A framework for the solution of the generalized realization problem

A.J. Mayo, A.C. Antoulas *

Department of Electrical, and Computer Engineering, Rice University, Houston, TX 77005, USA

Received 24 October 2006; accepted 4 March 2007
Available online 25 March 2007
Submitted by H. Schneider

Abstract

In this paper we present a novel way of constructing generalized state space representations \([E, A, B, C, D]\) of interpolants matching tangential interpolation data. The Loewner and shifted Loewner matrices are the key tools in this approach.

© 2007 Published by Elsevier Inc.

AMS classification: 15A18; 30E05; 65D05; 93D05; 93D07; 93D15; 93D20; 93D30

Keywords: Rational interpolation; Tangential interpolation; Bi-tangential interpolation; Realization; Loewner matrices; Shifted Loewner matrices; Hankel matrices; Descriptor systems; Generalized controllability matrices; Generalized observability matrices

1. Introduction

The realization problem consists in the simultaneous factorization of a (finite or infinite) sequence of matrices. More precisely, given \(h_t \in \mathbb{R}^{p \times m}, \ t = 1, 2, \ldots\), we wish to find \(A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, C \in \mathbb{R}^{p \times n}\) such that

\[
h_t = CA^{t-1}B, \quad t > 0.
\]

This amounts to the construction of a linear dynamical system in state space form:
such that its transfer function satisfies $H(s) = C(sI - A)^{-1}B = \sum_{t>0} h_t s^{-t}$. The $h_t$ are often referred to as the Markov parameters of $\Sigma$, and correspond to information about the transfer function at infinity; for details see Section 4.4 of [1].

The question thus arises as to whether such a problem can be solved if information about the transfer function at different points of the complex plane is provided. In this paper we will refer to this as the generalized realization problem. This problem is closely related to rational interpolation and can be stated as follows: given data obtained by sampling the transfer matrix of a linear system, construct a controllable and observable state space model of a system consistent with the data. The data will be in one of the following classes:

1. **Scalar data**: given the pairs of scalars $\{(z_i, y_i)|z_i, y_i \in \mathbb{C}, i = 1, \ldots, N\}$, construct $[E, A, B, C, D]$, of appropriate dimensions, such that $H(z_i) = C(z_iE - A)^{-1}B + D = y_i, i = 1, \ldots, N$.

2. **Matrix data**: given the pairs of scalars and matrices $\{|(z_i, Y_i)|z_i \in \mathbb{C}, Y_i \in \mathbb{C}^{n \times m}, i = 1, \ldots, N\}$, construct $[E, A, B, C, D]$, of appropriate dimensions, such that $H(z_i) = C(z_iE - A)^{-1}B + D = Y_i, i = 1, \ldots, N$.

3. **Tangential data**, that is matrix data sampled directionally. In this case the data is composed of the right interpolation data

$$\{(\lambda_i, r_i, w_i) | \lambda_i \in \mathbb{C}, r_i \in \mathbb{C}^{m \times 1}, w_i \in \mathbb{C}^{p \times 1}, i = 1, \ldots, \rho\},$$

or more compactly

$$A = \text{diag}[\lambda_1, \ldots, \lambda_\rho] \in \mathbb{C}^{p \times p}, \quad R = [r_1, \ldots, r_\rho] \in \mathbb{C}^{m \times p},$$

$$W = [w_1, \ldots, w_\rho] \in \mathbb{C}^{p \times p},$$

and of the left interpolation data

$$\{(\mu_j, \ell_j, v_j) | \mu_j \in \mathbb{C}, \ell_j \in \mathbb{C}^{1 \times p}, v_j \in \mathbb{C}^{1 \times m}, j = 1, \ldots, \nu\},$$

or more compactly

$$M = \text{diag}[\mu_1, \ldots, \mu_\nu] \in \mathbb{C}^{v \times v}, \quad L = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_\nu \end{bmatrix} \in \mathbb{C}^{v \times p}, \quad V = \begin{bmatrix} v_1 \\ \vdots \\ v_\nu \end{bmatrix} \in \mathbb{C}^{v \times m}.$$  

We wish to construct $[E, A, B, C, D]$, of appropriate dimensions, such that the associated transfer function $H(s) = C(sE - A)^{-1}B + D$, satisfies both the right constraints

$$H(\lambda_i)r_i = [C(\lambda_iE - A)^{-1}B + D]r_i = w_i, \quad i = 1, \ldots, \rho,$$

and the left constraints

$$\ell_jH(\mu_j) = \ell_j[C(\mu_jE - A)^{-1}B + D] = v_j, \quad j = 1, \ldots, \nu.$$  

Each one of these problems generalizes the previous one. The matching of derivatives can also be included and will be discussed towards the end, in Section 6.

The problem of rational interpolation is one that has been studied, in various forms, for over a century. For example, Pick [17] and Nevanlinna [16] were concerned with the construction of functions that take specified values in the disk, and are bounded therein. It has also been relevant to the engineering community for some time; for example, the use of interpolation in network and
system theory is explored in [21,22]. The problem of unconstrained rational interpolation, that is construction of a rational function without specifying a region of analyticity, has been studied at least since Belevitch [7]. For comprehensive accounts on this topic see the books [6] and [1]. The tangential interpolation problem from a model reduction viewpoint has been studied in [11].

In the last two decades, there have been several approaches to this problem. One is the generating system approach in which a rational matrix function is constructed, called the generating system, that parameterizes the set of all interpolants. Furthermore, if the generating system is column reduced, the allowable degrees of interpolants follow by inspection. Some references for this approach are [5,6,1].

The other approach seeks to directly construct state space models of such interpolants. The realization problem described earlier for instance, can solved using the (block) Hankel matrix

\[ H = \begin{bmatrix}
    h_1 & h_2 & h_3 & \cdots \\
    h_2 & h_3 & h_4 & \cdots \\
    h_3 & h_4 & h_3 & \cdots \\
    \vdots & \vdots & \vdots & \ddots
\end{bmatrix}. \]

The necessary and sufficient solvability condition given an infinite sequence of Markov parameters is that

\[ \text{rank}(H) = n < \infty. \]

The importance of the Hankel matrix is also reflected in the fact that it can be factored into a product of observability and controllability matrices.

For generalized realization problems, a key tool is the Loewner matrix, variously called the divided-difference matrix and null-pole coupling matrix. This was first used systematically for solving rational interpolation problems in [3]. The Loewner matrix is closely related to the Hankel matrix defined above and can be factored into a product of matrices of system theoretic significance. From this factorization, then, a state space model can be constructed; for details we refer to [2] and [3].

Returning to classical realization theory, the construction of a state space model proceeds as follows. Let \( \Delta \in \mathbb{R}^{n \times n} \) be a non-singular submatrix of \( \mathcal{H} \), let \( \sigma \Delta \in \mathbb{R}^{n \times n} \) be the matrix with the same rows, but columns shifted by \( m \) columns, let \( \Gamma \in \mathbb{R}^{n \times m} \) have the same rows as \( \Delta \), but the first \( m \) columns only, and, finally, let \( \Omega \in \mathbb{R}^{p \times n} \) be the submatrix of \( \mathcal{H} \) composed of the same columns as \( \Delta \), but its first \( p \) rows. Then

\[ A = \Delta^{-1} \sigma \Delta, \quad B = \Delta^{-1} \Gamma, \quad C = \Omega, \]

is a state space representation of a system that matches the data. Notice also that the associated transfer function \( H \) can be expressed as

\[ H(s) = \Omega(s\Delta - \sigma \Delta)^{-1} \Gamma. \]  

(8)

One of our aims is to generalize this expression for tangential interpolation data (see Lemma 5.1). Though connections between realization and interpolation have been known, the procedure presented here is the first to construct generalized state space realizations from tangential interpolation data. The key is the introduction of the shifted Loewner matrix associated with the data. Preliminary versions of the results below were presented in [15].

The structure of this paper is the following. In Section 2, we introduce some relevant background material. In particular, we briefly review some results on descriptor systems, as well as selected results on the rational interpolation problem (cf. [3,2,4]). In Section 3, we introduce
the main tools, namely, the Loewner and the shifted Loewner matrices as well as the associated Sylvester equations. We then show how to construct a family of interpolants when the Loewner matrix has full rank. The core of the results including an explicit algorithm for the construction of generalized state space realizations from tangential interpolation data, is presented in Section 5. Section 6 is devoted to data satisfying derivative constraints and Section 7 discusses the parametrization of interpolants in the single-input single-output case and makes contact with the generating system framework. We conclude with Section 8 which presents several illustrative examples.

2. Preliminaries

In the present paper we allow interpolants to be improper, and, in so doing, we diverge from [2]. There are two reasons for this: first, the Loewner matrix carries information about interpolants seemingly without regard to whether or not they are proper. It is natural, then, to consider all interpolants, and then, if the situation demands it, to specialize to proper interpolants. The generating system approach shares this philosophy. The second reason is that interpolation with generalized state-space (descriptor) systems is the natural way to proceed. To this end, we briefly review some facts about these systems. References on the subject are [9,18,20,8,19].

