Systems theoretic methods in decoding

Margreta Kuijper
The University of Melbourne
Dep. Electr. Electronic Eng.
VIC 3010
Australia
margreet@ee.unimelb.edu.au

Jan Willem Polderman
University of Twente
Fac. EEMCS, Math. Syst. Contr. Th.
P.O. Box 217, 7500 AE Enschede
The Netherlands
j.w.polderman@ewi.utwente.nl

Abstract

In this paper we show how ideas based on system theoretic modeling of linear behaviors may be used for decoding of linear codes. In particular we show how Sudan's bivariate interpolation approach to list decoding of RS codes allows a system theoretic interpretation.

1 Introduction

The purpose of this paper is to illustrate the potential of the system theoretic framework in decoding of linear codes. There has been a growing interest to establish relationships between the area of coding theory and the area of system theory, particularly the behavioral approach. In [8] it is shown how the theory on behavioral modeling leads to a transparent interpretation of various existing decoding methods as well as to the derivation of an insightful decoding algorithm. In particular, the Berlekamp-Massey algorithm is interpreted as behavioral modeling for single-input-single-output partial realization, a result that was first presented in [14]. A multivariable version of this algorithm is derived in [3, 15] and put to work in [6, 4, 5] to achieve increased error correction of the closely related BCH codes. In more recent work, it is shown how the Welch-Berlekamp algorithm [8, 7] can be interpreted as a special instance of behavioral modeling for single-input-single-output interpolation. This work has been extended to errors-and-erasures decoding in [13].

In this paper we place the list decoding approach of [18, 19] in a behavioral framework, see also [10]. We find that we are able to interpret this approach as behavioral modeling for multivariable interpolation. This enables us to generate insightful decoding algorithms in a straightforward way. In this paper we restrict ourselves to the early results of [18, 19], that is, interpolation with multiplicity one. The exposition is largely based on [11]. For more details and proofs we refer to [11, 12]. Our approach can also deal with the more general case where the data points are to be interpolated with multiplicities larger than one. See [12]. Finally, we are currently developing a similar framework for codes over finite rings, see [9].

A crucial difference between the area of coding theory and the area of system theory is the fact that the alphabets used in coding theory are finite whereas the alphabets used in system theory are usually infinite, typically the real/complex numbers. In this paper we address the implications of finite fields for behavioral modeling.

2 Behaviors over finite fields

In this section we review some basic concepts and results of the behavioral approach to a linear system over a field $\mathbb{F}$. For most of these it makes no difference whether $\mathbb{F}$ is

finite or infinite. The only exception pertains to the differentiation operator as we will see below. The reader is referred to [17] for more detailed discussions.

Following [17] a dynamical system is a triple $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$. Here $\mathbb{T}$ can be thought of as the time axis, $\mathbb{W}$ is the signal alphabet, and $\mathfrak{B}$, the behavior of the system, is a subset of $\mathbb{W}^{\mathbb{T}}$ (the set of functions from $\mathbb{T}$ to $\mathbb{W}$). Relevant choices for our purposes are $\mathbb{T} = \mathbb{Z}_+$, $\mathbb{W} = \mathbb{F}^q$, and $\mathfrak{B}$ a linear subspace of $\mathbb{W}^{\mathbb{T}}$.

We define σ, the shift operator, acting on elements in $\mathbb{W}^{\mathbb{T}}$ as $(\sigma w)(k) = w(k+1)$. An important class of systems are those whose behaviors are defined as the kernel of a polynomial matrix in σ. Let $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ be a $g \times q$ matrix in the indeterminate ξ and with coefficients in $\mathbb{F}$. Then we define the behavior corresponding to $R(\xi)$ as

$$\mathfrak{B} = \{\mathbf{w} : \mathbb{Z}_+ \to \mathbb{F}^q \mid R(\sigma)\mathbf{w} = 0\}. \tag{1}$$

It is easy to see that $\mathfrak{B}$ is linear. Moreover, $\mathfrak{B}$ is time-invariant, that is, for every trajectory $\mathbf{w}$ in $\mathfrak{B}$ the shifted trajectory $\sigma\mathbf{w}$ is also in $\mathfrak{B}$. The class of behaviors in q variables that admit a representation of the form $R(\sigma)\mathbf{w} = 0$ is denoted by $\mathfrak{L}^q$. Representations of the form $R(\sigma)\mathbf{w} = 0$ are, for obvious reasons, referred to as *kernel representations*. In the general theory of behaviors many other representations are of interest. For our purposes only kernel representations are of interest.

