text{Im}_{\pm}\bar{G} \supset_{\neq} \mathcal{C}$ where the inclusion is proper because of the assumption on $\mathcal{C}$. This means that the syndrome decoder $\bar{H}^T$ accepts as legal codewords even sequences in $\bar{\mathcal{C}} \setminus \mathcal{C}$, that are not elements of the code.

5 State space realization of encoders and decoders

Assigning an encoder via an L-polynomial matrix $G(z_1, z_2)$, corresponds to describe the algorithm which transforms an input information sequence into an output codeword, and hence to specify only what happens at the terminals of an encoding device. The realization problem consists in obtaining a mathematical model of some "machine" that implements the input-output map. In other words, a state-space realization shows how the encoding algorithm proceeds, by explicitly displaying the corresponding evolution of the memory function.
In general there is not a unique way to find out an algorithm which produces the input-output map of a convolutional encoder. So we have to introduce some a priori assumptions, like in the 1D case, on the class of the mathematical models to use for this

purpose. Moreover, as there is no natural notion of causality in the discrete plane, we need also to specify the partial ordering which underlays the recursive data processing. The class of 2D models more extensively investigated in the literature is that of 2D systems, for which the state equation updates according with a quarter plane causality notion. In this section we shall analyze to what extent 2D systems can be used for realizing 2D encoders and decoders.

A (quarter plane causal) 2D system $\Sigma = (A_1, A_2, B_1, B_2, C, D)$ is given by the following equations [11]

$$\begin{aligned} x(i+1,j+1) &= x(i,j+1)A_1 + x(i+1,j)A_2 \\ &+ u(i,j+1)B_1 + u(i+1,j)B_2 \\ w(i,j) &= x(i,j)C + u(i,j)D, \end{aligned} \tag{5.1}$$

(2)

where the local state $x(i,j)$ is a ν-dimensional vector over $\mathbf{F}$, input and output functions take values in $\mathbf{F}^k$ and $\mathbf{F}^n$ respectively, and A_1, A_2, B_1, B_2, C and D are matrices of suitable dimensions, with entries in $\mathbf{F}$.
When trying to implement a 2D encoder through a 2D system (5.1), a preliminary step is to eliminate the state variables, in order to make explicit the input-output relation it produces. To this purpose, some restrictive hypotheses are introduced on both the supports of the input signals and the initial conditions of the system, which are formalized as follows:

(i) [Past finite support of the input $\mathbf{u}$] For every $(l,m) \in \mathbf{Z} \times \mathbf{Z}$, the corresponding past cone $\{(i,j) : i \leq l, j \leq m\}$ intersects the support of $\mathbf{u}$ in a finite number of points.

(**ii**) [Zero initial conditions] For every $(l,m) \in \mathbf{Z} \times \mathbf{Z}$, $\mathrm{supp}(\mathbf{u}) \cap \{(i,j) : i \leq l, j \leq m\} = \emptyset$ implies $x(l,m) = 0$.

Under assumptions (i) and (ii), the system output $\mathbf{w}$ corresponding to the input sequence $\mathbf{u}$ is given by $\mathbf{w} = \mathbf{u}G_\Sigma$, where

$$G_\Sigma(z_1, z_2) = D + (B_1 z_1 + B_2 z_2)(I - A_1 z_1 + A_2 z_2)^{-1} C, \tag{5.2}$$

is the *transfer matrix* of Σ.

It's a well-known result [17] that every proper rational 2D matrix $G(z_1, z_2)$ is the transfer matrix of a suitable 2D state model Σ. Since a convolutional (in particular, basic) code $\mathcal{C}$ can always be thought of as the image of a $k \times n$ matrix G with elements in $\mathbf{F}[z_1, z_2]$, the submodule of the codewords of $\mathcal{C}$ with past finite support can be generated by an appropriate state model (5.1), whose transfer matrix G_Σ coincides with G. Actually, when condition (ii) is met, such codewords of $\mathcal{C}$ are obtained by applying to Σ all input sequences with past finite supports. Σ will be called a "realization" (or a state model) of the encoder G.