2.1. Generalized state-space (descriptor) systems

Consider the linear dynamical system $\Sigma$ described by a set of differential and algebraic equations

$$\Sigma: E\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t),$$

where $x(t)$ is the pseudo-state, $E, A \in \mathbb{R}^{k \times k}$, $B \in \mathbb{R}^{k \times m}$, $C \in \mathbb{R}^{p \times k}$, $D \in \mathbb{R}^{p \times m}$, are constant matrices with $E$ possibly singular. It is also required that $\det(sE - A) \neq 0$, that is the pencil $sE - A$ be regular. Systems of this type are variously called descriptor systems, singular systems and generalized state-space systems. We will not distinguish between a system and the quintuple of matrices $[E, A, B, C, D]$. In much of the literature on descriptor systems the $D$ matrix is taken to be $0$ [9]. There is ample reason for taking this approach: the $D$ matrix is in a sense never needed as every rational transfer matrix has a state space representation of the form $[E, A, B, C]$. However, we shall see that allowing the $D$ term to be non-zero has its advantages (see Section 5.4).

There are several notions of equivalence that can be applied to such systems. One such notion is restricted system equivalence (r.s.e.).

**Definition 2.1.** Two systems $[E_i, A_i, B_i, C_i, D_i], i = 1, 2$, are r.s.e. if there exist invertible matrices $P, Q \in \mathbb{R}^{k \times k}$ such that

$$PE_1Q = E_2, \quad PA_1Q = A_2, \quad PB_1 = B_2, \quad C_1Q = C_2, \quad D_1 = D_2.$$

This type of equivalence preserves transfer matrices; however the converse is not true: it can be the case that two systems that are not r.s.e. have the same transfer function. Every regular system is r.s.e. to a system of the following form:

$$\dot{x}_1 = A_1x_1 + B_1u, \quad y_1 = C_1x_1 + Du,$$

$$N\dot{x}_2 = x_2 + B_2u, \quad y_2 = C_2x_2,$$
and \( y = y_1 + y_2 \), where \( N \) is nilpotent with index \( \eta \), that is, \( \eta \) is the smallest positive integer such that \( N^\eta = 0 \). This is the Weierstrass canonical form. Thus, given any regular system \([E, A, B, C, D]\), there exists a system \([\tilde{E}, \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}]\) with the same transfer matrix such that

\[
\tilde{E} = \begin{bmatrix} I_{k_1} \\ N \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} A_1 \\ I_{k_2} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \quad \tilde{D} = D,
\]

and \( k_1 + k_2 = k \). The subsystem \([I_{k_1}, A_1, B_1, C_1, D]\) is referred to as the slow subsystem and \([N, I_{k_2}, B_2, C_2, 0]\) as the fast subsystem.

**Definition 2.2.** System (9) is controllable if, for any \( t_1 > 0 \), \( x(0) \in \mathbb{R}^k \) and \( w \in \mathbb{R}^k \), there exists an input \( u(t) \), which is at least \((\eta - 1)\)-times differentiable, such that \( x(t_1) = w \).

**Proposition 2.1.** The following statements are equivalent:

1. System (9) is controllable.
2. \( \text{rank}[B_1 \quad A_1 B_1 \quad \cdots \quad A_1^{k_1-1} B_1] = k_1 \), and \( \text{rank}[B_2 \quad NB_2 \quad \cdots \quad N^{\eta-1} B_2] = k_2 \).
3. \( \text{rank}[sE - A \quad B] = k \), for all finite \( s \in \mathbb{C} \), and \( \text{rank}[E \quad B] = k \).

**Definition 2.3.** System (9) is observable if the initial condition \( x(0) \) can be uniquely determined from \( u(t), y(t) \), for \( 0 \leq t < \infty \).

**Proposition 2.2.** The following statements are equivalent

1. System (9) is observable.
2. \( \text{rank}[C_1^s \quad A_1^s C_1^s \quad \cdots \quad (A_1^s)^{k_1-1} C_1^s] = k_1 \), and \( \text{rank}[C_2^s \quad N^s C_2^s \quad \cdots \quad (N^s)^{\eta-1} C_2] = k_2 \).
3. \( \text{rank}[sE - A] = k \), for all finite \( s \in \mathbb{C} \), and \( \text{rank} \begin{bmatrix} E \\ C \end{bmatrix} = k \).

A straightforward calculation shows that the transfer matrix associated to a descriptor system is completely determined by its controllable and observable parts.

**Proposition 2.3.** Every rational matrix \( H(s) \) has a realization \([E, A, B, C]\) which satisfies \( H(s) = C(sE - A)^{-1} B \); furthermore the realization is minimal if and only if the system is both controllable and observable.

A commonly used measure of complexity of a rational matrix function is its McMillan degree. Heuristically, this can be thought of as the number of poles of a transfer matrix (counted with multiplicity). For proper systems, McMillan degree is equal to the order of a minimal state space model of the transfer matrix. For descriptor systems this does not hold. However, for a controllable and observable state space representation the McMillan degree, \( n \), is equal to the rank of \( E [20,\text{13}] \). We can thus bound from below the McMillan degree in terms of \( k \), the order of the system and the number of inputs or outputs:

\[
n \geq \min\{k - m, k - p\}.
\]
2.2. Selected results on the interpolation problem

In [3], the scalar rational interpolation problem was considered. The main tool employed was the Loewner matrix: assume that we are given the rational function \( y(s) \), and a set \( Z = \{z_1, \ldots, z_N\} \) of points in the complex plane. Let \( Y = \{y_1, \ldots, y_N\}, \) with \( y_i := y(z_i), i = 1, \ldots, N, \) and partition \( Z \) and \( Y \) as follows:

\[
Z = \{\lambda_1, \ldots, \lambda_\rho\} \cup \{\mu_1, \ldots, \mu_\nu\} \quad \text{and} \quad Y = \{w_1, \ldots, w_\rho\} \cup \{v_1, \ldots, v_\nu\},
\]

where \( \rho + \nu = N \). The Loewner matrix \( L \) associated to the above partitioning is

\[
L = \begin{bmatrix} v_i - w_j \\ \mu_i - \lambda_j \end{bmatrix} \in \mathbb{C}^{\nu \times \rho}.
\]

The following lemma relates the rank of the Loewner matrix to the McMillan degree of the associated rational function:

**Lemma 2.1.** Given \( y(s) \) as above of McMillan degree \( n \), let \( L \) be any \( \nu \times \rho \) Loewner matrix corresponding to the set \( Z \times Y \). If \( \min\{\nu, \rho\} \geq n \), the rank of \( L \) is \( n \).

Suppose, then, that all we are given is the interpolation data \( Z \times Y \). The following theorem relates the degree of interpolants consistent with the data, to the rank of the Loewner matrix.

**Lemma 2.2.** Given the interpolation data \( Z \times Y \), let \( L \) be an almost square Loewner matrix, that is of size \( m \times m \) if \( N = 2m \), or of size \( m \times (m+1) \) if \( N = 2m + 1 \). Assume that rank \( L = n \).

1. if all \( n \times n \) submatrices of \( L \) are non-singular, there is a unique degree \( n \) interpolant.
2. Otherwise there is a family of rational interpolants of least degree \( N - n \).

This theorem implies that there is a *jump* in permitted degrees of interpolants. As shown in [4], the Loewner matrix has a system theoretic interpretation: it is minus the product of the generalized controllability and observability matrices. In particular, suppose that the interpolation data is obtained by sampling a proper rational function \( y(s) \) with minimal state space representation \([A, B, C, D]\). That is, \( y(\lambda_i) = w_i, i = 1, \ldots, \rho \) and \( y(\mu_i) = v_i, i = 1, \ldots, \nu \). Then

\[
\mathbb{L}_{ij} = \frac{C((\mu_i I - A)^{-1} - (\lambda_j I - A)^{-1})B}{\mu_i - \lambda_j} = -C(\mu_i I - A)^{-1}(\lambda_j I - A)^{-1}B.
\]

Therefore, \( \mathbb{L} = -\mathcal{O}\mathbb{R} \), where

\[
\mathcal{O} = \begin{bmatrix} C(\mu_1 I - A)^{-1} \\ \vdots \\ C(\mu_\nu I - A)^{-1} \end{bmatrix} \in \mathbb{R}^{\nu \times n} \quad \text{and} \quad \mathbb{R} = \begin{bmatrix} (\lambda_1 I - A)^{-1}B & \cdots & (\lambda_\rho I - A)^{-1}B \end{bmatrix} \in \mathbb{R}^{n \times \rho}.
\]

\( \mathcal{O} \) is the generalized observability matrix, and \( \mathbb{R} \) the generalized controllability matrix. The rank of these matrices is well understood, as the following theorem indicates.

**Lemma 2.3** [2]. Let \((A, B)\) be a controllable pair, and \( \lambda_1, \ldots, \lambda_\rho \) scalars that are not eigenvalues of \( A \). It follows that
\[
\text{rank}[(\lambda_1 I - A)^{-1} B \cdots (\lambda_\rho I - A)^{-1} B] = \text{size}(A),
\]
provided that \(\rho \geq n\).

This result yields a proof of Lemma 2.1: a Loewner matrix constructed from a function of McMillan degree \(n\), can be factored as a product of a rank \(n\) matrix with full column rank and a rank \(n\) matrix with full row rank, and hence the Loewner matrix must have rank \(n\).

As demonstrated in [2], the above results generalize to matrix interpolation data in a straightforward manner. For Loewner matrices, the \((i, j)\)th block is
\[
L_{ij} = V_i - W_j.
\]

The generalized controllability and observability matrices for matrix data, involve block rows and block columns.

