A behavioral approach to list decoding

Margreta Kuijper
Department of Electrical and Electronic Engineering
The University of Melbourne, VIC 3052 Australia
E-mail: m.kuijper@ee.mu.oz.au

Jan Willem Polderman
Faculty of Mathematical Sciences
University of Twente
P.O.Box 217, 7500 AE Enschede, The Netherlands
E-mail: J.W.Polderman@math.utwente.nl

Abstract

List decoding may be translated into a bivariate interpolation problem. The interpolation problem is to find a bivariate polynomial of minimal weighted degree that interpolates a given set of pairs taken from a finite field. We present a behavioral approach to this interpolation problem. With the data points we associate a set of trajectories. For this set of trajectories we construct the Most Powerful Unfalsified Model. The bivariate polynomial is then derived from a specific representation of the MPUM.

1 Introduction and problem statement

In this paper we present a behavioral interpretation of the list decoding approach that was proposed in [1]. We concentrate on the behavioral elements and keep the coding details to a minimum that is just sufficient to appreciate the lines of thought. A more elaborate treatment will be presented in a forthcoming paper. The paper is a follow up of [2, 3, 4] and works out the suggestion made there to put list decoding in the perspective of multivariable behavioral interpolation.

Briefly, an (n, κ) Reed-Solomon code is defined as follows. Let $\mathbb{F}$ be a finite field, say $\mathbb{F} = \{\xi_1, \ldots, \xi_n\}$. The message word is a κ-tuple $(m_0, m_1, \ldots, m_{\kappa-1}) \in \mathbb{F}^\kappa$. With this κ-tuple we associate the polynomial $m(\xi) = m_0 + m_1\xi + \ldots + m_{\kappa-1}\xi^{\kappa-1} \in \mathbb{F}[\xi]$. The codeword $\mathbf{c}$ is then the n-tuple of the evaluations of $m(\xi)$ at the elements of $\mathbb{F}$, that is, $\mathbf{c} = (m(\xi_1), \ldots, m(\xi_n))$.

It should be remarked that the codeword may also consist of the evaluation at the elements of a subset of $\mathbb{F}$. For simplicity and ease of notation we do not consider this possibility here. 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 a recent paper, [1], a list decoding scheme based on bi-variate interpolation was proposed. In list decoding a *list* of possible polynomials $m(\xi)$ is derived from the received word.

The idea put forward in [1] is as follows. Denote the received word by $\mathbf{r} = (\eta_1, \ldots, \eta_n)$. Let $Q(\xi, \eta) \in \mathbb{F}[\xi, \eta]$ be a bivariate polynomial of minimal $(1, \kappa - 1)$ *weighted degree*, defined below, such that $Q(\xi_i, \eta_i) = 0$ for $i = 1 \ldots, n$.

Definition 1.1. Let $Q(\xi, \eta) \in \mathbb{F}[\xi, \eta]$, say $Q(\xi, \eta) = \sum_{i \in I, j \in J} q_{ij} \xi^i \eta^j$. The (w_ξ, w_η) weighted degree of $Q(\xi, \eta)$ is defined as

$$\operatorname{wdeg} Q(\xi, \eta) = \max_{i \in I, j \in J} \{ i w_\xi + j w_\eta \mid q_{ij} \neq 0 \} \tag{1.1}$$

A convenient alternative description of the weighted degree is

$$\operatorname{wdeg} Q(\xi, \eta) = \max_{\deg p(\xi) = w_\eta} \deg Q(\xi^{w_\xi}, p(\xi)). \tag{1.2}$$

In fact, in most but not all cases, the weighted degree is just the normal degree of $Q(\xi^{w_\xi}, \xi^{w_\eta})$. 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. Let $\ell = \operatorname{wdeg} Q(\xi, \eta)$. Suppose now that the received word contains less than $n - \ell$ errors. Then there exists a polynomial $\tilde{m}(\xi)$ of degree less than κ such that $\tilde{m}(\xi_i) = \eta_i$ for at least $\ell + 1$ values of i. In fact, the original polynomial $m(\xi)$ does this, but there can be more. We conclude that $Q(\xi, \tilde{m}(\xi))$ has at least $\ell + 1$ zeros. On the other hand, $\deg Q(\xi, \tilde{m}(\xi))$ cannot exceed ℓ since by assumption $\operatorname{wdeg} Q(\xi, \eta) = \ell$ so that by (1.2) $\deg Q(\xi, \tilde{m}(\xi)) \leq \ell$. Since a polynomial of degree not exceeding ℓ can only have more than ℓ roots if it is the zero polynomial, it follows that $Q(\xi, \tilde{m}(\xi))$ is indeed the zero polynomial. But this implies that $\eta - \tilde{m}(\xi)$ divides $Q(\xi, \eta)$. In particular $\eta - m(\xi)$ divides $Q(\xi, \eta)$. The list decoding now consists of constructing a polynomial $Q(\xi, \eta)$ such that $Q(\xi_i, \eta_i) = 0$ and such that $\operatorname{wdeg} Q(\xi, \eta)$ is minimal. Once $Q(\xi, \eta)$ has been constructed all factors of the form $\eta - \tilde{m}(\xi)$ are extracted thus producing a list of candidate polynomials $\tilde{m}(\xi)$. These candidate polynomials are subsequently checked to produce a sublist of most likely message words. The present paper concentrates on the construction of the polynomial $Q(\xi, \eta)$ of minimal weighted degree.