To obtain a reliable realization Σ, however, it is not enough to check whether $G_\Sigma = G$, and some additional aspects have to be taken into account.

a) If the state-output transfer matrix $(I - A_1 z_1 - A_2 z_2)^{-1} C$ is not polynomial, local states x exist, which give rise to free output evolutions $x(I - A_1 z_1 - A_2 z_2)^{-1} C$ with infinite supports. Clearly such local states, when induced by noise, generate infinite error sequences in the encoding process.

b) If the input-state transfer matrix $(B_1 z_1 + B_2 z_2)(I - A_1 z_1 - A_2 z_2)^{-1}$ is not polynomial, finite support input sequences possibly produce infinite support sequences in the state space. Therefore Σ could remain indefinitely excited by a finite signal, even though the corresponding output dies out in a finite number of steps.

Both the previous drawbacks can be avoided if the inverse matrix $(I - A_1 z_1 - A_2 z_2)^{-1}$ is polynomial or, equivalently [10], if the characteristic polynomial of the system is unitary, i.e.

$$\det(I - A_1 z_1 - A_2 z_2) = 1. \tag{5.3}$$

2D systems satisfying condition (5.3) are called "finite memory", since they reach the zero state in a finite number of steps after zeroing the input signal. Since every matrix $G(z_1, z_2) \in \mathbf{F}[z_1, z_2]^{k \times n}$ admits a finite memory realization [18], it follows that every polynomial encoder, and a fortiori every basic polynomial encoder, can be synthesized by resorting to a 2D system with finite memory.

c) When implementing a complete transmission system, we have to realize both the encoder and the corresponding decoder via finite memory 2D state models (5.1), and hence to use an encoder-decoder pair with elements in $\mathbf{F}[z_1, z_2]$. So, the code $\mathcal{C}$ has to be the image of a matrix G which is $\ell ZP not only over \mathrm{F}_\pm$ (which amounts to assume that $\mathcal{C}$ is basic) but also over $\mathbf{F}[z_1, z_2]$. Namely, the ideal generated in $\mathbf{F}[z_1, z_2]$ by the maximal order minors of G is the ring $\mathbf{F}[z_1, z_2]$ itself. This condition guarantees the existence of a right inverse G^{-1} with elements in $\mathbf{F}[z_1, z_2]$, so that both G and G^{-1} have finite memory realizations.

The following proposition characterizes 2D basic codes which admit an encoder, $\ell ZPin\mathbf{F}[\mathrm{z}_1, z_2]$.

Proposition 5.1 [Basic Encoders in $\mathbf{F}[z_1, z_2]$] *Let $\mathcal{C}$ be a basic code of length n and rank k. The following are equivalent:*
(i) there exists a basic encoder $G_+(z_1, z_2) \in \mathbf{F}[z_1, z_2]^{k \times n}$ which is $\ell ZPin\mathbf{F}[z_1, z_2]$;
(ii) for any basic encoder $G(z_1, z_2) \in \mathcal{F}_\pm^{k \times n}$, with maximal order minors $m_i(G)$, $i = 1, 2, ...N = \binom{n}{k}$, there is a pair $(l, m) \in \mathbf{Z} \times \mathbf{Z}$ such that

$$\bigcup_{i=1}^{N} \mathrm{supp}\Big(m_i(G) z_1^l z_2^m\Big) \subseteq \mathbf{N} \times \mathbf{N} \tag{5.4}$$

and the ideal in $\mathbf{F}[z_1, z_2]$ generated by $m_i(G) z_1^l z_2^m$, $i = 1, 2, ..., N$ is the whole ring, i.e.

$$\Big(m_1(G) z_1^l z_2^m, m_2(G) z_1^l z_2^m, ..., m_N(G) z_1^l z_2^m\Big) = \mathbf{F}[z_1, z_2]; \tag{5.5}$$

(iii) given a basic encoder $G_+(z_1, z_2) \in \mathbf{F}[z_1, z_2]^{k \times n}$, in any factorization

$$G_+(z_1, z_2) = T_+(z_1, z_2) \bar{G}_+(z_1, z_2),$$

where the matrix $T_+(z_1,z_2)$ is a greatest left factor (g.l.f.) of $G_+(z_1,z_2)$ over $\mathbf{F}[z_1,z_2]$, $\bar{G}_+(z_1,z_2)$ is $\ell ZP over \mathbf{F}[z_1,z_2]$.