**Special case.** Suppose that we consider scalar interpolation data consisting of a single point with multiplicity, namely: \((s_0; \phi_0, \phi_1, \ldots, \phi_{N-1})\), i.e. the value of a function at \(s_0\) and that of a number of its derivatives are provided. It can be shown (see [3]) that the Loewner matrix in this case becomes
\[
L = \begin{bmatrix}
\phi_1 & \phi_2 & \phi_3 & \cdots \\
\frac{\phi_1}{s_0} & \phi_1 & \phi_3 & \cdots \\
\frac{\phi_2}{s_0} & \frac{\phi_2}{s_0} & \phi_1 & \cdots \\
\frac{\phi_3}{s_0} & \frac{\phi_3}{s_0} & \frac{\phi_3}{s_0} & \phi_1 & \cdots \\
\vdots & \vdots & \vdots & \ddots & \ddots
\end{bmatrix},
\]
and has Hankel structure. Thus the Loewner matrix generalizes the Hankel matrix when interpolation at finite points is considered. For connections between Hankel and Loewner matrices see also [10].

3. The Loewner and the shifted Loewner matrices

The Loewner matrix and the shifted Loewner matrix are both of dimension \(v \times \rho\), and are defined in terms of the data (3) and (5) as follows
\[
L = \begin{bmatrix}
\frac{v_1 V_1 - \ell_1 W_1}{\mu_1 - \lambda_1} & \cdots & \frac{v_1 V_\rho - \ell_1 W_\rho}{\mu_1 - \lambda_\rho} \\
\frac{v_2 V_1 - \ell_2 W_1}{\mu_2 - \lambda_1} & \cdots & \frac{v_2 V_\rho - \ell_2 W_\rho}{\mu_2 - \lambda_\rho} \\
\vdots & \ddots & \vdots \\
\frac{v_\nu V_1 - \ell_\nu W_1}{\mu_\nu - \lambda_1} & \cdots & \frac{v_\nu V_\rho - \ell_\nu W_\rho}{\mu_\nu - \lambda_\rho}
\end{bmatrix}, \quad \sigma L = \begin{bmatrix}
\frac{\mu_1 V_1 - \lambda_1 \ell_1 W_1}{\mu_1 - \lambda_\rho} & \cdots & \frac{\mu_1 V_\rho - \lambda_\rho \ell_1 W_\rho}{\mu_1 - \lambda_\rho} \\
\frac{\mu_2 V_1 - \lambda_1 \ell_2 W_1}{\mu_2 - \lambda_\rho} & \cdots & \frac{\mu_2 V_\rho - \lambda_\rho \ell_2 W_\rho}{\mu_2 - \lambda_\rho} \\
\vdots & \ddots & \vdots \\
\frac{\mu_\nu V_1 - \lambda_1 \ell_\nu W_1}{\mu_\nu - \lambda_\rho} & \cdots & \frac{\mu_\nu V_\rho - \lambda_\rho \ell_\nu W_\rho}{\mu_\nu - \lambda_\rho}
\end{bmatrix}.
\]

Notice that each entry shown above, for instance \(\frac{v_i V_j - \ell_i W_j}{\mu_i - \lambda_j}\), is scalar, and is obtained by taking appropriate inner products of left and right data. If we assume the existence of a rational matrix function, \(H(s)\), that generates the data then, as the name suggests, the shifted-Loewner matrix is the Loewner matrix corresponding to \(sH(s)\). These matrices satisfy the following Sylvester equations
\[
L A - M L = LW - VR, \quad \sigma L A - M \sigma L = LWA - MVR
\]
The first consequence of these equations is

**Proposition 3.1.** There holds: \( \sigma L - \mathbb{L}A = VR \) and \( \sigma L - M \mathbb{L} = LW \).

**Proof.** Multiplying the equation for \( \mathbb{L} \) by \( A \) on the right and subtracting the equation for \( \sigma \mathbb{L} \) we obtain

\[
(\sigma L - \mathbb{L}A - VR)A - M(\sigma L - \mathbb{L}A - VR) = 0.
\]

If \( A \) and \( M \) have no eigenvalues in common, the solution to this Sylvester equation is zero, which yields the desired equality. Correspondingly, if we multiply the equation for \( \mathbb{L} \) on the left by \( M \) and subtract from it the equation for \( \sigma \mathbb{L} \), by the same reasoning we obtain the second equality. \( \square \)

We can also study the Loewner matrix by assuming the existence of an underlying interpolant, say \( H(s) = C(sE - A)^{-1}B + D \). Following [2], we make use of the fact that the Loewner matrix has a system theoretically significant factorization, namely it can be factored as a product of the tangential generalized controllability and tangential generalized observability matrices. For descriptor systems, this factorization is

\[
\mathbb{L} = -YEX, \quad \text{where } Y = \begin{bmatrix}
\ell_1 C(\mu_1 E - A)^{-1} \\
\vdots \\
\ell_v C(\mu_v E - A)^{-1}
\end{bmatrix},
\]

\[
X = \begin{bmatrix}
(\lambda_1 E - A)^{-1}B_1 \\
\vdots \\
(\lambda_\rho E - A)^{-1}B_\rho
\end{bmatrix}.
\]

Likewise,

\[
\sigma \mathbb{L} = -YAX + LDR.
\]

\(Y\) and \(X\) are themselves solutions to Sylvester equations:

\[
-AX + EXA = BR, \quad \text{ (13)}
\]

\[
-YA + MYE = LC. \quad \text{ (14)}
\]

**Remark 3.1.** In [2] state space models are constructed by factoring the Loewner matrix \( \mathbb{L} \) to define \( Y \) and \( X \), and then use these equations (with \( E = I \)) to construct \( A, B, C, D \).

The following is the descriptor system version of a theorem proved in [2].

**Lemma 3.1.** Let \((E, A, B)\) be a controllable triple of order \( k \), and \( \lambda_1, \ldots, \lambda_\rho \) scalars that are not generalized eigenvalues of \((A, E)\). Then the rank of the generalized controllability matrix is \( k \). That is,

\[
\text{rank}[(\lambda_1 E - A)^{-1}B \ldots (\lambda_\rho E - A)^{-1}B] = k,
\]

provided that \( \rho \geq k \).

This is easily proved by showing that this matrix is of the same rank as the controllability matrix given in [9].

A straight forward corollary to this is a well known result for scalar rational functions, that if the dimension of the Loewner matrix exceeds the degree of the function, then the Loewner matrix
has rank equal to the degree. However, without making assumptions on $R$ (beyond the stipulation that it be full rank), the extension of this result to tangential generalized controllability matrices does not hold.

**Remark 3.2.** Eqs. (13), (14) can be solved by means of (generalized) eigenvalue problems. Let $X_1, Y_1 \in \mathbb{R}^{k \times \rho}$, $X_2, Y_2 \in \mathbb{R}^{\rho \times \rho}$, satisfy
\[
\begin{bmatrix}
A & BR \\
L & A
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2
\end{bmatrix}
= 
\begin{bmatrix}
E & I \\
I & E
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2
\end{bmatrix}
A_1,
\]
\[
\begin{bmatrix}
Y_1^* & Y_2^*
\end{bmatrix}
\begin{bmatrix}
A \\
L^*C
\end{bmatrix}
= 
\begin{bmatrix}
E & I \\
I & E
\end{bmatrix}.
\]

If we choose $A_1$ to have the same eigenvalues as $A$, and since the spectra of $(A, E)$ and $A$ are assumed disjoint, $A_1 X_2 = X_2 A_1$, implies that $X_2$ is non-singular. Using a similar argument we conclude that $Y_2$ is also non-singular. Therefore the solution of the two equations is given by $X = X_1 X_2^{-1}$, $Y = Y_2^{-1} Y_1$. Expressing the solution of Sylvester equations in terms of a generalized eigenvalue problem is of interest in large-scale applications, where multiple matrix inversions are to be avoided.

**4. Construction of interpolants when $\mathbb{L}$ is full rank**

If the Loewner matrix $\mathbb{L} \in \mathbb{C}^{v \times \rho}$ has full rank, say full row rank $v \leq \rho$, the following result gives a family of proper interpolants of dimension and McMillan degree $\rho$.

**Theorem 4.1.** Let $\mathbb{L} \in \mathbb{C}^{v \times \rho}$, be the Loewner matrix constructed using the right and left tangential interpolation data (3), (5). Assume that rank $\mathbb{L} = v$, and let $\mathbb{L}^\#$, be a right inverse of $\mathbb{L}$. The following quadruple is a realization of McMillan degree $\rho$.

\[
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix} = \begin{bmatrix}
A + \mathbb{L}^\#(V - LD)R \\
-(W - DR)
\end{bmatrix}
\begin{bmatrix}
\mathbb{L}^\#(V - LD) \\
D
\end{bmatrix},
\]

where $D \in \mathbb{R}^{\rho \times m}$, is a matrix parameter.