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 n trajectories. Then we construct a *weighted*

degree row reduced matrix $R(\xi)$ that represents $\mathfrak{B}$. The notion of weighted degree row reduced matrix 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, of course, 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 $(1, \kappa - 1)$ weighted degree that interpolates the data points (ξ_i, η_i) for $i = 1, \ldots, n$.

The outline of the paper is as follows. In Section 2 we review the elements from the behavioral approach that are relevant for our problem. Wherever appropriate we leave the field $\mathbb{F}$ unspecified, however, there are a few instances where it is essential to realize that we are using finite fields. This is particularly true for the characterization of autonomous behaviors defined by a square matrix of polynomials over $\mathbb{F}$. Section 3 presents the behavioral solution of the interpolation problem. Section 4 treats a more general problem, namely the construction of a polynomial $Q(\xi,\eta)$ that interpolates the data points (ξ_i, η_i) with multiplicity. What is meant by that is explained in detail in Section 4. Finally, in Section 5, some conclusions are drawn.

2 Representations of linear time-invariant behaviors

In this section we review some basic concepts of the behavioral approach to linear systems. In the sequel $\mathbb{F}$ is a field. Later in the paper $\mathbb{F}$ will be a finite field of characteristic p. 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 $\mathbb{W}^{\mathbb{T}}$. 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 $(\sigma w)(k) = w(k+1)$. Furthermore σ^j is defined as $\sigma^j \mathbf{w} = \sigma^{j-1}(\sigma \mathbf{w})$. 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

$$\mathfrak{B} = \{\mathbf{w} : \mathbb{Z}_+ \to \mathbb{F}^q \mid R(\sigma)\mathbf{w} = 0\}. \tag{2.3}$$

It is easy to see that $\mathfrak{B}$ is linear. Moreover, $\mathfrak{B}$ is time-invariant. By that we mean that for every $\mathbf{w} \in \mathfrak{B}$ we have that also $\sigma\mathbf{w} \in \mathfrak{B}$. The class of behaviors in q variables that admit a representation of the form $R(\sigma)\mathbf{w} = 0$ is denoted by $\mathfrak{L}^q$.

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. *Let $R_i(\xi) \in \mathbb{F}^{g_i\times q}[\xi]$ and denote the corresponding behaviors by $\mathfrak{B}_i$, $i = 1, 2$. 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 $U(\xi) \in \mathbb{F}^{g\times g}$ is said to be *unimodular* if there exists $V(\xi) \in \mathbb{F}^{g\times g}$ 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 2.1. *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 2.1 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 key in the application of the behavioral approach to coding theory is the *row reduced form*.

Definition 2.1. 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 $D(\xi) = \mathrm{diag}(\xi^{d_1}, \ldots, \xi^{d_g})$ 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 strictly 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*.

Theorem 2.2. *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 degree row reduced*.

Definition 2.2. Let $n_1, \ldots, n_q$ be nonnegative integers. Define $N(\xi) = \mathrm{diag}(\xi^{n_1}, \ldots, \xi^{n_q})$. The matrix $R(\xi) \in \mathbb{F}^{g\times q}[\xi]$ is called $(n_1, \ldots, n_q)$ weighted degree row reduced if $R(\xi)N(\xi)$ is row reduced.

Theorem 2.3. *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 $(n_1, \ldots, n_q)$ row reduced.*

Proof. Let $N(\xi)$ be as in Definition 2.2. According to Theorem 2.2 there exists a unimodular matrix $U(\xi)$ such that $U(\xi)R(\xi)N(\xi)$ is row reduced. But then, by definition, $U(\xi)R(\xi)$ is $(n_1, \ldots, n_q)$ weighted degree row reduced. □

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

The next two results show that row reducedness indicates minimality. This observation is crucial in the behavioral interpretation of the decoding scheme of [1].

Lemma 2.2. *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.*

Proof. Denote the row degrees of $R(\xi)$ by $d_1, \ldots, d_g$ and define $D(\xi) = \operatorname{diag}(\xi^{d_1}, \ldots, \xi^{d_g})$. Since $R(\xi)$ is row reduced there exist matrices $R_0 \in \mathbb{F}^{g\times q}$ and $R_1(\xi) \in \mathbb{F}^{g\times q}[\xi]$ such that

$$R(\xi) = D(\xi)R_0 + R_1(\xi), \tag{2.4}$$

with R_0 of full row rank and $D(\xi)^{-1}R_1(\xi)$ strictly proper. Let $a(\xi) \in \mathbb{F}^{1\times g}[\xi]$, $a(\xi) \neq 0$, and define $r(\xi) = a(\xi)R(\xi)$ and $b(\xi) = a(\xi)D(\xi)$. Denote the row degree of $b(\xi)$ by d. Obviously $d \geq \min(d_1, \ldots, d_g)$. And hence we can write

$$\frac{b(\xi)}{\xi^d} = b_0 + b_1(\xi), \tag{2.5}$$

with $b_0 \neq 0$ and $b_1(\xi)$ strictly proper. Therefore

$$\frac{r(\xi)}{\xi^d} = \underbrace{b_0 R_0}_{\neq 0} + \underbrace{\underbrace{b_1(\xi)R_0}_{\text{strictly proper}} + \underbrace{\frac{b(\xi)}{\xi^d} D(\xi)^{-1}R_1(\xi)}_{\text{strictly proper}}}_{\text{strictly proper}}. \tag{2.6}$$