PROOF **(i)** $\Rightarrow$ **(ii)** If $G_+(z_1,z_2) \in \mathbf{F}[z_1,z_2]^{k\times n}$ is a basic encoder, $\ell ZP in \mathbf{F}[z_1,z_2]$, any equivalent basic encoder $G(z_1,z_2)$ differs from $G_+(z_1,z_2)$ in a unimodular matrix $U(z_1,z_2) \in \mathcal{F}_\pm^{k\times k}$, namely $G = UG_+$. As det U is a unit in $\mathcal{F}_\pm$, that is det $U = z_1^{\nu_1} z_2^{\nu_2}$, it follows that the i-th maximal order minor of G is obtained from the corresponding i-th minor of G_+ as $m_i(G) = z_1^{\nu_1} z_2^{\nu_2} m_i(G_+)$. So, assuming $l = -\nu_1$ and $m = -\nu_2$, we satisfy both (5.4) and (5.5).

(ii) $\Rightarrow$ **(iii)** Let G_+ be a basic encoder with elements in $\mathbf{F}[z_1,z_2]$, and consider the pair (l,m) such that the maximal order minors $m_i(G_+)$, $i = 1,2,...,N$, satisfy (5.4) and (5.5). Clearly l and m are nonpositive integers, and (5.5) implies that $m_i(G_+)$, $i = 1,2,...N$, generate in $\mathbf{F}[z_1,z_2]$ the principal ideal $\left(z_1^{-l} z_2^{-m}\right)$, i.e.

$$\left(m_1(G_+), m_2(G_+), ..., m_N(G_+)\right) = \left(z_1^{-l} z_2^{-m}\right).$$

By extracting from G_+ a g.l.f. $T_+(z_1,z_2)$, we obtain $G_+ = T_+ \bar{G}_+$. As the determinant of a g.l.f. of G_+ is the g.c.d. of its maximal order minors $m_i(G_+)$, $i = 1,2,...,N$, then det $T_+(z_1,z_2) = z_1^{-l} z_2^{-m}$.
Obviously, the maximal order minors of $\bar{G}_+$ are $m_1(G_+) z_1^l z_2^m$, $m_2(G_+) z_1^l z_2^m$, ..., $m_N(G_+) z_1^l z_2^m$, and therefore $\bar{G}_+$ is $\ell ZP in \mathbf{F}[z_1,z_2]$.

(iii) $\Rightarrow$ **(i)** It's immediate that G_+ and $\bar{G}_+$ are equivalent encoders of $\mathcal{C}$. Being $\ell ZP in \mathbf{F}[z_1,z_2]$, $\bar{G}_+$ is $\ell ZP als oin \mathrm{F}_\pm$ and therefore basic. ■

Example 4 Let $\mathbf{F} = GF(2)$. The following encoder

$$G(z_1,z_2) = \begin{bmatrix} 0 & z_1+1 & z_2 \\ 1 & 0 & 0 \end{bmatrix}$$

has maximal order minors z_1+1, z_2 and 0. Since z_2 is a unit in $\mathcal{F}_\pm$, G is $\ell ZP in \mathrm{F}_\pm$ and the code $\mathcal{C}$ generated by G is basic.
Neither G, nor (by the above proposition) any other equivalent encoder of $\mathcal{C}$, is $\ell ZP in \mathbf{F}[z_1,z_2]$.

Using a basic code $\mathcal{C}$ with the aforementioned properties in a trasmission system, requires to preliminarily design an encoder G $\ell ZP in \mathbf{F}[z_1,z_2]$, and to compute a decoder G^{-1} with elements in $\mathbf{F}[z_1,z_2]$. Correspondingly, two finite memory realizations for both G and G^{-1} have to be constructed, by resorting to 2D realization algorithms available in the literature [18].
Most of the computational effort of the above procedure is devoted to obtain the transfer matrix of the decoder, and to realize it as a state model. On the other hand, when relaxing the requirement that G^{-1} has to be a polynomial matrix, a considerable simplification is achieved by resorting to the inverse system technique. Actually, if $\Sigma = (A_1, A_2, B_1, B_2, C, D)$ is a realization of G, then $D = G(0,0)$ is right invertible, and for each right inverse of D, the corresponding inverse system