**Proof.** From the first equation at (11) follows:

\[
\mathbb{L}(A - sI) - (M - sI)\mathbb{L} = \mathbb{L}(W - DR) - (V - LD)R.
\]

For $s = \mu_j$, there holds $e^*_j(M - \mu_j I) = 0$. Hence

\[
e^*_j\mathbb{L}(A - \mu_j I) = e^*_jL(W - DR) - e^*_j(V - LD)R
\]

\[
\Rightarrow e^*_j\mathbb{L}(A - \mu_j I) + e^*_j(V - LD)R = \ell_j(W - DR)
\]

\[
\Rightarrow e^*_j\mathbb{L}(A - \mu_j I) + \mathbb{L}^\#(V - LD)R = \ell_j(W - DR)
\]

\[
\Rightarrow e^*_j\mathbb{L}(A - \mu_j I) = -\ell_jC
\]

\[
\Rightarrow e^*_j\mathbb{L}^\#(V - LD) = \ell_j^*C(\mu_j I - A)^{-1}B
\]

Throughout, $(\cdot)^*$, denotes transposition followed by complex conjugation, if the matrix is complex.
\[ (v_j - \ell_j D) = \ell_j C(\mu_j I - A)^{-1} B \]
\[ \ell_j [C(\mu_j I - A)^{-1} B + D] = v_j \]

which proves that the left tangential interpolation conditions are satisfied. To prove that the right tangential conditions are also satisfied we notice that the solution \( x_i \) of the equation

\[ [A - \lambda_i I + \mathbb{L}(V - LD)R]x_i = \mathbb{L}(V - LD)Re_i, \]

is \( x_i = e_i \); thus

\[ C(\lambda_i I - A)^{-1} Br_i = (W - DR)[A - \lambda_i I + \mathbb{L}(VR - LDR)]^{-1}\mathbb{L}(V - LD)r_i = w_i - Dr_i, \]

which implies the desired right tangential interpolation conditions. Notice that interpolation does not depend on \( D \). \( \square \)

It readily follows that the condition for this realization to be well defined is that none of the interpolation points be an eigenvalue of \( A \). This can also be achieved by appropriately restricting \( D \), so that the following equations are satisfied:

\[ \det[A - sI + \mathbb{L}(V - LD)R] \neq 0 \]

for all \( s = \lambda_i \) and \( s = \mu_j \). We refer to the examples section for an illustration of this issue. Furthermore, since \( A \) is diagonal, the constructed system is controllable provided that none of the rows of \( \mathbb{L}(V - LD) \) is zero; this can always be achieved by appropriate choice of \( D \); similarly, observability can be guaranteed by appropriate choice of \( D \).

Finally, it should be noticed that the poles of the constructed systems will vary. The freedom lies in choosing a generalized inverse of \( \mathbb{L} \) and \( D \). Thus for a fixed \( \mathbb{L} \), the choice of poles of the constructed system is obtained by means of constant output feedback of the system \( \begin{pmatrix} A + \mathbb{L}(VR - LDR) \\ \delta \end{pmatrix} \). Consequently, in the case \( m = p = 1 \), the eigenvalues of the constructed system are obtained by means of a rank one perturbation of the right interpolation points \( A \):

\[ A = A + \mathbb{L}(V - \delta L)R, \quad \delta \in \mathbb{R}, \]

where \( \delta = D \). It should be noticed that a similar expression for the \( A \) matrix of the interpolant appears in [12].

5. The general tangential interpolation problem

In this section we will present conditions for the solution of the general tangential interpolation problem by means of state space matrices \([E, A, B, C, D]\). The first fundamental result is as follows. Expression (16) for the transfer function below generalizes (8).

**Lemma 5.1.** Assume that \( \rho = \nu \) and let \( \det(x\mathbb{L} - \sigma \mathbb{L}) \neq 0 \), for all \( x \in \{\lambda_i\} \cup \{\mu_j\} \). Then \( E = -\mathbb{L} \), \( A = -\sigma \mathbb{L} \), \( B = V \), and \( C = W \) is a minimal realization of an interpolant of the data. That is, the function

\[ H(s) = W(\sigma \mathbb{L} - s \mathbb{L})^{-1} V, \]  

interpolates the data.
Proof. Multiplying the first equation of (12) by \( s \) and subtracting it from the second we get
\[
(\sigma L - s I)A - M(\sigma L - s I) = LW(A - s I) - (M - s I)VR.
\]
(17)
Multiplying this equation by \( e_i \) on the right and setting \( s = \lambda_i \), we obtain
\[
(\lambda_i I - M)(\sigma L - \lambda_i L)e_i = (\lambda_i I - M)Vr_i \Rightarrow (\sigma L - \lambda_i L)e_i = W(\sigma L - \lambda_i L)^{-1}Vr_i.
\]
Therefore \( w_i = H(\lambda_i)r_i \). This proves the right tangential interpolation property. To prove the left tangential interpolation property, we multiply the above equation by \( e_\ast j \) on the left and set \( s = \mu_j \):
\[
\ell_j W \Rightarrow e_\ast j V = \ell_j W(\sigma L - \mu_j L)^{-1}V + D
\]
Therefore \( v_j = \ell_j H(\mu_j) \).

We show now that the realization is controllable. Suppose that there exists \( v^* \) such that
\[
v^*[\sigma L - \lambda L V] = 0 \quad \text{for some } \lambda \in \mathbb{C}.
\]
Then \( v^*\sigma L = v^*\lambda L \) and \( v^*V = 0 \). Using Proposition 3.1 it follows that either \( \lambda \in \text{spec}(A) \) or \( v^*L = 0 \). Both of these are contradictions, so no such \( v^* \) exists and \( \text{rank} [\sigma L - \sigma L V] = k \). Similarly, one can prove that \( \text{rank} [L V] = k \), and so the realization given is controllable. Observability is proved in the same manner, so the realization is minimal. \( \square \)

The \( D \) term
Eqs. (12) can also be written by adding and subtracting the term \( LDR \), where \( D \in \mathbb{R}^{p \times m} \), is at this stage a free matrix parameter:
\[
\begin{align*}
(\sigma L - LDR)A - M(\sigma L - LDR) &= L(W - DR)A - (V - LD)R, \\
(\sigma L - LDR)A - M(\sigma L - LDR) &= L(W - DR)A - M(V - LD)R.
\end{align*}
\]
(18)
(19)
Thus, in general all realizations of degree \( k \) are parameterized as follows:
\[
H(s) = (W - DR)[(\sigma L - LDR) - s L]^{-1}(V - LD) + D
\]
For \( D = 0 \) we recover (16).

5.1. Analysis of the problem in the general case
Recall that given the data (3) and (5), the Loewner and the shifted Loewner matrices are constructed by solving the Sylvester equations (12). We assume that \( \text{rank} L = n < \nu, \rho \), that is, we assume that the Loewner matrix is not full (row or column) rank. The following result holds.

Lemma 5.2. Given are the right and left tangential interpolation data (3) and (5), respectively.

Matrices
\( Y \in \mathbb{C}^{\nu \times k}, \quad X \in \mathbb{C}^{k \times \rho}, \quad E \in \mathbb{C}^{k \times k} \text{ and } A \in \mathbb{C}^{k \times k}, \)

are constructed which satisfy the following relationships

(a) \(-YEX = \mathbb{1},\)
(b) \(-YAX = \sigma \mathbb{1},\) and
(c) \(\text{rank}(xE - A) = k, \text{ for all } x \in \{\lambda_i\} \cup \{\mu_j\}.\)

Furthermore matrices \(B\) and \(C\) are also constructed, which satisfy the relationships

(d) \(-AX + EXA = BR,\) and
(e) \(-YA + MYE = LC.\)

The quadruple \((E, A, B, C)\) is a state space representation of an interpolant.

Proof. The definitions of \(B, C\) imply, respectively \(YB = V, CX = W.\) Left (tangential) interpolation follows because the string of equalities below holds:

\[ \ell_i C(\mu_i E - A)^{-1}B = e_i^*YB = e_i^*V = v_i, \]

where the first equality holds because \(Y\) is a generalized observability matrix. A similar argument shows that the right (tangential) interpolation conditions are satisfied. \(\square\)

Thus in order to construct interpolants consistent with the data, this lemma suggests the following

**Problem 5.1.** Given are the right interpolation data

\[ \{(\lambda_i, r_i, w_i)|\lambda_i \in \mathbb{C}, r_i \in \mathbb{C}^{m \times 1}, w_i \in \mathbb{C}^{p \times 1}, i = 1, \ldots, \rho\}, \]

and the left interpolation data

\[ \{(\mu_i, \ell_i, v_i)|\mu_i \in \mathbb{C}, \ell_i \in \mathbb{C}^{1 \times p}, v_i \in \mathbb{C}^{1 \times m}, i = 1, \ldots, \nu\}. \]

We seek a positive integer \(k\) and matrices

\[ X \in \mathbb{C}^{k \times \rho}, Y \in \mathbb{C}^{\nu \times k}, E \in \mathbb{C}^{k \times k}, A \in \mathbb{C}^{k \times k}, B \in \mathbb{C}^{k \times m}, C \in \mathbb{C}^{p \times k}, \]

which satisfy conditions (a)–(e) of Lemma 5.2.

5.2. Construction of interpolants

In this section we discuss the construction of state space models of the form \([E, A, B, C]\), with \(D = 0.\) The following is the main assumption pertaining to the construction proposed:

\[
\text{rank}(x \mathbb{1} - \sigma \mathbb{1}) = \text{rank}[\mathbb{1} \quad \sigma \mathbb{1}] = \text{rank}\left[\begin{array}{c} \mathbb{1} \\ \sigma \mathbb{1} \end{array}\right] =: k, \quad x \in \{\lambda_i\} \cup \{\mu_j\}. \tag{20}
\]

If these conditions are satisfied, for some \(x \in \{\lambda_i\} \cup \{\mu_j\},\) we compute the short singular value decomposition.
\[ x \perp - \sigma \perp = Y \Sigma X, \]  
where \( \text{rank}(x \perp - \sigma \perp) = \text{rank}(\Sigma) = \text{size}(\Sigma) =: k \), \( Y \in \mathbb{C}^{\nu \times k} \) and \( X \in \mathbb{C}^{k \times \rho} \).