The conclusion is that the row degree of $r(\xi)$ is equal to d. This proves the statement. □

Corollary 2.1. *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.*

Proof. Let $a(\xi) \in \mathbb{F}^{1\times g}[\xi]$, $a(\xi) \neq 0$, and define $r(\xi) = a(\xi)R(\xi)$. Let the diagonal matrix $W(\xi) = \operatorname{diag}(1, \xi^{\kappa-1}, \ldots, \xi^{(q-1)(\kappa-1)})$. Since $R(\xi)$ is weighted row reduced $R(\xi)W(\xi)$ is row reduced. By Lemma 2.2 the row degree of $r(\xi)W(\xi)$ is larger than or equal to the minimal row degree of $R(\xi)W(\xi)$. As a consequence the weighted row degree of $r(\xi)$ is at least the minimal weighted row degree of $R(\xi)$. □

Given a trajectory $\mathbf{w} : \mathbb{Z}_+ \to \mathbb{F}^q$, or a finite number of trajectories $\mathbf{w}_j : \mathbb{Z}_+ \to \mathbb{F}^q$, $j = 1, \ldots, N$. One may want to find a dynamical system whose behavior contains these specific trajectories. From a modeling perspective it appears sensible to look for the smallest behavior that contains the N trajectories. The following result states that this smallest behavior exists.

Theorem 2.4. *Let $\mathbf{w}_j : \mathbb{Z}_+ \to \mathbb{F}^q$, $j = 1, \ldots, N$ be given. Then there exists a unique behavior $\mathfrak{B}$, referred to as the Most Powerful Unfalsified Model (MPUM for short) of the $\mathbf{w}_j$'s, in the class $\mathfrak{L}^q$ with the properties: 1. $\mathbf{w}_j \in \mathfrak{B}$, $j = 1, \ldots, N$. 2. If $\mathbf{w}_j \in \mathfrak{B}'$, $j = 1, \ldots, N$ for some $\mathfrak{B}' \in \mathfrak{L}^q$ then $\mathfrak{B} \subset \mathfrak{B}'$.*

As remarked, behaviors are represented by polynomial matrices. The question arises how, for a given polynomial matrix, the behavior can be determined explicitly. In order to enable results that make sense for finite fields we first need to introduce the concept of Hasse derivative [6, 7] (called *hyperderivative* in [9, p. 303]). Let $P(\xi) = \sum p_i \xi^i$ be a polynomial with coefficients in a field $\mathbb{F}$. Then the polynomial $D^j_{\mathrm{H}} P(\xi) := \sum \binom{i}{j} p_i \xi^{i-j}$ is called the jth *Hasse derivative* of $P(\xi)$. Note that $j!$ times $D^j_{\mathrm{H}} P(\xi)$ equals the usual "formal derivative" $\frac{d^j P(\xi)}{d\xi^j}$. In fact, the jth Hasse derivative yields exactly the jth order Taylor coefficient of $P(\xi)$. In finite fields, say of characteristic p, the Hasse derivative is much more useful than the formal derivative because whenever $j \geq p$ we have $j! = 0$ and hence all jth formal derivatives vanish. A key property of Hasse derivatives which we use in the sequel is given by the following lemma:

Lemma 2.3. *The polynomial $(\xi - \lambda)^m$ divides $P(\xi)$ if and only if λ is a root of the first $m-1$ Hasse derivatives of $P(\xi)$.*

Proof. Follows immediately from the "Repeated Factor Test" of [7], see also [9]. □

Let us now continue to determine an explicit expression for a behavior in terms of its polynomial representation. Our key players will be trajectories $\boldsymbol{w}_i^j : \mathbb{Z}_+ \to \mathbb{F}$ defined by

$$w_i^j(k) := \begin{array}{ll} \binom{k}{j} \lambda_i^{k-j} & \text{for } k \geq j \\ 0 & \text{for } k < j \end{array},$$

where $\lambda_i \in \mathbb{F}$. Note that the trajectory $\boldsymbol{w}_i^{j+1}$ is the Hasse derivative of trajectory $\boldsymbol{w}_i^j$. We first treat the simplest case, namely where $R(\xi)$ is scalar.