$$\Sigma^{-1}(D^{-1}) = (A_1 - CD^{-1}B_1, A_2 - CD^{-1}B_2, D^{-1}B_1, D^{-1}B_2, -CD^{-1}, D^{-1})$$

is a realization of a proper rational inverse of G. So, an interesting question is to investigate what conditions on Σ and G guarantee that the inverse system $\Sigma^{-1}(D^{-1})$ is finite memory, and therefore realizes a decoder of $\mathcal{C}$. The following proposition shows how the fact that $\Sigma^{-1}(D^{-1})$ inherits the finite memory property of Σ, only depends on the encoder G and possibly on the constant matrix D^{-1}, whereas the particular structure of the state space realization does not play any role.

Proposition 5.2 [19,20] *Let $\Sigma = (A_1, A_2, B_1, B_2, C, D)$ be a finite memory 2D system which realizes a $k \times n$ encoder $G(z_1, z_2)$, $\ell ZP over \mathbf{F}[z_1, z_2]$.*
For every right inverse D^{-1} of D, the following statements are equivalent
(i) the inverse system $\Sigma^{-1}(D^{-1})$ is finite memory;
(ii) $G(z_1, z_2)D^{-1}$ is a unimodular matrix in $\mathbf{F}[z_1, z_2]$;
(iii) G can be row bordered into an $n \times n$ matrix

$$V(z_1, z_2) = \begin{bmatrix} G(z_1, z_2) \\ K \end{bmatrix},$$

unimodular in $\mathbf{F}[z_1, z_2]$, by any constant full row rank $(n-k) \times n$ matrix K such that $KD^{-1} = 0$ ■

When $\mathcal{C}$ admits an encoder G which satisfies the above equivalent conditions, the inverse system technique can be applied to any other basic encoder $\tilde{G}$ of $\mathcal{C}$, $\ell ZP in \mathbf{F}[z_1, z_2]$. In fact, by Lemma A.6, G and $\tilde{G}$ differ in an $\mathbf{F}[z_1, z_2]-$unimodular matrix $V(z_1, z_2)$, that is $\tilde{G} = VG$. Therefore, if G meets condition (ii) of the above proposition,

$$\tilde{G}(z_1, z_2)\tilde{D}^{-1} = \Big[V(z_1, z_2)G(z_1, z_2)\Big]\Big[D^{-1}V(0, 0)^{-1}\Big]$$

is unimodular too.

Before concluding this section, we aim to mention the problem of finding, and realizing through a finite memory system, a syndrome decoder of a given encoder G, $\ell ZP in \mathbf{F}[z_1, z_2]$. In the general case, this requires to construct a $\ell ZP polynomial matrix H^T(z_1, z_2)$, of suitable dimensions, such that $GH^T = 0$, and to implement a finite memory realization of it.
When the encoder G fulfills the equivalent requirements of Proposition 5.2, the problem becomes considerably simpler. Actually, let $G^{-1}(z_1, z_2)$ denote the decoder realized by the inverse system $\Sigma^{-1}(D^{-1})$ and $[\, D^{-1} \quad L \,]$ the inverse matrix of $\begin{bmatrix} D \\ K \end{bmatrix}$, K as in (iii) of Proposition 5.2. It's quite easy to prove that the polynomial matrix

$$H^T(z_1, z_2) := \Big[I \; - \; G^{-1}(z_1, z_2)G(z_1, z_2)\Big] \; L \tag{5.6}$$

is a syndrome decoder of $\mathcal{C}(G)$. Moreover, the block scheme of FIG.4

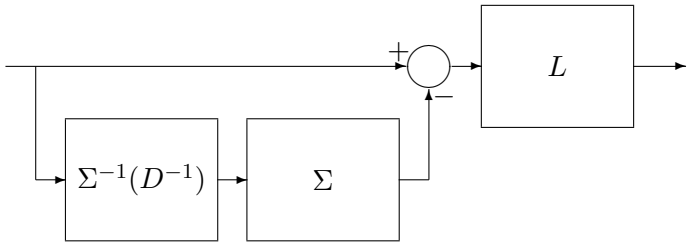


FIG.4

immediately suggests how to obtain from Σ and $\Sigma^{-1}(D^{-1})$ the following finite memory realization of H^T:

$$\Sigma_{H^T} = \Big(\begin{bmatrix} A_1 - CD^{-1}B_1 & -CD^{-1}B_1 \\ 0 & A_1 \end{bmatrix}, \begin{bmatrix} A_2 - CD^{-1}B_2 & -CD^{-1}B_2 \\ 0 & A_2 \end{bmatrix},$$

$$[\, -D^{-1}B_1 \quad -D^{-1}B_1 \,], [\, -D^{-1}B_2 \quad -D^{-1}B_2 \,], \begin{bmatrix} CD^{-1}DH \\ -CH \end{bmatrix}, \Big[(I - D^{-1}D)H \Big] \Big).$$

6 References

1. G.D.Forney, "Convolutional codes I: algebraic structure", IEEE Trans. Inf. Th., vol.16, pp.720-738, 1970

2. B.C.Lévy, "2D polynomial and rational matrices, and their applications for the modeling of 2D dynamical systems", Tech.Rep. No.M735-11, Stanford Electronics Laboratories, Stanford University, 1981

3. D.C.Youla, P.F.Pickel, "The Quillen-Suslin theorem", IEEE Trans. Circ. and Sys., vol.31, pp.513-518, 1984

4. D.C.Youla, G.Gnavi, "Notes on n-dimensional system theory", IEEE Trans. Circ. and Sys., vol.26, pp.105-111, 1979

5. M.Morf, B.C.Lévy, S.Y.Kung, T.Kailath "New results in 2D systems theory: Part I and II", Proc. of the IEEE, vol.65, pp.861-872 and 945-961, 1977

6. J.C.Willems, "Models for dynamics", Dynamics Reported, vol.2, pp.171-267, 1989

7. P.Rocha, "Structure and representation of 2D systems", Ph.D.Thesis, Rijksuniversiteit Groningen, 1990

8. P.Rocha, J.C.Willems, "Controllability of 2D systems", IEEE Trans. Aut. Contr., vol.36, pp.413-423, 1991

9. E.Fornasini, S.Zampieri,"A note on the state space realization of 2D FIR transfer functions", Sys. Contr. Letters, vol.16, pp.17-22, 1990

10. M.Bisiacco, "State and output feedback stabilizability of 2D systems", IEEE Trans. Circ. and Sys., vol.32, pp.1246-54, 1985

11. E.Fornasini, G.Marchesini, "Properties of pairs of matrices and state models for 2D systems; pt.I and pt.II", Proc. 7th Int. Conf. on Multivariate Analysis, Penn State Univ., May 1992

12. Ph.Piret, "Convolutional codes", The MIT Press, 1991

13. S.Lang, "Algebra", Addison Wesley Pubbl. Comp., 1967

14. E.Fornasini, P.Rocha, S.Zampieri, "State space realization of 2D finite dimensional behaviours", to appear on SIAM Journal of Control

15. E.Fornasini, M.E.Valcher, "An algebraic approach to 2D convolutional codes. Pt. I", Tech.Rep., 1992

16. W.Greub, "Linear algebra", Springer - Verlag, 1975

17. E.Fornasini, G.Marchesini, "State space realization of two-dimensional filters", IEEE Trans. on Aut. Contr., vol.21, pp.484-492, 1976

18. M.Bisiacco, E.Fornasini, Marchesini, "Dynamic regulation of 2D systems: a state-space approach", Lin. Alg. Appl., vol.122/124, pp.195-218, 1989

19. E.Fornasini, M.E.Valcher "Polynomial inverses of 2D transfer matrices and finite memory realizations via inverse systems", to appear on Multidim. Sys. and Sign. Proc.