**Theorem 5.1.** With the quantities above, a minimal realization \([E, A, B, C]\), of an interpolant is given as follows:

\[
\begin{align*}
E &:= -Y^* L X^* \\
A &:= -Y^* \sigma L X^* \\
B &:= Y^* V \\
C &:= WX^*
\end{align*}
\]  

(22)

The proof of this theorem is a special case of the proof of Theorem 5.2 and is thus omitted.

Assumption (20) that is imposed demands that the model constructed have full rank tangential generalized controllability and observability matrices. As noted in the previous section, this condition need not hold, so there are cases in which the algorithm does not yield an interpolant when one does exist. However, for scalar rational functions, this condition must hold (see Corollary 5.1).

Notice that if the directions are not restricted to a subspace, as the amount of data increases assumption (20) will always be satisfied. Thus generically assumption (20) is always satisfied.

Before moving on, we note the following easily verifiable condition for the uniqueness of the associated rational matrix.

**Lemma 5.3.** If \( k - n = \min\{m, p\} \), the transfer matrix associated to the constructed model is the unique interpolant of McMillan degree \( n \).

**Proof.** The rank \( k \) of the singular value decomposition guarantees that no interpolant of order less than \( k \) exists. As the tangential generalized observability and controllability matrices have rank equal to the order of the system, every model of the same order and McMillan degree must be equivalent to the constructed model, hence must have the same transfer matrix. Any model of higher order violates (at least) one of the conditions

\[
\text{rank}[E \ B] = k, \quad \text{rank} \begin{bmatrix} E \\ C \end{bmatrix} = k,
\]

and hence is not minimal. \( \square \)

5.3. Rank properties of \( \perp \) and \( \sigma \perp \)

To prove that the above procedure produces an interpolant, we need the following fact.

**Proposition 5.1.** Let \( V \in \mathbb{C}^{N \times n} \). If the vector \( v \) belongs to the column span of the matrix \( V \), for any matrix \( W \in \mathbb{C}^{N \times n} \) such that \( W^* V = I_n \), \( VW^* v = v \).

The first result which helps explain why assumption (20) has been imposed is

**Lemma 5.4.** Let \([E, A, B, C]\) be a minimal state space representation of order \( k \). Given also are \( L, R, A, M \), such that the associated tangential generalized observability matrix \( Y \) has full
column rank $k$, and the corresponding tangential generalized controllability matrix $X$ has full row rank $k$; none of the interpolation points are poles, and $\rho, \nu \geq k$. Then (20) holds.

**Proof.** Clearly, $\text{rank}(xL - \sigma L) = k$. Notice that

\[
\begin{bmatrix}
L & \sigma L
\end{bmatrix} = -Y \begin{bmatrix} EX & AX \end{bmatrix} = -Y \begin{bmatrix} E & A \end{bmatrix} \begin{bmatrix} X & X \end{bmatrix},
\]

where $Y$ has by assumption full column rank; we claim that $[EX \ AX]$ has full row rank, for if we assume that there exists a row vector of appropriate dimension $z$ which lies in the left kernel of this matrix, then $zEX = 0$ and $zAX = 0$, which implies that $z(xE - A)X = 0$, for all $x \in \left\{ \lambda_i \right\} \cup \left\{ \mu_j \right\}$; but since $(xE - A)$ is invertible and $X$ has full row rank, $z$ must be zero, which proves that $[EX \ AX]$ has full row rank $k$. Consequently the product of a full column rank matrix, and a full row rank matrix, has the required property $\text{rank}(\begin{bmatrix} L & \sigma L \end{bmatrix}) = k$. The second equality can be proven similarly. □

This lemma implies that in the scalar case condition (20) is necessary and sufficient for the existence of interpolants.

**Corollary 5.1.** Given the scalar data $L, R, V, W$, an interpolant $E, A, B, C$ of order $k \leq \rho, \nu$, exists if, and only if, assumption (20) holds.

**Proof.** The sufficiency of (20) follows from Theorem 5.1. Its necessity follows from Lemmas 5.4 and 3.1. □

**Lemma 5.5.** Let $L, \sigma L$, be the Loewner and the shifted Loewner matrices satisfying (20) for some set of $\rho + \nu$ points. The left and right kernels of $xL - \sigma L$ are the same for each value of $x$.

**Proof.** We will prove this only for the right kernel. First, it is clear that $\text{ker}\left( \begin{bmatrix} L & \sigma L \end{bmatrix} \right) \subseteq \text{ker}(xL - \sigma L)$. As the dimensions of the kernels are the same, equality must hold. This is true for each value of $x$, and so, $\text{ker}(xL - \sigma L)$ does not vary for the different values of $x$. □

**Corollary 5.2.** Let $L, \sigma L$ be two matrices satisfying (20) for some set of $\rho + \nu$ points. Then \[
\text{colspan}(L), \text{colspan}(\sigma L) \subseteq \text{colspan}(xL - \sigma L).
\]

**Corollary 5.3.** Recall the short SVD of $xL - \sigma L$, given by (21). Then $YY^*\sigma L X^*X = \sigma L$, $YY^*LX^*X = L$, $YY^*V = V$, and $WX^*X = W$.

This follows directly from the Proposition 5.1 and the fact, implied by Proposition 3.1, that $\text{colspan} V \subseteq \text{colspan} Y$, and $\text{rowspan} W \subseteq \text{rowspan} X$.

**5.4. Taking advantage of the $D$ term**

Recall that Eqs. (19) are satisfied for all $D$. In this case the shifted Loewner matrix is replaced by $\sigma L - LDR$ while $V, W$ are replaced by $V - LD, W - DR$. Assumption (20) becomes in this case...
\[
\text{rank}(x \mathbb{L} - \sigma \mathbb{L} + \text{LDR}) = \text{rank}\begin{bmatrix}
\mathbb{L} & \sigma \mathbb{L} - \text{LDR}
\end{bmatrix} = \text{rank}\begin{bmatrix}
\sigma \mathbb{L} - \text{LDR}
\end{bmatrix} = k, \quad x \in \{\lambda_i\} \cup \{\mu_j\},
\] (23)

where this relationship is satisfied for some \(D \in \mathbb{C}^{p \times m}\). We now restate the construction procedure.

**Theorem 5.2.** For any \(x \in \{\lambda_i\} \cup \{\mu_j\}\), take the short singular value decomposition of
\[
x \mathbb{L} - \sigma \mathbb{L} + \text{LDR} = Y \Sigma X,
\]
where \(\text{rank}(x \mathbb{L} - \sigma \mathbb{L} + \text{LDR}) = \Sigma = \text{size} = k\), and \(Y \in \mathbb{C}^{v \times k}\) and \(X \in \mathbb{C}^{k \times \rho}\). Define
\[
\begin{align*}
E &:= -Y^*X^*, \\
A &:= -Y^*(\sigma \mathbb{L} - \text{LDR})X^*, \\
B &:= Y^*(V - \text{LD}), \\
C &:= (W - \text{DR})X^*.
\end{align*}
\] (24)

Then \([E, A, B, C, D]\), is a desired realization.

**Proof.** First, note that
\[
-AX + EXA = Y^*(\sigma \mathbb{L} - \text{LDR})X^* X - Y^*X^*X A
= Y^*(\sigma \mathbb{L} - \text{LDR} - \mathbb{L} A)
= Y^*(V - \text{LD})R = BR
\]
where the second equality follows from Proposition 5.1, by noting that \((\sigma \mathbb{L} - \text{LDR})XX^* = \mathbb{L} - \text{LDR}\) and \(\mathbb{L}XX^* = \mathbb{L}\).

Likewise, \(-YA + MYE = LC\). Thus \(X\) and \(Y\) are, respectively, the generalized controllability and observability matrices for the system \([E, A, B, C, D]\).

To prove that the left interpolation conditions are met we do the following calculation:
\[
\ell_i C(\mu_i E - A)^{-1} B + \ell_i D = e_i^* Y B + \ell_i D
= e_i^* Y Y^*(V - \text{LD}) + \ell_i D
= e_i^* (V - \text{LD}) + \ell_i D = v_i,
\]
where the first equality holds because \(Y\) is the generalized observability matrix, and the second one holds because \(\text{colspan}(V - \text{LD}) \in \text{colspan}(x \mathbb{L} - \sigma \mathbb{L})\). The same reasoning holds for the right interpolation conditions.