It appears that different matrices $R_1(\xi)$ and $R_2(\xi)$ may define the same behavior. The following result classifies the set of matrices that define a given behavior $\mathfrak{B}$.

Lemma 2.1. *For $i = 1, 2$ let $R_i(\xi) \in \mathbb{F}^{g_i\times q}[\xi]$ and denote the corresponding behaviors by $\mathfrak{B}_i$. If $\mathfrak{B}_1 \subset \mathfrak{B}_2$, then there exists a matrix $F(\xi) \in \mathbb{F}^{g_2\times g_1}[\xi]$ such that $R_2(\xi) = F(\xi)R_1(\xi)$.*

A matrix $U(\xi) \in \mathbb{F}^{g\times g}[\xi]$ is said to be *unimodular* if there exists $V(\xi) \in \mathbb{F}^{g\times g}[\xi]$ such that $U(\xi)V(\xi) = V(\xi)U(\xi) = I$, equivalently, if $\det U(\xi)$ is a nonzero constant in $\mathbb{F}$. A direct consequence of the above lemma is the following.

Theorem 2.2. *Let $R_i(\xi) \in \mathbb{F}^{g\times q}[\xi]$ define the same behavior ($i = 1, 2$), i.e., $R_1(\sigma)\mathbf{w} = 0$ if and only if $R_2(\sigma)\mathbf{w} = 0$. Then there exists a unimodular matrix $U(\xi) \in \mathbb{F}^{g\times g}[\xi]$ such that $R_2(\xi) = U(\xi)R_1(\xi)$.*

Theorem 2.2 makes it possible to choose out of the many representations of a given behavior one that is particularly convenient for the application at hand. Examples are upper or lower triangular forms. Also, by means of appropriate unimodular premultiplication one may create zero rows to end up with a matrix in which the remaining nonzero rows are independent over $\mathbb{F}[\xi]$. The nonzero rows form a matrix with fewer rows and is said to be of *full row rank*. A form that is crucial in the application of the behavioral approach to coding theory is the *row reduced form*.

Definition 2.3. Let $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ and denote the rows of $R(\xi)$ by $r_i(\xi)$, $i = 1, \ldots, g$. The *row degrees* $d_1, \ldots, d_g$ are defined as $d_i = \max_{j=1,\ldots,q} \deg r_{ij}(\xi)$. Define the diagonal matrix $D(\xi) = \operatorname{diag}(\xi^{d_1}, \ldots, \xi^{d_g})$ and write $R(\xi) = D(\xi)R_0 + R_1(\xi)$ with $D(\xi)^{-1}R_1(\xi)$ strictly proper, meaning that in every entry of $D(\xi)^{-1}R_1(\xi)$ the degree of the denominator exceeds the degree of the numerator. Then, $R(\xi)$ is said to be *row reduced* if R_0 is of full row rank as a matrix over $\mathbb{F}^{g\times q}$. The matrix R_0 is called the *leading row coefficient matrix*.

The next two theorems are well-known results from behavioral theory.

Theorem 2.4. *Let $R(\xi) \in \mathbb{F}^{q\times q}[\xi]$ be a square matrix with row degrees $d_1, \ldots, d_q$. Denote the sum of these row degrees by d. Then $R(\xi)$ is row reduced if and only if* $\deg \det R(\xi) = d$.

Theorem 2.5. *Let $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ be of full row rank. There exists a unimodular matrix $U(\xi)$ such that $U(\xi)R(\xi)$ is row reduced.*

In the sequel we use a modified version of row reducedness of which the above is a special case. This is the notion of "weighted row reduced".

Definition 2.6. Let $n_1, \ldots, n_q$ be nonnegative integers. Define

$$N(\xi) = \operatorname{diag}(\xi^{n_1}, \ldots, \xi^{n_q}). \tag{2}$$

The *weighted row degrees* of $R(\xi)$ are defined as the row degrees of $R(\xi)N(\xi)$. The matrix $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ is called $(n_1, \ldots, n_q)$ *weighted row reduced* if $R(\xi)N(\xi)$ is row reduced.

The following two theorems are generalizations of Theorem 2.4 and Theorem 2.5.