Theorem 2.5. *Let $R(\xi) \in \mathbb{F}[\xi]$ be a polynomial of degree n and let $\mathfrak{B} = \{\boldsymbol{w} : \mathbb{Z}_+ \to \mathbb{F} \mid R(\sigma)\boldsymbol{w} = 0\}$. Then $\mathfrak{B}$ is an n-dimensional subspace of $\mathbb{F}^{\mathbb{Z}_+}$. If the roots of $R(\xi)$ are distinct and belong to $\mathbb{F}$, say $R(\xi) = \prod_{i=1}^N (\xi - \lambda_i)^{m_i}$, with $\lambda_i \in \mathbb{F}$, then*

$$\mathfrak{B} = \mathrm{span}\{\boldsymbol{w}_i^j \mid i = 1, \ldots, N; j = 0, \ldots, m_i - 1\},$$

that is, a trajectory $\mathbf{w} \in \mathfrak{B}$ if and only if there exist coefficients $\xi_{ij} \in \mathbb{F}$ such that

$$w(k) = \sum_{i=1}^{N} \sum_{j=0}^{m_i - 1} \binom{k}{j} \xi_{ij} \lambda_i^{k-j} \tag{2.7}$$

Proof. It is easy to see that $R(\sigma)\boldsymbol{w}_i^0 = R(\lambda_i)\boldsymbol{w}_i^0$. Analogous to the proof in [5, Chapter 3] for the case $\mathbb{F} = \mathbb{C}$, this expression is now differentiated $m_i - 1$ times at λ_i. The only difference is that the Hasse derivative is used instead of the formal derivative for reasons as outlined above. □

Remark 2.1. In terms of the Hasse derivative the expression (2.7) may conveniently be written as:

$$w(k) = \sum_{i=1}^{N} \sum_{j=0}^{m_i - 1} \xi_{ij} D^j_{\mathrm{H}}(\lambda_i^k). \tag{2.8}$$

Example 2.1. Take $\mathbb{F} = \mathbb{Z}/3$ and $R(\xi) = (\xi - 2)^4$. Then every solution of $R(\sigma)\mathbf{w} = 0$ is of the form

$$w(k) = \sum_{j=0}^{3} \xi_j \binom{k}{j} 2^k = \xi_0 2^k + \xi_1 k 2^k + \xi_2 \frac{k^2 - k}{2} 2^k + \xi_3 \frac{1}{6}(k^3 - 3k^2 + 2k)2^k, \quad \xi_j \in \mathbb{Z}/3. \tag{2.9}$$

The multivariable case, $q > 1$, is somewhat more involved, but basically analogous to the scalar case, as apparent from the following theorem.

Theorem 2.6. *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 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*

$$w(k) = \sum_{i=1}^{n} b_i \lambda_i^k. \tag{2.10}$$

Here, $b_i \in \mathbb{F}^{q\times q}$ such that $R(\lambda_i)b_i = 0$. More generally, if $\det R(\xi) = \prod_{i=1}^{N}(\xi - \lambda_i)^{m_i}$, with $\lambda_j \in \mathbb{F}$, then all trajectories in $\mathfrak{B}$ are of the form

$$w(k) = \sum_{i=1}^{N} \sum_{j=0}^{m_i-1} b_{ij} \binom{k}{j} \lambda_i^{k-j} = \sum_{i=1}^{N} \sum_{j=0}^{m_i-1} b_{ij} D_{\mathrm{H}}^j(\lambda^k) \tag{2.11}$$

where $b_{ij} \in \mathbb{F}^q$ satisfy the linear restrictions:

$$\sum_{j=\ell}^{m_i-1} \left[D_{\mathrm{H}}^{j-\ell} R(\lambda_i) \right] b_{ij} = 0 \quad \ell = 0, \ldots, m_i - 1, \quad i = 1, \ldots, N. \tag{2.12}$$

Example 2.2. Let $R(\xi) \in \mathbb{Z}/3[\xi]$ be given by

$$R(\xi) = \begin{bmatrix} \xi^4 + \xi^3 + \xi + 1 & \xi^5 + \xi^4 + \xi^2 + 2\xi + 1 \\ \xi^5 + \xi^4 + \xi^2 + \xi & \xi^6 + \xi^5 + \xi^3 + 2\xi + 1 \end{bmatrix}. \tag{2.13}$$

Then $\det R(\xi) = \xi^6 + 2\xi^5 + 2\xi^4 + 2\xi^3 + +2\xi^2 + 2\xi + 1 = (\xi - 1)^2(\xi - 2)^4$. Using Theorem 2.6 it follows that all solutions of $R(\sigma)\boldsymbol{w} = 0$ are of the form

$$w(k) = \begin{bmatrix} \xi_{11} \\ \xi_{12} \end{bmatrix} + \begin{bmatrix} 0 \\ 2\xi_{11} \end{bmatrix} k + \begin{bmatrix} \xi_{21} \\ 0 \end{bmatrix} 2^k + \begin{bmatrix} \xi_{22} \\ 0 \end{bmatrix} k 2^{k-1} + \begin{bmatrix} \xi_{23} \\ 0 \end{bmatrix} \binom{k}{2} 2^{k-2} + \begin{bmatrix} \xi_{24} \\ 0 \end{bmatrix} \binom{k}{3} 2^{k-3} \tag{2.14}$$

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. Combining Theorems 2.4 and 2.5 we are able to find a representation of the MPUM of exponential trajectories.