The proof of minimality follows that given for the analogous statement in Lemma 5.1. □

### 6. Tangential interpolation with derivative constraints

We will now show that the approach to interpolation given in this paper can also be applied to the problem of interpolation with derivative constraints. For the sake of brevity we will outline the necessary changes. Thus, we will prove up to the analogue of Lemma 5.1, after which things proceed as in the simple interpolation case. We distinguish between the cases in which the left and right interpolation points have trivial intersection, and when they do not. The former case will be referred to as two-sided tangential interpolation, and the latter case as bi-tangential interpolation.
6.1. Two-sided tangential interpolation

For ease of exposition we will assume that there is only a single interpolation point for the right data. The right interpolation data now takes the following form:

\[
(\lambda, r_i, W_i) | \lambda \in \mathbb{C}, r_i \in \mathbb{C}^{m \times 1}, W_i \in \mathbb{C}^{p \times 1}, i = 1, \ldots, m).
\]

(25)

We represent the \(j\)th column of \(W_i\) by \( \frac{1}{(j-1)!} w_i^{j-1} \) i.e.

\[
W_i = \begin{bmatrix}
  w_i^0 & \cdots & \frac{1}{(\rho_i - 1)!} w_i^{\rho_i - 1}
\end{bmatrix}.
\]

We wish then to construct a state space representation \([E, A, B, C, D]\) such that the associated transfer matrix \(H(s)\) satisfies

\[
d^k X_{\hat{i}+k} = (-1)^k ((\lambda E - A)^{-1} E)^k (\lambda E - A)^{-1} B r_i, \quad k = 0, \ldots, \rho_i - 1.
\]

Lemma 6.1. Let \(X\) be an appropriately sized matrix such that

\[
-AX + EXA = BR,
\]

where \(sE - A\) is a regular pencil that is invertible on the spectra of \(A\) with \(A\) and \(R\) defined as above. Then

\[
X_{\hat{i}+k} = (-1)^k ((\lambda E - A)^{-1} E)^k (\lambda E - A)^{-1} B r_i, \quad k = 0, \ldots, \rho_i - 1.
\]

Proof. We prove this by induction. For \(k = 0\), we do the following calculation:

\[
B r_i = -AX e_i + EX \lambda e_i = (\lambda E - A) X e_i.
\]

Suppose then that the lemma is true for \(k = l\), thus

\[
X_{\hat{i}+l} = (-1)^l ((\lambda E - A)^{-1} E)^l (\lambda E - A)^{-1} B r_i.
\]
We show that it is also true for \( k = l + 1 \).

\[
0 = B_\text{Re}_{i+l+1} = -A_\text{Xe}_{i+l+1} + E_\text{Xe}_{i+l+1}
\]

\[
= -A_\text{Xe}_{i+l+1} + \lambda E_\text{Xe}_{i+l+1} - E_\text{Xe}_{i+l}.
\]

Therefore,

\[
\text{Xe}_{i+l+1} = -(\lambda E - A)^{-1}E_\text{Xe}_{i+l}
\]

\[
= -(\lambda E - A)^{-1}E(-1)^{l}(\lambda E - A)^{-1}E(-1)^{l}(\lambda E - A)^{-1}B_\text{Re}_{i}.
\]

\[\square\]

**Lemma 6.2.** Assume that \( \bar{\rho} = \bar{\nu} \) and let \( \det(xL - \sigma L) = 0 \), for \( x \in \{\lambda, \mu\} \). Then \( E = -L, A = -\sigma L, B = V, \) and \( C = W \) is a state space representation of an interpolant of the data. That is, the function \( H(s) = W(\sigma L - s L)^{-1}V \) interpolates the data.

**Proof.** First, note that \( \frac{d^k}{ds^k} (sE - A)^{-1} = (-1)^kk!(sE - A)^{-1}E(sE - A)^{-1} \). Using the relations given in Proposition 3.1, we see that \( X = I_{\bar{\rho}} \) satisfies the previous lemma. Then

\[
H^{(k)}(\lambda)_{i} = W(-1)^{k}k!(\sigma L - \lambda L)^{-1}(-\lambda)^{k}(\sigma L - \lambda L)^{-1}V_{r_{i}}
\]

\[
= k!W(-1)^{k}(\sigma L - \lambda L)^{-1}(-\lambda)^{k}(\sigma L - \lambda L)^{-1}V_{\text{Re}_{i+k}}
\]

\[
= k!W_{e_{i+k}} = k!\frac{1}{k!}w_{i}^{k} = w_{i}^{k}.
\]

The proof of the left interpolation conditions follows analogously. \[\square\]

### 6.2. Bi-tangential interpolation

For ease of exposition, we will consider interpolation at a single interpolation point \( \lambda \). As we are restricting to a single interpolation point, we can without loss of generality restrict to \( p \) left directions and \( m \) right directions. Other than that \( L, R, V, W, A \) and \( M \) are defined as in the previous case. Clearly, \( M \) and \( A \) no longer have disjoint spectra, and so Eqs. (12) no longer have unique solutions. We would like to specify the undetermined entries in \( L \) and \( \sigma L \) such that relationships analogous to those given by Proposition 3.1 hold. We determine the value of these entries as follows: let \( H(s) \) be a rational matrix function that matches the given data i.e.

\[
\frac{d^k}{ds^k} H(s)_{r_i} \bigg|_{s=\lambda} = w_{i}^{k}, \quad i \in \{1, \ldots, m\}, \quad k \in \{0, \ldots, \rho_i - 1\},
\]

\[
\ell_i \frac{d^k}{ds^k} H(s) \bigg|_{s=\lambda} = v_{i}^{l}, \quad i \in \{1, \ldots, p\}, \quad k \in \{0, \ldots, \nu_i - 1\}.
\]

Define then

\[
W_{i}(\epsilon) := \left[ H(\lambda + \epsilon)_{r_i} \quad H^{(1)}(\lambda + \epsilon)_{r_i} \quad \frac{1}{2!}H^{(2)}(\lambda + \epsilon)_{r_i} \quad \cdots \quad \frac{1}{(\rho_i - 1)!}H^{(\rho_i - 1)}(\lambda + \epsilon)_{r_i} \right]
\]
and
\[ A_\epsilon := A + \epsilon I. \]

For any \( \epsilon \neq 0 \), if we replace \( W \) by \( W_\epsilon \) and \( A \) by \( A_\epsilon \), Eqs. (12) now have unique solutions that we denote by \( L_\epsilon \) and \( \sigma L_\epsilon \). We are then interested in these in the limit as \( \epsilon \) tends to zero. As before, let
\[
\hat{i} := \sum_{j=1}^{i-1} \rho_j + 1 \quad \text{and} \quad \hat{j} = \sum_{i=1}^{j-1} v_i + 1.
\]

**Lemma 6.3.** Let \( L := \lim_{\epsilon \to 0} L_\epsilon \), and \( \sigma L := \lim_{\epsilon \to 0} \sigma L_\epsilon \). Then
\[
L_{(2i+u)j+v} = \frac{1}{(u + v + 1)!} \ell_i H^{(u+v+1)}(\lambda) r_j, \quad \text{and}
\]
\[
\sigma L_{(2i+u)j+v} = \frac{1}{(u + v + 1)!} \left[ \lambda \ell_i H^{(u+v+1)}(\lambda) r_j + (u + v + 1) \ell_i H^{(u+v)}(\lambda) r_j \right],
\]
for \( u = 0, \ldots, v_i - 1, \ v = 0, \ldots, \rho_j - 1, \ i = 1, \ldots, p, \ j = 1, \ldots, m. \)

These lemmas are easily proved using Taylor series. They indicate that the additional constraints required are bi-tangential constraints on higher order derivatives of the desired function. For example, in the \( W \) data matrix, we are given only up to the \( \rho_i \)th derivative in the direction \( r_i \); likewise, in the \( V \) data matrix, we are given only up to the \( v_j \)th derivative in the direction \( \ell_j \). However, we see that we require a term involving a derivative of order \( \rho_i + v_j - 1 \). Therefore, we will assume that we are also given data that stipulates the values for the following quantities:
\[
\ell_p H^{(q)}(\lambda) r_v \quad \text{for} \ u = 1, \ldots, p, \ v = 1, \ldots, m, \ q = \max\{\rho_v, v_u\}, \ldots, \rho_v + v_u - 1.
\]

(27)

**Remark 6.1.** Suppose that we sample a rational matrix function \( H(s) \) and appropriate derivatives thereof at \( \lambda \in \mathbb{C} \) along left directions \( \ell_1, \ldots, \ell_p \) and right directions \( r_1, \ldots, r_m \) Let \( \hat{L}_k \) be a matrix whose rows are those directions \( \ell_i \) such that \( v_i \geq k \). Let also \( v := \max\{v_1, \ldots, v_p\} \). Define \( \hat{R}_k \) and \( \rho \) analogously. Then we can find permutation matrices \( P \) and \( Q \) such that
\[
P L Q = \begin{bmatrix}
\hat{L}_1 H^{(1)}(\lambda) \hat{R}_1 \\
\frac{1}{2!} \hat{L}_2 H^{(2)}(\lambda) \hat{R}_2 \\
\vdots \\
\frac{1}{v!} \hat{L}_v H^{(v)}(\lambda) \hat{R}_v \\
\end{bmatrix} \\
\frac{1}{2} \hat{L}_1 H^{(2)}(\lambda) \hat{R}_1 \\
\frac{1}{3} \hat{L}_2 H^{(3)}(\lambda) \hat{R}_2 \\
\vdots \\
\frac{1}{(v+1)!} \hat{L}_v H^{(v+1)}(\lambda) \hat{R}_v \\
\end{bmatrix} \\
\frac{1}{(p+1)!} \hat{L}_2 H^{(p+1)}(\lambda) \hat{R}_p \\
\frac{1}{(p+2)!} \hat{L}_2 H^{(p+2)}(\lambda) \hat{R}_p \\
\end{bmatrix},
\]

(28)

and
\[
P \sigma L Q = \lambda P L Q + 
\begin{bmatrix}
\hat{L}_1 H^{(0)}(\lambda) \hat{R}_1 \\
\hat{L}_2 H^{(1)}(\lambda) \hat{R}_1 \\
\vdots \\
\frac{1}{(v-1)!} \hat{L}_v H^{(v-1)}(\lambda) \hat{R}_1 \\
\end{bmatrix} \\
\frac{1}{2!} \hat{L}_1 H^{(2)}(\lambda) \hat{R}_2 \\
\frac{1}{3!} \hat{L}_2 H^{(3)}(\lambda) \hat{R}_2 \\
\vdots \\
\frac{1}{(v-1)!} \hat{L}_v H^{(v-1)}(\lambda) \hat{R}_2 \\
\end{bmatrix} \\
\frac{1}{(p-1)!} \hat{L}_1 H^{(p-1)}(\lambda) \hat{R}_p \\
\frac{1}{(p-2)!} \hat{L}_2 H^{(p-2)}(\lambda) \hat{R}_p \\
\end{bmatrix},
\]

(29)
Notice that if \( v = v_1 = \cdots = v_p \) and \( \rho = \rho_1 = \cdots = \rho_m \), then these matrices have block Hankel structure. Furthermore, if the directions are the canonical basis, then these expressions are those of the Loewner and shifted Loewner matrices for MIMO interpolation with derivative constraint.