20. E.Fornasini, M.E.Valcher, "On the structure of finite memory and separable 2D systems", Proc. 2nd IFAC Workshop on System Structure and Control, Prague Sept.1992, pp.400-403

7 Appendix

Lemma A.1 *(i) Let* $G(z_1, z_2)$ *be a* $k \times n$ $\ell FP matrix with elements in F_{\pm}$. *If* $\mathbf{v} \in {\mathcal{F}_{\pm}}^{1\times n}$ *is a linear combination over* $\mathbf{F}(z_1, z_2)$ *of the rows of* $G(z_1, z_2)$, *i.e.*

$$\mathbf{v} = \mathbf{a}G, \quad \mathbf{a} \in \mathbf{F}(z_1, z_2)^{1\times k}, \tag{A.1}$$

then $\mathbf{a}$ *can be chosen in* ${\mathcal{F}_{\pm}}^{1\times k}$.

(ii) The same statement holds when $\mathcal{F}_{\pm}$ *is replaced by* $\mathbf{F}[z_1, z_2]$.

PROOF (**i**) *Since* G *is* $\ell FP, there exist [5] two polynomials* $h(z_1) \in \mathbf{F}[z_1, z_1^{-1}]$ *and* $k(z_2) \in \mathbf{F}[z_2, z_2^{-1}]$, *and two L-polynomial matrices* $X(z_1, z_2)$ *and* $Y(z_1, z_2)$, *such that*

$$GX = h(z_1)I_k \quad \text{and} \quad GY = k(z_2)I_k. \tag{A2}$$

It entails no loss of generality supposing that the row vector $\mathbf{a}$ *has irreducible entries,* a_i*. So, letting* β_0 *the l.c.m. of the denominators of* a_i*, (A1) can be rewritten as*

$$\beta_0 \mathbf{v} = [\beta_1 \ ... \ \beta_k]G \quad \beta_i \in \mathcal{F}_{\pm}, \ i = 1, 2, ... \tag{A3}$$

Postmultiplying both members of (A3) by $X(z_1, z_2)$ *and* $Y(z_1, z_2)$*, we obtain*

$$\begin{aligned} \beta_0 \mathbf{v} X &= [\beta_1 \ ... \ \beta_k]GX = [\beta_1 \ ... \ \beta_k]h(z_1) \\ \beta_0 \mathbf{v} Y &= [\beta_1 \ ... \ \beta_k]GY = [\beta_1 \ ... \ \beta_k]k(z_2). \end{aligned}$$

As $\beta_0, \beta_1, ..., \beta_k$ *have no common factors, it follows that* $\beta_0(z_1, z_2) \mid h(z_1)$ *and* $\beta_0(z_1, z_2) \mid k(z_2)$*, and therefore* $\beta_0(z_1, z_2)$ *is a unit in* $\mathcal{F}_{\pm}$*.*

(**ii**) *Obvious* ■

Corollary A.2 *Let* $G(z_1, z_2)$ *be in* $\mathcal{F}_{\pm}^{k \times n}$*, with row rank* $\bar{k}$ *over* $\mathbf{F}(z_1, z_2)$*. There exist two L-polynomial matrices,* $\bar{G}(z_1, z_2)$*,* $\bar{k} \times n$ ℓFP*, and* $T(z_1, z_2)$*,* $k \times \bar{k}$ *with full column rank, such that*

$$G(z_1, z_2) = T(z_1, z_2)\bar{G}(z_1, z_2). \tag{A4}$$

PROOF *Let* $G'(z_1, z_2)$ *be a matrix obtained by selecting in* $G(z_1, z_2)$ $\bar{k}$ *rows linearly independent over* $\mathbf{F}(z_1, z_2)$*, and* $Q(z_1, z_2)$ *a g.l.f. of* $G'(z_1, z_2)$*. Then* $G' = Q\bar{G}$*. Every row in* G *is a linear combination over* $\mathbf{F}(z_1, z_2)$ *of the rows of* $\bar{G}(z_1, z_2)$*, and, by Lemma A.1, the coefficients of the combination can be chosen in* $\mathcal{F}_{\pm}$*.*
Therefore $G = T\bar{G}$*,* $T(z_1, z_2)$ *being the* $k \times \bar{k}$ *matrix of the combinators. As rank* G *is* $\bar{k}$*, rank* T *cannot be less than* $\bar{k}$ ■