Theorem 2.7. *Let $R(\xi) \in \mathbb{F}^{q\times q}[\xi]$ be a square polynomial matrix of full row rank and let $n_1, \ldots, n_q$ be nonnegative integers. Let $N(\xi)$ be defined as in (2). Then $R(\xi)$ is $(n_1, \ldots, n_q)$ weighted row reduced if and only if $\deg\det R(\xi) + \deg\det N(\xi)$ equals the sum of the weighted row degrees of $R(\xi)$.*

Theorem 2.8. *Let $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ be of full row rank and let $n_1, \ldots, n_q$ be nonnegative integers. There exists a unimodular matrix $U(\xi)$ such that $U(\xi)R(\xi)$ is $(n_1, \ldots, n_q)$ weighted row reduced.*

Notice that $(0, \ldots, 0)$ weighted row reduced is just row reduced. We mainly consider $(0, \kappa-1, 2(\kappa-1), \ldots, (q-1)(\kappa-1))$ weighted row reduced. We shall refer to this special case as simply weighted row reduced whenever there is little danger of confusion.

The next two results show that row reducedness indicates minimality. This observation turns out to be crucial in the behavioral interpretation of the decoding scheme of [18].

Lemma 2.9. *Let $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ be row reduced and let $m \in \mathbb{Z}_+$ be the minimal row degree of $R(\xi)$. Then every linear combination over $\mathbb{F}[\xi]$ of the rows of $R(\xi)$ has row degree at least m.*

Corollary 2.10. *Let $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ be weighted row reduced and let $m \in \mathbb{Z}_+$ be the minimal weighted row degree of $R(\xi)$. Then every linear combination over $\mathbb{F}[\xi]$ of the rows of $R(\xi)$ has weighted row degree at least m.*

As remarked, behaviors are represented by polynomial matrices. The question arises how, for a given polynomial matrix, the behavior can be determined explicitly. Let us now continue to determine an explicit expression for a behavior over a finite field in terms of its polynomial representation. Our key players are trajectories $\boldsymbol{w}_i : \mathbb{Z}_+ \to \mathbb{F}$ defined by

$$w_i(k) := \lambda_i^k$$

where $\lambda_i \in \mathbb{F}$.

Theorem 2.11. *Let $R(\xi) \in \mathbb{F}^{q\times q}[\xi]$, let $\det R(\xi)$ be a polynomial of degree n, and let $\mathfrak{B} = \{\mathbf{w} : \mathbb{Z}_+ \to \mathbb{F} \mid R(\sigma)\mathbf{w} = 0\}$. Then $\mathfrak{B}$ is an n-dimensional subspace of $(\mathbb{F}^q)^{\mathbb{Z}_+}$. If the roots of $\det R(\xi)$ are mutually distinct and belong to $\mathbb{F}$, say $\det R(\xi) = \prod_{i=1}^{n}(\xi-\lambda_i)$, with $\lambda_i \in \mathbb{F}$, then all trajectories in $\mathfrak{B}$ are of the form*

$$w(k) = \sum_{i=1}^{n} b_i \lambda_i^k,$$

with $b_i \in \mathbb{F}^q$ such that $R(\lambda_i)b_i = 0$.

In the above we investigated explicit expressions for trajectories satisfying a given polynomial representation. In the sequel we are interested in the converse, namely building representations from given trajectories. For this purpose we use the theory of exact modeling of behaviors as first introduced in [20]. We here recall a few of the main ideas. Given a finite number of trajectories $\mathbf{w}_j : \mathbb{Z}_+ \to \mathbb{F}^q$ $(j = 1, \ldots, N)$ we may seek to build a system whose behavior contains these specific trajectories. A behavior $\mathcal{B}$ is called an *unfalsified model* for the *data set* $\mathbf{D} = \{\boldsymbol{w_1}, \ldots, \boldsymbol{w_N}\}$ if $\mathbf{D} \subseteq \mathcal{B}$. A model $\mathcal{B}_1$ is called *more powerful* than a model $\mathcal{B}_2$ if $\mathcal{B}_1 \subseteq \mathcal{B}_2$. From a modeling perspective it appears sensible to look for the smallest behavior that contains the N trajectories. A model $\mathcal{B}^*$ is called the *most powerful unfalsified model (MPUM)* for $\mathbf{D}$, if $\mathcal{B}^*$ is unfalsified for $\mathbf{D}$ and $\mathbf{D} \subseteq \mathcal{B} \implies \mathcal{B}^* \subseteq \mathcal{B}$. In [21] it is shown that for $\mathbf{D} = \{\boldsymbol{w_1}, \ldots, \boldsymbol{w_N}\}$ such an MPUM exists. In fact, a general procedure for the iterative construction of a kernel representation for $\mathcal{B}^*$ is presented. We here recall this procedure; its workings can be easily understood from Lemma 2.1.