Example 2.3. Let $a \in \mathbb{F}$ and $w(k) = a^k$. The MPUM of $\boldsymbol{w}$ is represented by $R(\xi) = \xi - a$. If $w_i(k) = \xi_i^k$, with $\xi_i \in \mathbb{F}$, $i = 1, .., n$, then the MPUM of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is represented by $R(\xi) = \prod_{i=1}^n (\xi - \xi_i)$.

In the multivariable case, for $w(k) = va^k$ with $v \in \mathbb{F}^q$ and $a \in \mathbb{F}$ a representation of the MPUM of $\boldsymbol{w}$ should satisfy: $\det R(\xi) = \xi - a$ and $R(a)v = 0$.

Given a finite set of trajectories $\mathbf{w}_1, \ldots, \mathbf{w}_n$ in $\mathbb{F}^{\mathbb{Z}_+}$ a well-known recursive technique to construct the MPUM of these trajectories is the following. Let $R_m(\xi)$ represent the MPUM of $\mathbf{w}_1, \ldots, \mathbf{w}_m$. Define $\tilde{\mathbf{w}}_{m+1} := R_m(\sigma)\mathbf{w}_{m+1}$ and let $\tilde{R}_{m+1}(\xi)$ be a representation of the MPUM of $\tilde{\mathbf{w}}_{m+1}$. Define $R_{m+1}(\xi) := \tilde{R}_{m+1}(\xi)R_m(\xi)$. Then $R_{m+1}(\xi)$ represents the MPUM of $\mathbf{w}_1, \ldots, \mathbf{w}_{m+1}$.

3 Minimal interpolation

The problem treated in this section is as follows. Given n pairs $(\xi_i, \eta_i) \in \mathbb{F}^2$, $i = 1, \ldots n$. Find a polynomial $Q(\xi, \eta) \in \mathbb{F}[\xi, \eta]$ of minimal $(1, \kappa - 1)$ weighted degree, see Definition 1.1, such that $Q(\xi_i, \eta_i) = 0$ for $i = 1, \ldots, n$.

To solve the above problem we can nicely apply the behavioral theory. The outline is as follows. First we write $Q(\xi, \eta)$ as $Q(\xi, \eta) = \sum_{j=0}^M d_j(\xi)\eta^j$. What we are aiming at is $Q(\xi_i, \eta_i) = 0$, i.e., $\sum_{j=0}^M d_j(\xi_i)\eta_i^j = 0$. Recalling Theorem 2.4 the behavioral interpretation is almost straightforward. Namely, we are looking for a polynomial vector $d(\xi) = [d_0(\xi), \ldots, d_M(\xi)]$ such that

$$\begin{bmatrix} d_0(\xi_i) & \cdots & d_M(\xi_i) \end{bmatrix} \begin{bmatrix} 1 \\ \eta_i \\ \vdots \\ \eta_i^M \end{bmatrix} = 0 \text{ for } i = 1, \ldots, n. \tag{3.15}$$

In the light of Theorem 2.5 this is equivalent to

$$\underbrace{\begin{bmatrix} d_0(\sigma) & \cdots & d_M(\sigma) \end{bmatrix}}_{d(\sigma)} \underbrace{\left(\begin{bmatrix} 1 \\ \eta_i \\ \vdots \\ \eta_i^M \end{bmatrix} \xi_i^k \right)}_{w_i(k)} = 0 \quad \text{for } i = 1, \ldots, n. \tag{3.16}$$

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 $d(\sigma)\mathbf{w}_i = 0$ for $i = 1, \ldots, n$. Notice that if $d_M(\xi) \neq 0$, then $\operatorname{wdeg} Q(\xi, \eta) \geq M(\kappa - 1)$. Of course, for $M = 0$ there exists a trivial solution, namely $Q(\xi, \eta) = \prod_{i=1}^n (\xi - \xi_i)$. This solution has weighted degree n. The minimal weighted degree does therefore not exceed n. It follows that we can take

$$M = \max\{j \in \mathbb{N} \mid j \leq \frac{n}{\kappa - 1}\}. \tag{3.17}$$

Remark 3.1. A tighter upperbound for the minimal weighted degree can be expressed in terms of both n and κ. It is based on a counting argument, see [1, Lemma 7]. This upperbound can then be used to derive a possibly smaller choice of M.