**Proposition 6.1.** \( \mathcal{L} \) and \( \mathcal{\sigma} \mathcal{L} \) satisfy \( \mathcal{\sigma} \mathcal{L} - \mathcal{L} A = \mathbf{V} \mathbf{R} \) and \( \mathcal{\mathcal{\sigma} \mathcal{L}} - \mathcal{M} \mathcal{L} = \mathbf{W} \), where

\[
A = \text{diag}[\mathbf{J}_{\rho_1}(t), \ldots, \mathbf{J}_{\rho_p}(t)], \quad M = \text{diag}[\mathbf{J}_{v_1}^*(t), \ldots, \mathbf{J}_{v_v}(t)],
\]

and \( \mathbf{W}, \mathbf{V}, \mathcal{L} \) and \( \mathcal{\mathcal{\sigma} \mathcal{L}} \) are as above.

**Proof.** We prove that \( \mathcal{\sigma} \mathcal{L} - \mathcal{L} A = \mathbf{V} \mathbf{R} \). Note that if we multiply the left hand side by \( \epsilon_j \mathcal{\sigma} \mathcal{L} + \mathcal{L} \mathcal{\sigma} \mathcal{L} \) the definitions of \( \mathcal{L} \) and \( \mathcal{\sigma} \mathcal{L} \) imply that we get 0 unless \( v = 0 \). Clearly, \( \Re \mathbf{e}_j = 0 \) if \( v \neq 0 \), so we just have to check the equality when \( v = 0 \).

\[
(\mathcal{\sigma} \mathcal{L} - \mathcal{L} A) \mathbf{e}_j = \begin{bmatrix}
\frac{1}{0!} \ell_i \mathbf{H}^{(0)} \mathbf{r}_j \\
\frac{1}{(1)!} \ell_i \mathbf{H}^{(1)} \mathbf{r}_j \\
\vdots \\
\frac{1}{(v_1 - 1)!} \ell_i \mathbf{H}^{(v_1 - 1)} \mathbf{r}_j \\
\vdots \\
\frac{1}{(v + v_p - 1)!} \ell_i \mathbf{H}^{(v_p - 1)} \mathbf{r}_j
\end{bmatrix} = \mathbf{V} \mathbf{r}_j = \mathbf{V} \Re \mathbf{e}_j. \quad \square
\]

We conclude by proving the analogue of Lemma 5.1.

**Lemma 6.4.** Assume that \( \tilde{\rho} = \tilde{v} \) and let \( \det(\lambda \mathcal{L} - \mathcal{\sigma} \mathcal{L}) \neq 0 \). Then

\[
\mathbf{E} = -\mathcal{L}, \quad \mathbf{A} = -\mathcal{\sigma} \mathcal{L}, \quad \mathbf{B} = \mathbf{V}, \quad \mathbf{C} = \mathbf{W},
\]

is a minimal realization of an interpolant of the data. That is, the function

\[
\mathbf{H}(s) = \mathbf{W}(\mathcal{\sigma} \mathcal{L} - s \mathcal{L})^{-1} \mathbf{V}
\]

interpolates the data.

**Proof.** By the same reasoning as in the previous section, \( \mathbf{H}(s) \) satisfies the two-sided tangential interpolation conditions i.e. \( \mathbf{H}^{(k)}(\lambda) \mathbf{r}_j = \mathbf{w}_i^k, k \leq \rho_i, i = 1, \ldots, m, \) and \( \ell_i \mathbf{H}^{(k)}(\lambda) = \mathbf{v}_j^k, k \leq v_i, i = 1, \ldots, p. \) As before, \( \frac{d^k}{ds^k} (s \mathbf{E} - \mathbf{A})^{-1} = (-1)^k k! (s \mathbf{E} - \mathbf{A})^{-1} \mathbf{E}^k (s \mathbf{E} - \mathbf{A})^{-1}. \) Choose \( a, b \) such that \( a + b = k \) and \( a \leq \rho_v - 1, b \leq v_a - 1. \) Then

\[
\frac{d^k}{ds^k} (s \mathbf{E} - \mathbf{A})^{-1} = (-1)^k k! (s \mathbf{E} - \mathbf{A})^{-1} [E(s \mathbf{E} - \mathbf{A})^{-1}]^{a-1} E(E(s \mathbf{E} - \mathbf{A})^{-1} \mathbf{E}^b (s \mathbf{E} - \mathbf{A})^{-1}. \]

Then

\[
\ell_i \mathbf{H}^{(k)}(\lambda) \mathbf{r}_j = \ell_i [(-1)^k k! (\mathcal{\sigma} \mathcal{L} - \mathcal{\lambda} \mathcal{L})^{-1} ((-\mathcal{L})(\mathcal{\sigma} \mathcal{L} - \mathcal{\lambda} \mathcal{L})^{-1})^{a-1} \\
\times ((-\mathcal{L})(\mathcal{\sigma} \mathcal{L} - \mathcal{\lambda} \mathcal{L})^{-1} (-\mathcal{L}))^{b} (\mathcal{\sigma} \mathcal{L} - \mathcal{\lambda} \mathcal{L})^{-1}] \mathbf{r}_j.
\]

On the right, using Proposition 6.1 we see that \( \mathbf{X} = \mathbf{I}_\tilde{\rho} \) satisfies Lemma 6.1. Thus

\[
(-1)^q ((\mathcal{\sigma} \mathcal{L} - \mathcal{\lambda} \mathcal{L})^{-1} (-\mathcal{L}))^q (\mathcal{\sigma} \mathcal{L} - \mathcal{\lambda} \mathcal{L})^{-1} \mathbf{V} \Re i^{i+q} = \mathbf{e}_{i+q}.
\]
Likewise, on the left, we have
\[ e^*_i p = \left( (\sigma - \lambda L) - (\sigma - \lambda R) \right)^{p-1} e^*_i p. \]

Thus
\[ \ell_i H^{(k)}(\lambda) r_j = k! \frac{1}{(p+q)!} H_{ij}^{(p+q)}(\lambda) = H_{ij}^{(p+q)}(\lambda). \]

**Remark 6.2.** It is worth noting at this point that for the bi-tangential interpolation problem involving one point, assumption (20), simplifies as follows. It is assumed here for simplicity, that the number of left and right directions remain constant; this however is not crucial for the validity of the arguments and the corresponding procedures.

Let \( \eta_k = \frac{1}{k!} H^{(k)}(\lambda) R \in \mathbb{R}^{\nu \times \rho} \), and define the block Hankel matrix
\[
\mathcal{L} = \begin{bmatrix}
\eta_0 & \cdots & \eta_{p-1} & \eta_p \\
\eta_1 & \cdots & \eta_p & \eta_{p+1} \\
\vdots & \ddots & \vdots & \vdots \\
\eta_{v-1} & \cdots & \eta_{v+p-2} & \eta_{v+p-1} \\
\eta_v & \cdots & \eta_{v+p-1} & \eta_{v+p}
\end{bmatrix},
\]

be a matrix composed of \((v+1)\) block rows and \((\rho+1)\) block columns. Then (20) reduces to the \((v+1) \times \rho\) and the \(v \times (\rho+1)\), principal block sub-matrices of \( \mathcal{L} \) having the same rank. If \( \lambda = 0 \), then the minimal realization has dimension equal to the rank of the \( v \times \rho \) principal block submatrix of \( \mathcal{L} \) (i.e. of \( \sigma I \)), while if \( \lambda = \infty \), the minimal dimension is the rank of \( I \) (i.e. the submatrix of \( \mathcal{L} \) composed of block rows 2, \ldots, \( v \) and block columns 2, \ldots, \( \rho \) (which generalizes the well known result for scalar and matrix realization)).

**Remark 6.3.** As discussed above, in order to be able to construct generalized state space realizations from bi-tangential interpolation data, we also require that the values of additional derivatives of the underlying function be available, see e.g. (27). Instead of requiring additional values however, we can regard these as free parameters and try and construct an appropriate (e.g. a minimal order) interpolant over all possible values of these parameters. Here is a simple example illustrating this point. We are sampling at \( s = \lambda \) the \( 2 \times 2 \) rational function \( H(s) \) with entries \( h_{ij}^{(1)} \), their first derivatives at the same point being denoted by \( h_{ij}^{(1,1)} \); the value of this function is sampled along the right directions \([e_1, e_2]\), while the first derivative at \( s = \lambda \) is sampled along the direction \( e_1 \); similarly the left directions are \([e^*_1, e^*_2], e^*_1\), for the value and first derivative, respectively. In the absence of additional information about the derivatives of \( H \) at \( s = \lambda \), the ensuing Loewner and shifted Loewner matrices contain unknown entries:

\[
\sigma R = \begin{bmatrix}
h_{11}^{(1)} & h_{12}^{(1)} & h_{11}^{(1)} \\
h_{21}^{(1)} & h_{22}^{(1)} & h_{21}^{(1)} \\
h_{11}^{(1)} & h_{12}^{(1)} & \alpha
\end{bmatrix}, \quad R = \begin{bmatrix}
h_{11}^{(1)} & h_{12}^{(1)} & \alpha \\
h_{21}^{(1)} & \beta & \gamma \\
\alpha & \delta & \epsilon
\end{bmatrix}.
\]

The unknowns \( \alpha, \beta, \gamma, \delta, \) and \( \epsilon \) are to be determined so that assumption (20) is satisfied, for instance, with the smallest possible rank. This problem is considered in [14].
7. Parametrization of interpolants

For the scalar case, we will show that the parameter $\mathbf{D}$ denoted by $\delta$ here, parameterizes all interpolants of the given degree. Note that in the scalar case, $\mathbf{L}$ is taken to be the column vector of appropriate length with ones in all of its entries. Likewise, $\mathbf{R}$ is the row vector with ones in all of its entries. When specifying an interpolation condition, we will therefore not include the direction. Recall that in the scalar case there is a family of interpolants if, and only if, the Loewner matrix is square ($\rho = \nu$) and non-singular. Therefore this assumption will hold for this section. Noting that in this case $s\perp - \sigma \perp + \delta \mathbf{LR}$ is a rank one perturbation of $s\perp - \sigma \perp$, the following holds.