Procedure 2.12. *([21]) Initially define*

$$R_0(\xi) := I \text{ (where } I \text{ is the identity matrix).}$$

Proceed iteratively as follows for $k = 1, \ldots, N$. Define, after receiving $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$, the k-th error trajectory $\boldsymbol{e}_k$ *as*

$$\boldsymbol{e}_k := R_{k-1}(\sigma)\boldsymbol{w}_k.$$

Compute a kernel representation $V_k(\sigma)\boldsymbol{w} = 0$ of the MPUM for $\{\boldsymbol{e}_k\}$. Then define

$$R_k(\xi) := V_k(\xi)R_{k-1}(\xi).$$

Theorem 2.13. *([21]) For $k = 1, \ldots, N$, the kernel representation $R_k(\sigma)\boldsymbol{w} = 0$ of the above procedure, represents the MPUM for $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$.*

Remark 2.14. For general trajectories Procedure 2.12 may be cumbersome to run. However, for exponential trajectories the procedure is easy and convenient to perform. The trajectories to which we apply Procedure 2.12 are exponential as we shall see in Section 4.

3 Behavioral interpolation for decoding

Let $\{\xi_1, \cdots, \xi_n\}$ be a subset of a finite field $\mathbf{F}$ with the ξ_is mutually distinct. An (n, κ) RS code is defined as a set of codewords of the form $\boldsymbol{c} = (m(\xi_1), \ldots, m(\xi_n))$, with $m(\xi) \in \mathbb{F}[\xi]$ running through the set of polynomials of degree $< \kappa$. The codeword $\mathbf{c}$ is transmitted through a channel where errors may occur so that the received word $\mathbf{r}$ is not necessarily equal to the transmitted codeword $\mathbf{c}$. The decoding problem consists of reconstructing the original polynomial $m(\xi)$ from the received word $\mathbf{r}$. In list decoding a list of possible polynomials $m(\xi)$ is generated. The breakthrough idea of [18] is to use bivariate polynomials for list decoding.

Definition 3.1. Let $Q(\xi, \eta) = \sum_{i \in I, j \in J} q_{ij} \xi^i \eta^j \in \mathbb{F}[\xi, \eta]$ be a bivariate polynomial. The (w_ξ, w_η) weighted degree of $Q(\xi, \eta)$ is defined as

$$\operatorname{wdeg}_{(w_\xi, w_\eta)} Q(\xi, \eta) = \max_{i \in I, j \in J} \{i w_\xi + j w_\eta \mid q_{ij} \neq 0\}.$$

Lemma 3.2. *Let (ξ_i, η_i) $(i = 1, \ldots, n)$ be elements of $\mathbb{F}^2$ with the ξ_is mutually distinct. Let $Q(\xi, \eta) \in \mathbb{F}[\xi, \eta]$ be a bivariate polynomial with $\operatorname{wdeg}_{(1, \kappa-1)} Q(\xi, \eta) = \ell$ such that $Q(\xi_i, \eta_i) = 0$ for $i = 1 \ldots, n$. Let $\tilde{m}(\xi)$ be a polynomial of degree $< \kappa$ such that $\#\{i \mid \tilde{m}(\xi_i) = \eta_i\} \geq \ell + 1$. Then $\eta - \tilde{m}(\xi)$ divides $Q(\xi, \eta)$.*

Proof. The univariate polynomial $Q(\xi, \tilde{m}(\xi))$ clearly has at least $\ell + 1$ zeros. On the other hand, $\deg Q(\xi, \tilde{m}(\xi))$ cannot exceed ℓ since by assumption $\mathrm{wdeg}_{(1,\kappa-1)} Q(\xi, \eta) = \ell$. Since a polynomial of degree not exceeding ℓ can only have more than ℓ zeros if it is the zero polynomial, it follows that $Q(\xi, \tilde{m}(\xi))$ is the zero polynomial. This now implies that $\eta - \tilde{m}(\xi)$ divides $Q(\xi, \eta)$. □

In the sequel we are only concerned with the $(1, \kappa-1)$ weighted degree and therefore we refer to it as just the weighted degree. The next corollary expresses the main idea of the Sudan list decoding approach.