The idea is now to find a representation $\tilde{R}(\xi)$ of the MPUM of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ and subsequently transform $\tilde{R}(\xi)$ into a weighted degree row reduced matrix $R(\xi)$. 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 3.1. *Let $\mathfrak{B}$ be the MPUM of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ defined in (3.16) with M defined by (3.17). Let $R(\xi) \in \mathbb{F}^{(M+1)\times(M+1)}[\xi]$ be a weighted degree 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 $(1, \kappa - 1)$ 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

$$\begin{bmatrix} \tilde{d}_0(\sigma) & \cdots & \tilde{d}_M(\sigma) \end{bmatrix} \mathbf{w}_i = 0 \text{ for } i = 1, \ldots, n. \tag{3.18}$$

It follows from the definition of MPUM that $\tilde{d}(\sigma)\mathbf{w} = 0$ for all $\mathbf{w} \in \mathfrak{B}$. 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 2.1 that the weighted row degree of $\tilde{d}(\xi)$ is larger than or equal to the weighted row degree of $d(\xi)$. It follows that the $(1, \kappa - 1)$ weighted degree of $\tilde{Q}(\xi, \eta)$ is larger than or equal to the $(1, \kappa - 1)$ weighted degree of $Q(\xi, \eta)$. □

Example 3.1. As an example we take $\mathbb{F} = \mathbb{Z}/7$ and $\kappa = 3$. The pairs that we want to interpolate are $(0, 6), (1, 3), (2, 4), (3, 6)), (4, 4), (5, 2), (6, 5)$. We want to find a polynomial $Q(\xi, \eta)$ of minimal $(1, \kappa - 1)$ weighted degree that interpolates the given data points. Following the exposition above we take $M = 3$ and define seven trajectories in $(\mathbb{Z}/7)^4$ as follows:

$$w_1(k) = \begin{bmatrix} 1 \\ 6 \\ 6^2 \\ 6^3 \end{bmatrix} 0^k =: Y_1 \xi_1^k \quad \cdots \quad w_7(k) = \begin{bmatrix} 1 \\ 5 \\ 5^2 \\ 5^3 \end{bmatrix} 6^k =: Y_7 \xi_7^k. \tag{3.19}$$

Next we construct a representation of the MPUM of $\mathbf{w}_1, \cdots, \mathbf{w}_7$. According to Theorem 2.6 we want to find a matrix $R(\xi) \in (\mathbb{Z}/7)^{4\times 4}[\xi]$ such that $\det R(\xi) = (\xi - \xi_1) \cdots (\xi - \xi_7)$ and $R(\xi_i)Y_i = 0$ for $i = 1, \ldots 7$. We construct $R(\xi)$ as follows: $R_{11}(\xi) = (\xi - \xi_1) \cdots (\xi - \xi_7)$; $R_{ii}(\xi) = 1$ for $i = 2, \ldots, 4$; $R_{ij}(\xi) = 0$ for $i = 2, \ldots, 4$, $j = 2, \ldots 4$, $i \neq j$. Finally, for the remaining entries we take interpolating Lagrange polynomials, that is, polynomials of degree

at most six whose coefficients solve $R(\xi_i)Y_i = 0$. This yields

$$R(\xi) = \begin{bmatrix} \xi\,(\xi-1)\,(\xi-2)\,(\xi-3)\,(\xi-4)\,(\xi-5)\,(\xi-6) & 0 & 0 & 0 \\ 1+2\,\xi+4\,\xi^2+5\,\xi^3+3\,\xi^4+\xi^5+2\,\xi^6 & 1 & 0 & 0 \\ 6+6\,\xi+2\,\xi^2+4\,\xi^3+\xi^4+5\,\xi^5+2\,\xi^6 & 0 & 1 & 0 \\ 1+4\,\xi+2\,\xi^2+2\,\xi^3+6\,\xi^4+\xi^5+6\,\xi^6 & 0 & 0 & 1 \end{bmatrix}. \tag{3.20}$$

Then we transform $R(\xi)$ into weighted row reduced form and obtain

$$R_{\text{wred}}(\xi) = \begin{bmatrix} 6\,\xi^4+3\,\xi^3+\xi^2+3\,\xi+4 & 2\,\xi^2+4 & 0 & 0 \\ 2\,\xi^4+\xi^2+1 & 6\,\xi^2+4\,\xi+2 & 1 & 0 \\ 5\,\xi^4+2\,\xi^3+4\,\xi+6 & 3\,\xi^3+2\,\xi^2+5\,\xi & 1 & 0 \\ 6\,\xi^6+\xi^5+6\,\xi^4+2\,\xi^3+2\,\xi^2+4\,\xi+1 & 0 & 0 & 1 \end{bmatrix}. \tag{3.21}$$

The $(1,2)$ weighted row degrees of $R_{\text{wred}}(\xi)$ are $4, 4, 5$, and 6 respectively. It follows that both the first and the second row have minimal weighted row degree. Both yield polynomials $Q(\xi,\eta)$ of minimal $(1,\kappa-1)$ weighted degree that interpolate the data. These are $Q_1(\xi,\eta) = 6\xi^4+3\xi^3+\xi^2+3\xi+4+(2\xi^2+4)\eta$ and $Q_2(\xi,\eta) = 2\xi^4+\xi^2+1+(6\xi^2+4\xi+2)\eta+\eta^2$.