Proposition A.3 *Let* $G(z_1, z_2)$ *be in* $\mathcal{F}_{\pm}^{k \times n}$*,* $k \leq n$*.* G *has an L-polynomial right inverse if and only if* G *is* ℓZP*.*

PROOF *If* G *is* ℓZP*, the ideal generated by its maximal order minors,* $m_i(G)$*,* $i = 1, 2, ...$ $\binom{n}{k}$*, coincides with the whole ring* $\mathcal{F}_{\pm}$*. It follows that there exists* $\alpha_i \in \mathcal{F}_{\pm}$*,* $i = 1, 2, ...$ $\binom{n}{k}$*, such that* $\sum_i \alpha_i m_i(G) = 1$*.*
Consider the identity, $m_i(G) I_k = G S_i \mathrm{adj}(G S_i)$*, where* S_i *denotes the selection matrix corresponding to the minor* $m_i(G)$*. Then we have*

$$I_k = \sum_i \alpha_i m_i(G) I_k = G(z_1, z_2)[\sum_i \alpha_i(z_1, z_2) S_i \mathrm{adj}(G(z_1, z_2) S_i].$$

Clearly, the L-polynomial matrix $K(z_1, z_2) := \sum_i \alpha_i(z_1, z_2) S_i \mathrm{adj}(G(z_1, z_2) S_i)$ *is a right inverse of* G*.*

The converse is a direct consequence of the Binet-Cauchy formula ■

Proposition A.4 *Let* $G(z_1, z_2) \in \mathcal{F}_{\pm}^{k \times n}$ *and* $H^T(z_1, z_2) \in \mathcal{F}_{\pm}^{n \times (n-k)}$ *be* ℓFP *and* rFP *matrices, respectively,* 0*.*
The corresponding maximal order minors of G *and* H^T *are equal, modulo a unit of the ring* $\mathcal{F}_{\pm}$*.*

PROOF *Since G is ℓFP, there exist [5] two matrices, X_1 and X_2, with elements in $\mathcal{F}_\pm$, and two polynomials, $g_1(z_1) \in \mathbf{F}[z_1, z_1^{-1}]$ and $g_2(z_2) \in \mathbf{F}[z_2, z_2^{-1}]$, such that*

$$G_1 X_1 = g_1(z_1) I_k \quad \text{and} \quad G_2 X_2 = g_2(z_2) I_k.$$

Consider for instance $m_1(G)$, the maximal order minor of G corresponding to the selection of the first k columns of G. Complete G into a square matrix by resorting to a $(n-k) \times n$ matrix, whose columns are all zero except for the last $n-k$, which consitute the identity matrix. Thus

$$\begin{bmatrix} G & \\ 0 & I_{n-k} \end{bmatrix} [\, X_1 \mid H^T \,] = \begin{bmatrix} g_1(z_1) I_k & 0 \\ Q & M_1(H^T) \end{bmatrix}, \tag{A5}$$

where $M_1(H^T)$ is the $(n-k) \times (n-k)$ submatrix of H^T obtained by selecting the last $n-k$ rows. Assuming $R_1(z_1, z_2) := [\, X_1 \;\; H^T \,]$ and $\mu_1 := \det M_1(H^T)$, we get

$$m_1(G) \det R_1 = \Big(g_1(z_1)\Big)^k \mu_1(H^T).$$

Now replace X_2 with X_1 in (A5) and let $R_2 := [\, X_2 \;\; H^T \,]$. We obtain

$$m_1(G) \det R_2 = \Big(g_2(z_2)\Big)^k \mu_1(H^T).$$

So $m_1(G) \mid \Big(g_1(z_1)\Big)^k \mu_1(H^T)$ and $m_1(G) \mid \Big(g_2(z_2)\Big)^k \mu_1(H^T)$. Since $\Big(g_1(z_1)\Big)^k$ and $\Big(g_2(z_2)\Big)^k$ are coprime, then

$$m_1(G) \mid \mu_1(H^T). \tag{A6}$$

Dually, as H^T is rFP, there exist two matrices, Y_1 and Y_2, with elements in $\mathcal{F}_\pm$, and two polynomials, $h_1(z_1) \in \mathbf{F}[z_1, z_1^{-1}]$ and $h_2(z_2) \in \mathbf{F}[z_2, z_2^{-1}]$, such that

$$Y_1 H^T = h_1(z_1) I_{n-k} \quad \text{and} \quad Y_2 H^T = h_2(z_2) I_{n-k}.$$