Corollary 3.3. *Let $Q(\xi, \eta) \in \mathbb{F}[\xi, \eta]$ be a bivariate polynomial of weighted degree ℓ such that $Q(\xi_i, \eta_i) = 0$ for $i = 1 \ldots, n$. Let $\boldsymbol{r} = (\eta_1, \ldots, \eta_n)$ be a received word. Denote the corresponding transmitted message polynomial by $m(\xi)$. If $\boldsymbol{r}$ contains less than $n - \ell$ errors then $\eta - m(\xi)$ divides $Q(\xi, \eta)$.*

The main idea of Sudan's list decoding approach is to construct a polynomial $Q(\xi, \eta)$ such that $Q(\xi_i, \eta_i) = 0$. It makes sense to minimize the weighted degree of this polynomial as this maximizes the number of errors that can be corrected that way. In the decoding process all factors of the form $\eta - \tilde{m}(\xi)$ are subsequently extracted to produce a list of candidate polynomials $\tilde{m}(\xi)$ of degree $< \kappa$. The next step is then to produce a sublist of most likely message words by computing the corresponding codewords and comparing with $\boldsymbol{r}$. It has been shown in the literature [16] that in most cases this sublist consists of only one polynomial. This is due to the geometric structure of the code.

Roughly, our approach is structured as follows. We write the polynomial $Q(\xi, \eta)$ to be constructed as $Q(\xi, \eta) = \sum_{j=0}^{M} d_j(\xi)\eta^j$ for an appropriate choice of M. With the n data points (ξ_i, η_i) $(i = 1, \ldots, n)$ we associate n trajectories $\mathbf{w}_i : \mathbb{Z}_+ \to \mathbb{F}^{M+1}$. We then determine the Most Powerful Unfalsified Model $\mathfrak{B}$ of these trajectories. Then we construct a *weighted row reduced* matrix $R(\xi)$ that represents $\mathfrak{B}$ (the notion of "weighted row reduced" is defined in Section 2). From $R(\xi)$ we select a row $d(\xi)$ of minimal weighted row degree and finally we define $Q(\xi, \eta) = \sum_{j=0}^{M} d_j(\xi)\eta^j$, where the $d_i(\xi)$s are the entries of $d(\xi)$. It turns out that $Q(\xi, \eta)$ constructed in this way is a bivariate polynomial of minimal weighted degree that interpolates the data points (ξ_i, η_i) for $i = 1, \ldots, n$.

4 Minimal interpolation as behavioral modeling

In this section we reformulate the problem statement as introduced in Section 1 in terms of behavioral modeling. It turns out that we can apply the behavioral theory in a straightforward way as follows. First we write $Q(\xi, \eta)$ as $Q(\xi, \eta) = \sum_{j=0}^{M} d_j(\xi)\eta^j$. The upper limit M has to be chosen with care, for too small an M may result in a $Q(\xi, \eta)$ that is not of minimal weighted degree. We comment on the choice of M later. What we are aiming at is $Q(\xi_i, \eta_i) = 0$, i.e., $\sum_{j=0}^{M} d_j(\xi_i)\eta_i^j = 0$. Another way of stating this is that we are looking for a polynomial vector $d(\xi) = [d_0(\xi), \ldots, d_M(\xi)]$ such that

$$\begin{bmatrix} d_0(\xi_i) & \cdots & d_M(\xi_i) \end{bmatrix} \begin{bmatrix} 1 \\ \eta_i \\ \vdots \\ \eta_i^M \end{bmatrix} = 0 \text{ for } i = 1, \ldots, n.$$

In the light of Theorem 2.11 this is equivalent to

$$\underbrace{\begin{bmatrix} d_0(\sigma) & \cdots & d_M(\sigma) \end{bmatrix}}_{d(\sigma)} \underbrace{\left(\begin{bmatrix} 1 \\ \eta_i \\ \vdots \\ \eta_i^M \end{bmatrix} \xi_i^k \right)}_{w_i(k)} = 0 \quad \text{for } i = 1, \ldots, n. \tag{3}$$