4 Minimal interpolation with multiplicity

The problem that we study in this section is an extension of that in the previous section. It is motivated by an extension of Sudan's approach, which enables the correction of more errors, see [8]. Again we are given n pairs $(\xi_i,\eta_i) \in \mathbb{F}^2$, $i = 1,\ldots,n$. Again, we want to find a polynomial $Q(\xi,\eta) \in \mathbb{F}[\xi,\eta]$ of minimal $(1,\kappa-1)$ weighted degree such that $Q(\xi_i,\eta_i) = 0$ for $i = 1,\ldots,n$. The difference with the previous interpolation problem is that we want (ξ_i,η_i) to be roots of $Q(\xi,\eta)$ of multiplicity $s \geq 1$. The notion of interpolation with multiplicity is explained below.

Definition 4.1. Let $Q(\xi,\eta) \in \mathbb{F}[\xi,\eta]$, say $Q(\xi,\eta) = \sum_{i=0}^{N_x}\sum_{j=0}^{N_y} q_{ij}\xi^i\eta^j$. The pair $(0,0) \in \mathbb{F}^2$ is a root of $Q(\xi,\eta)$ of multiplicity $s \in \mathbb{N}$ if $q_{i,s-1-i} = 0$ for all $i = 0,\ldots,s-1$ and $q_{i',s-i'} \neq 0$ for some $i' \in \{0,\ldots,s\}$. The pair $(a,b) \in \mathbb{F}^2$ is a root of $Q(\xi,\eta)$ of multiplicity $s \in \mathbb{N}$ if $(0,0)$ is a root of $Q(\xi+a,\eta+b)$ of multiplicity s.

The property that (a,b) is a multiple root of $Q(\xi,\eta)$ can be expressed in terms of the Hasse derivatives of $Q(\xi,\eta)$.

Definition 4.2. Let $Q(\xi,\eta) \in \mathbb{F}[\xi,\eta]$, $Q(\xi,\eta) = q_1(\xi)q_2(\eta)$. The (ℓ_1,ℓ_2)th Hasse derivative is defined as $D_{\text{H}}^{\ell_1}q_1(\xi)D_{\text{H}}^{\ell_2}q_2(\eta)$. The Hasse derivative of a general polynomial is defined through the property $D_{\text{H}}^{\ell_1,\ell_2}(Q_1(\xi,\eta)+Q_2(\xi,\eta)) = D_{\text{H}}^{\ell_1,\ell_2}(Q_1(\xi,\eta)) + D_{\text{H}}^{\ell_1,\ell_2}(Q_2(\xi,\eta))$.

Theorem 4.1. *Let $Q(\xi,\eta) \in \mathbb{F}[\xi,\eta]$ and $(a,b) \in \mathbb{F}^2$. Then (a,b) is a root of $Q(\xi,\eta)$ of multiplicity s if and only if*

$$\begin{aligned} \left(D_{\mathrm{H}}^{m-\ell,\ell} Q\right)(a,b) &= 0 \quad m = 0,\ldots,s-1, \quad \ell = 0,\ldots,m \\ \left(D_{\mathrm{H}}^{s-\ell,\ell} Q\right)(a,b) &\neq 0 \quad \textit{some} \quad 0 \leq \ell \leq s \end{aligned} \tag{4.22}$$

Proof. The proof is a direct application of the Taylor expansion of $Q(\xi,\eta)$ about (a,b). It is important to note that the coefficients in (4.22) are well defined elements in $\mathbb{F}$. □

We are now ready to give a behavioral interpretation to the interpolation-with-multiplicity-problem in the same vain as in Section 3. To that end write $Q(\xi,\eta) = \sum_{j=0}^{M} d_j(\xi)\eta^j$. The requirement that $Q(\xi,\eta)$ interpolates (ξ_i,η_i) with multiplicity at least s is equivalent to:

$$\begin{bmatrix} D_{\mathrm{H}}^{m-\ell}(d_0)(\sigma) & \cdots & D_{\mathrm{H}}^{m-\ell}(d_M)(\sigma) \end{bmatrix} \left(\left(D_{\mathrm{H}}^{\ell} \begin{bmatrix} 1 \\ \eta_i \\ \vdots \\ \eta_i^M \end{bmatrix} \right) \xi_i^k \right) = 0 \tag{4.23}$$
$$m = 0,\ldots,s-1, \quad \ell = 0,\ldots,m$$

This is easily seen to be equivalent to

$$\underbrace{\begin{bmatrix} d_0(\sigma) & \cdots & d_M(\sigma) \end{bmatrix}}_{d(\sigma)} \underbrace{\left(\left(D_{\mathrm{H}}^{\ell} \begin{bmatrix} 1 \\ \eta_i \\ \vdots \\ \eta_i^M \end{bmatrix} \right) D_{\mathrm{H}}^{m-\ell} \xi_i^k \right)}_{w_{m\ell i}(k)} = 0$$
$$m = 0,\ldots,s-1, \quad \ell = 0,\ldots,m \tag{4.24}$$

Apparently we are looking for a vector $d(\xi)$ of minimal weighted degree such that $d(\sigma)\mathbf{w}_{m\ell i} = 0$ for $m = 0,\ldots,s-1$, $\ell = 0,\ldots,m$, and $i = 1,\ldots,n$. Notice that whereas (4.24) guarantees interpolation with multiplicity *at least* s, the additional requirement that $Q(\xi,\eta)$ is of minimal weighted degree implies that the multiplicity is *exactly* s.