We can proceed as before, getting

$$\begin{bmatrix} G \\ Y_1 \end{bmatrix} \begin{bmatrix} I_k & \\ & H^T \\ 0 & \end{bmatrix} = \begin{bmatrix} M_1(G) & 0 \\ T & h_1(z_1) I_{n-k} \end{bmatrix},$$

where $M_1(G)$ is the $k \times k$ submatrix of G obtained by selecting its first k columns. Assuming $S_1 := \begin{bmatrix} G \\ Y_1 \end{bmatrix}$, we get $\det S_1 \;\; \mu_1(H^T) = m_1(G)\Big(h_1(z_1)\Big)^{n-k}$, and, analogously, $\det S_2 \;\; \mu_1(H^T) = m_1(G)\Big(h_2(z_2)\Big)^{n-k}$, where $S_2 := \begin{bmatrix} G \\ Y_2 \end{bmatrix}$.
Therefore $\mu_1(H^T)$ is a common factor of $m_1(G)\Big(h_1(z_1)\Big)^{n-k}$ and $m_1(G)\Big(h_2(z_2)\Big)^{n-k}$, and then

$$\mu_1(H^T) \mid m_1(G). \tag{A7}$$

(A6) and (A7) together imply that $m_1(G)$ and $\mu_1(H^T)$ differ in a unit of $\mathcal{F}_\pm$. Similarly we can show that the same result holds for any other pair of corresponding minors in G and in H^T ■

Lemma A.5 *Let $H^T(z_1, z_2) \in \mathcal{F}_\pm^{n\times(n-k)}$. The map $H^T : \mathcal{F}_\pm^n \to \mathcal{F}_\pm^{n-k} \ : \ \mathbf{w} \mapsto \mathbf{w}H^T$ is onto if and only if H^T is rZP.*

PROOF *Assume that H^T is onto. Then there exist $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{n-k} \in \mathcal{F}_\pm^n$ such that $\mathbf{w}_i H^T = \mathbf{e}_i = [\ 0...0\ 1\ 0...0\ ]$. Letting $W = \mathrm{col}\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{n-k}\}$, we have $WH^T = I_{n-k}$. So, by Proposition A3, $H^T(z_1, z_2)$ is rZP.*

Conversely, if H^T is rZP, it admits a left inverse $W \in \mathcal{F}_\pm^{(n-k)\times n}$. So, for every $\mathbf{p} \in \mathcal{F}_\pm^{n-k}$ we have $\mathbf{p} = \mathbf{p}I_{n-k} = (\mathbf{p}W)H^T$, which implies that $\mathbf{p}$ is the image, under H^T, of an L-polynomial vector ■

Lemma A.6 *Suppose that $\bar{G}_1(z_1, z_2)$ and $\bar{G}_2(z_1, z_2)$ are $k \times n$ polynomial matrices, $\ell FPin\mathbf{F}[z_1, z_2]$. If there exists an $\mathcal{F}_\pm-$unimodular matrix $U(z_1, z_2)$ such that*

$$\bar{G}_1(z_1, z_2) = U(z_1, z_2)\bar{G}_2(z_1, z_2) \tag{A8}$$

then

$$\bar{G}_1(z_1, z_2) = V(z_1, z_2)\bar{G}_2(z_1, z_2) \tag{A9}$$

for some $\mathbf{F}[z_1, z_2]-$unimodular matrix V.

PROOF *By assumption (A8), every row of $\bar{G}_1$ is a linear combination over $\mathbf{F}(z_1, z_2)$ of the rows of $\bar{G}_2$. By Lemma A.1, there exists a matrix $V(z_1, z_2) \in \mathbf{F}[z_1, z_2]^{k\times k}$ such that $\bar{G}_1 = V\bar{G}_2$. Since $\bar{G}_1$ is ℓFP,$\det V$ is a nonzero constant and V is $\mathbf{F}[z_1, z_2]$ − unimodular* ■