Apparently the aim is to find an integer M and a polynomial vector $d(\xi) \in \mathbb{F}^{1\times(M+1)}[\xi]$ of minimal weighted degree such that $d(\sigma)w_i = 0$ for $i = 1, \ldots, n$. Of course a trivial solution is $Q(\xi, \eta) = \prod_{j=1}^{n}(\xi - \xi_i)$. This solution has weighted degree n and may serve as an upperbound on the minimal weighted degree. As a consequence we can take

$$M = \max\{j \in \mathbb{N} \mid j \leq \frac{n}{\kappa - 1}\}. \tag{4}$$

The idea is now to find a representation $\tilde{R}(\xi)$ of the MPUM of $w_1, \ldots, w_n$ that is weighted row reduced. It then turns out that for $d(\xi)$ we can take a row of $R(\xi)$ of minimal weighted degree. We explain this in more detail below.

Theorem 4.1. *Let $\mathfrak{B}$ be the MPUM of $w_1, \ldots, w_n$ defined in (3) with M defined by (4). Let $R(\xi) \in \mathbb{F}^{(M+1)\times(M+1)}[\xi]$ be a weighted row reduced representation of $\mathfrak{B}$ and let $d(\xi) = \begin{bmatrix} d_0(\xi) & \cdots & d_M(\xi) \end{bmatrix}$ be a row of $R(\xi)$ of minimal weighted degree. Define $Q(\xi, \eta) = \sum_{j=0}^{M} d_j(\xi)\eta^j$. Then $Q(\xi, \eta)$ is a polynomial of minimal weighted degree with $Q(\xi_i, \eta_i) = 0$ for $i = 1, \ldots, n$.*

Proof. Let $\tilde{Q}(\xi) \in \mathbb{F}[\xi, \eta]$ be such that $\tilde{Q}(\xi_i, \eta_i) = 0$ for $i = 1, \ldots, n$. Write $\tilde{Q}(\xi, \eta) = \sum_{j=0}^{M} \tilde{d}_j(\xi)\eta^j$ and $\tilde{d}(\xi) = \begin{bmatrix} \tilde{d}_0(\xi) & \cdots & \tilde{d}_M(\xi) \end{bmatrix}$. Then

$$\begin{bmatrix} \tilde{d}_0(\sigma) & \cdots & \tilde{d}_M(\sigma) \end{bmatrix} w_i = 0 \text{ for } i = 1, \ldots, n.$$

It follows from the definition of MPUM that $\tilde{d}(\sigma)w = 0$ for all $w \in \mathfrak{B}$. It follows from Lemma 2.1 that there exists $F(\xi) \in \mathbb{F}^{1\times(M+1)}[\xi]$ such that $\tilde{d}(\xi) = F(\xi)R(\xi)$. It now follows from Corollary 2.10 that the weighted row degree of $\tilde{d}(\xi)$ is larger than or equal to the weighted row degree of $d(\xi)$. This means that the weighted degree of $\tilde{Q}(\xi, \eta)$ is larger than or equal to the weighted degree of $Q(\xi, \eta)$. □

5 Conclusions

In this paper we have presented a systems theoretic approach to list decoding using the concept of behavior. The contribution lies in the behavioral solution to the bivariate interpolation problem associated to the decoding problem. Particularly attractive from a conceptual point of view is the modularity of the proposed solution. With the received word a set of trajectories is associated. These trajectories in turn generate a behavior $\mathfrak{B}$. This behavior may be represented as the kernel of a matrix of polynomials in the shift. After transforming this matrix into weighted row reduced form a row of minimal weighted row degree is selected. Finally, the interpolating bivariate polynomial is obtained from that row in a straightforward fashion. In [11] we also present an algorithm that is iterative in the data points. At each step weighted row

reducedness is guaranteed so that the transformation to weighted row reduced at the end is superfluous. The algorithm is in fact a reformulation of [16]. It produces as a by product a parametrization of all interpolants. Some questions remain. Although we only need one row of minimal weighted row degree, we compute the complete weighted row reduced representation of the behavior $\mathfrak{B}$. Having that matrix at our disposal, the question arises whether the other rows may help in the decoding process. In particular, the row of minimal weighted degree may not be unique and we may ask ourselves what additional useful information about the transmitted codeword is possibly carried by the other rows.

References

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

[2] T. Kailath. *Linear Systems*. Prentice Hall, Englewood Cliffs, N.J, 1980.