We now proceed in exactly the same way as in Section 3. That is, we construct a weighted degree row reduced matrix $R(\xi)$ that represents the MPUM of the trajectories $\mathbf{w}_{m\ell i}$. From $R(\xi)$ we select a row $d(\xi) = \begin{bmatrix} d_0(\xi) & \ldots & d_M(\xi) \end{bmatrix}$ of minimal weighted row degree. The desired polynomial $Q(\xi,\eta) = \sum_{j=0}^{M} d_j(\xi)\eta^j$ interpolates the data points (ξ_i,η_i) with multiplicity s and has minimal weighted degree. 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. Analogously to the $s = 1$ case it follows that the following choice of M suffices:

$$M = \max\{j \in \mathbb{N} \mid j \leq \frac{sn}{\kappa - 1}\}. \tag{4.25}$$

Example 4.1. Take $\mathbb{F} = \mathbb{Z}/5$ and $\kappa = 3$. The data points that we want to interpolate are $(\xi_1, \eta_1) = (0, 1), (\xi_2, \eta_2) = (1, 0), (\xi_3, \eta_3) = (2, 2), (\xi_4, \eta_4) = (3, 3), (\xi_5, \eta_5) = (4, 1)$. The multiplicity with which we interpolate is taken to be $s = 2$. Following (4.25) we take $M = 5$. According to (4.24) we define three trajectories for each data point i ($i = 1, \ldots, 5$), namely

$$w_{i1}(k) = \begin{bmatrix} 1 \\ \eta_i \\ \eta_i^2 \\ \vdots \\ \eta_i^5 \end{bmatrix} \xi_i^k \quad w_{i2}(k) = \begin{bmatrix} 0 \\ 1 \\ 2\eta_i \\ \vdots \\ 5\eta_i^4 \end{bmatrix} \xi_i^k \quad w_{i3}(k) = \begin{bmatrix} 1 \\ \eta_i \\ \eta_i^2 \\ \vdots \\ \eta_i^5 \end{bmatrix} k\xi_i^{k-1}. \tag{4.26}$$

Next we determine a representation $R(\xi) \in \mathbb{Z}/5^{6\times 6}[\xi]$ of the MPUM of the fifteen trajectories. Subsequently, $R(\xi)$ is transformed into $(1, 2)$ weighted row reduced form $R_{\mathrm{w}}(\xi)$. Then we select a row of $R_{\mathrm{w}}(\xi)$ of minimal weighted row degree. The calculations were done by Maple and yielded:

$$d(\xi) = \begin{bmatrix} \xi^6 + 2\xi^4 + \xi^3 + 3\xi + 3 & 3\xi^4 + 3\xi^3 + 3\xi^2 + \xi & \xi + 1 & 1 & 0 & 0 \end{bmatrix} \tag{4.27}$$

as a row of minimal weighted row degree. The corresponding interpolating bivariate polynomial of minimal weighted degree is hence given by

$$Q(\xi, \eta) = \xi^6 + 2\xi^4 + \xi^3 + 3\xi + 3 + (3\xi^4 + 3\xi^3 + 3\xi^2 + \xi)\eta + (\xi + 1)\eta^2 + \eta^3 \tag{4.28}$$

It straightforward to check that $Q(\xi_i, \eta_i) = 0$ and that moreover $Q(\xi + \xi_i, \eta + \eta_i)$ has terms of ordinary degree two and higher so that $Q(\xi, \eta)$ indeed interpolates the given data with multiplicity two.

A bivariate polynomial of minimal weighted degree that interpolates with multiplicity just one is given by

$$\tilde{Q}(\xi, \eta) = \xi^3 + 4 + (4\xi + 1)\eta. \tag{4.29}$$

5 Conclusions

In this paper we presented a behavioral approach to a bivariate interpolation problem over a finite field. The relevance of the interpolation problem to decoding was touched upon only briefly. A more elaborate treatment including efficient algorithms for the recursive determination of weighted degree row reduced polynomial matrices will be presented in a forthcoming paper.

References

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

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

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

[4] M. Kuijper. Behavioral interpolation for coding and control. In *Proc. 39th IEEE Conf. Decision and Control*, pages 2488–2493, Sydney, Australia, 2000.

[5] 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.

[6] H. Hasse. Theorie der höheren Differentiale in einem algebraischen Funktionkörper mit volkommenen Konstantenkörper bei beliebiger Charakteristik. *J. Reine und Angewante Mathematik*, 175:50–54, 1936.

[7] G. Castagnoli, J.L. Massey, P. A. Schoeller, and N. von Seemann. On repeated-root cyclic codes. *IEEE Trans. Inf. Th.*, 37:50–54, 1991.

[8] V. Guruswami and M. Sudan. Improved decoding of Reed-Solomon and algebraic-geometric codes. *IEEE Trans. Inf. Th.*, 45:1757–1768, 1999.

[9] R. Lidl and H. Niederreiter. *Finite fields*. Cambridge University Press, second edition, 1997.