[3] M. Kuijper. An algorithm for constructing a minimal partial realization in the multivariable case. *Systems & Control Letters*, 31:225–233, 1997.

[4] M. Kuijper. The Berlekamp-Massey algorithm, error-correction, keystreams and modelling. In G. Picci and D.S. Gilliam, editors, *Dynamical Systems, Control, Coding Computer Vision: New trends, Interfaces, and Interplay*, Progress in Systems and Control Theory, pages 321–342. Birkhäuser, 1999.

[5] M. Kuijper. Further results on the use of a generalized B-M algorithm for BCH decoding beyond the designed error-correcting capability. In *Proceedings of the 13th Symposium on Applied Algebra Algebraic Algorithms, and Error-Correcting Codes (AAECC)*, pages 98–99, Hawaii, USA, 1999.

[6] M. Kuijper. Parametrizations and finite options,. In J.W. Polderman and H.L. Trentelman, editors, *The Mathematics of Systems and Control: from Intelligent Control to Behavioural Systems (Festschrift on the occasion of the 60th birthday of Jan C. Willems)*, pages 59–72, 1999.

[7] M. Kuijper. A system-theoretic derivation of the Welch-Berlekamp algorithm. In *Proceedings 2000 IEEE International Symposium in Information Theory*, page 418, Sorrento, Italy, 2000.

[8] M. Kuijper. Algorithms for decoding and interpolation. In Brian Marcus and Joachim Rosenthal, editors, *Codes, Systems, and Graphical Models*, volume 123 of *The IMA Volumes in Mathematics and its Applications*, pages 265–282. Springer-Verlag, 2001.

[9] M. Kuijper, R. Pinto, and J.W. Polderman. The predictable degree property and row reducedness for systems over a finite ring. *Linear Algebra and its Applications*, 2007. To appear.

[10] M. Kuijper and J.W. Polderman. A behavioral approach to list decoding. In D. Gilliam and J. Rosenthal, editors, *Proceedings of 15th International Symposium on Mathematical Theory of Networks and Systems*, pages 1–13, University of Notre Dame, Indiana, USA, 2002.

[11] M. Kuijper and J.W. Polderman. Behavioral models for list decoding. *Journal of Mathematical and Computer Modeling of Dynamical Systems (MCMDS)*, 8:429–443, 2002.

[12] M. Kuijper and J.W. Polderman. Reed-Solomon list decoding from a system theoretic perspective. *IEEE Trans. Inf. Th.*, IT-50:259–271, 2004.

[13] M. Kuijper, M. van Dijk, H. Hollmann, and A J. Oostveen. A unifying system-theoretic framework for errors-and-erasures Reed-Solomon decoding. In S. Boztas and I.E. Shparlinski, editors, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, LN in Computer Science 2227, pages 343–352. Springer, 2001.

[14] M. Kuijper and J.C. Willems. On constructing a shortest linear recurrence relation. *IEEE Trans. Aut. Control*, 42:1554–1558, 1997.

[15] M. Kuijper and J.C. Willems. An algorithm for computing a shortest linear recurrence relation for a sequence of matrices: generalizing the Berlekamp-Massey algorithm. In *Proceedings 1998 IEEE International Symposium on Information Theory (ISIT'98, MIT Cambridge, USA)*, page 441, 1998.

[16] R. Nielsen and T. Hoeholdt. Decoding Reed-Solomon codes beyond half the minimum distance. In J. Buchmann, T. Hoeholdt, T. Stichtenoth, and H. Tapia-Recillas, editors, *Coding Theory, Cryptography and Related Areas*, pages 221–236, Berlin, 2000. Springer-Verlag.

[17] J.W. Polderman and J.C. Willems. *Introduction to mathematical systems theory: a behavioral approach*, volume 26 of *Texts in Applied Mathematics*. Springer, New York NY, USA, 1997.

[18] M. Sudan. Decoding of Reed-Solomon codes beyond the error correction bound. *J. Compl*, 13:180–193, 1997.

[19] M. Sudan. Decoding of Reed-Solomon codes beyond the error correction diameter. In *Proceedings of the 35th Allerton Conference on Communication Control and Computing*, 1997.

[20] J.C. Willems. From time series to linear system. part II: Exact modelling. *Automatica*, 22:675–694, 1986.

[21] J.C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Aut. Control*, 36:259–294, 1991.