**Lemma 7.1.** With $\Phi = (\sigma \perp - s \perp)^{-1}$, there holds:

$$
(W - \delta \mathbf{R})(\Phi^{-1} - \delta \mathbf{LR})^{-1}(V - \delta \mathbf{L}) + \delta = \delta \frac{(1 - \mathbf{W}\Phi\mathbf{L})(1 - \mathbf{R}\Phi\mathbf{V})}{(1 - \mathbf{R}\Phi\mathbf{L})} + \mathbf{W}\Phi\mathbf{V}. \tag{30}
$$

It is hereby assumed that the pencil $s\perp - \sigma \perp$ is regular.

For the proof we will make use of the Sherman–Morrison–Woodbury formula:

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B}(\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1},$$

which in the case that $\mathbf{B} = \mathbf{b}$ is a single column, $\mathbf{C} = \mathbf{c}$ a single row, and consequently $\mathbf{D} = \delta$, is a scalar, becomes

$$(\mathbf{A} + \delta\mathbf{bc})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{bc}\mathbf{A}^{-1}}{\delta^{-1} + \mathbf{c}\mathbf{A}^{-1}\mathbf{b}} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{bc}\mathbf{A}^{-1}}{\delta^{-1} + \mathbf{c}\mathbf{A}^{-1}\mathbf{b}}. \tag{31}
$$

**Proof.** We have

$$
(W - \delta \mathbf{R})(\Phi^{-1} + (-\delta)\mathbf{LR})^{-1}(V - \delta \mathbf{L}) + \delta = \underbrace{(W - \delta \mathbf{R})\Phi(V - \delta \mathbf{L})}_{a} \underbrace{+ \delta \frac{(W - \delta \mathbf{R})\Phi\mathbf{L}\Phi(V - \delta \mathbf{L})}{(1 - \delta \mathbf{R}\Phi\mathbf{L})}}_{b} + \delta,
$$

$$a = \mathbf{W}\Phi\mathbf{V} - \delta \mathbf{R}\Phi\mathbf{V} - \delta \mathbf{W}\Phi\mathbf{L} + \delta^{2}\mathbf{R}\Phi\mathbf{L},
$$

$$b = (1 - \delta \mathbf{R}\Phi\mathbf{L}) \cdot \mathbf{b} = \delta \mathbf{W}\Phi\mathbf{L}\Phi\mathbf{V} - \delta^{2}\mathbf{R}\Phi\mathbf{L}\Phi\mathbf{V} - \delta^{2}\mathbf{W}\Phi\mathbf{L}\Phi\mathbf{V} + \delta^{3}\mathbf{R}\Phi\mathbf{L}\Phi\mathbf{V}, \tag{31}
$$

$$
(1 - \delta \mathbf{R}\Phi\mathbf{L}) \cdot a = (1 - \delta \mathbf{R}\Phi\mathbf{L}) \cdot (\mathbf{W}\Phi\mathbf{V} - \delta \mathbf{R}\Phi\mathbf{V} - \delta \mathbf{W}\Phi\mathbf{L} + \delta^{2}\mathbf{R}\Phi\mathbf{L})

= \mathbf{W}\Phi\mathbf{V} - \delta \mathbf{R}\Phi\mathbf{V} - \delta \mathbf{W}\Phi\mathbf{L} + \delta^{2}\mathbf{R}\Phi\mathbf{L}

- \delta \mathbf{R}\Phi\mathbf{L} \cdot (\mathbf{W}\Phi\mathbf{V} - \delta \mathbf{R}\Phi\mathbf{V} - \delta \mathbf{W}\Phi\mathbf{L} + \delta^{2}\mathbf{R}\Phi\mathbf{L})

= \mathbf{W}\Phi\mathbf{V} - \delta \mathbf{R}\Phi\mathbf{V} - \delta \mathbf{W}\Phi\mathbf{L} + \delta^{2}\mathbf{R}\Phi\mathbf{L}

- \delta \mathbf{R}\Phi\mathbf{L}\mathbf{W}\Phi\mathbf{V} + \delta^{2}\mathbf{R}\Phi\mathbf{L}\Phi\mathbf{V} + \delta^{2}\mathbf{R}\Phi\mathbf{L}\mathbf{W}\Phi\mathbf{V} - \delta^{3}\mathbf{R}\Phi\mathbf{L}\Phi\mathbf{V}, \tag{32}
$$

$$\delta(1 - \delta \mathbf{R}\Phi\mathbf{L}) = \delta - \delta^{2}\mathbf{R}\Phi\mathbf{L}. \tag{33}
$$
Thus (31) + (32) + (33) yield
\[ \begin{align*}
\{ W\Phi V - \delta R\Phi V - \delta W\Phi L + \delta^2 R\Phi L - \delta R\Phi LW\Phi V \\
+ \delta^2 R\Phi LR\Phi V + \delta^2 R\Phi LW\Phi L - \delta^3 R\Phi LR\Phi L \\
+ [\delta W\Phi LR\Phi V - \delta^2 R\Phi LR\Phi V - \delta^2 W\Phi LR\Phi L + \delta^3 R\Phi LR\Phi L] + [\delta - \delta^2 R\Phi L].
\end{align*} \]

All terms in $\delta^2$ and $\delta^3$ cancel and thus
\[ W\Phi V - \delta R\Phi V - \delta W\Phi L + \delta W\Phi LR\Phi V + \delta = W\Phi V(1 - \delta R\Phi L) + \delta(1 - W\Phi L)(1 - R\Phi V). \]

This proves (30). \qed

7.1. Properties of the quantities involved

The following properties hold.

**Proposition 7.1**

- $W\Phi V$ generically interpolates the pairs $(\lambda_i, w_i)$ and $(\mu_j, v_j)$.
- $W\Phi L$ generically interpolates the pairs $(\mu_j, 1)$. Furthermore, if the triple $(-\|, -\|, L)$ is not controllable, then $W\Phi L \equiv 1$.
- $R\Phi V$ generically interpolates the pairs $(\lambda_i, 1)$. Similarly, if the triple $(R, -\|, -\|)$ is not observable we have $R\Phi V \equiv 1$.

In either of these cases, the set of interpolants parameterized by $\delta$ as above consists of a single element, namely $W\Phi V$.

**Proof.** The generic case is that the finite eigenvalues of the pair $(\sigma \|, \|)$ do not coincide with the interpolation points. The validity of the first property was shown earlier. To prove the second one we note that $e^j LW\Phi L = W\Phi L$, and then substitute $\sigma \| - M\|$ for $LW$ using Proposition 3.1.

A similar calculation yields the third condition. If the triple $(-\|, -\|, L)$ is not controllable, then the rational function $W\Phi L$ has pole-zero cancellation. By assumption the zeros of the common factor do not occur at the interpolation points, so the interpolation conditions are still met. Therefore, the rational function $W\Phi L - 1$ of degree strictly less than $\rho$, has $\rho$ zeros, hence vanishes identically. Therefore $W\Phi L \equiv 1$. The proof is analogous in the case that $(R, \|, \sigma \|)$ is not observable. It is clear that for every value of $\delta$ the expression on the right of (30) reduces to $W\Phi V$. \qed

7.2. Connection with the generating system framework

In [5] it was shown that (in the scalar case) given an interpolant of the data $p(s)/q(s), p(s), q(s) \in C(s)$, and a second rational function $\hat{p}(s)/\hat{q}(s), \hat{p}(s), \hat{q}(s) \in C(s)$, such that the roots of the determinant of the rational matrix $\Theta = \begin{bmatrix} p & \hat{p} \\ q & \hat{q} \end{bmatrix}$ are precisely the interpolation points, any interpolant $\frac{a(s)}{b(s)}$, of the data can be expressed as $\begin{bmatrix} n(s) \\ d(s) \end{bmatrix} = \Theta(s) \begin{bmatrix} a(s) \\ b(s) \end{bmatrix}$, for polynomials $a, b$, of appropriate degree. Therefore $\Theta(s)$ is called the generating matrix for the given interpolation problem. In particular
If \( \Theta \) happens to be column reduced with same column indices, say \( k \), then all interpolants of McMillan degree \( k \), are obtained by choosing \( a \) and \( b \) constant.

We will now show that the expression in Lemma 7.1 yields indeed a generating matrix. Eq. (30) can be written as a linear fraction:

\[
\frac{W \Phi V + \delta[(1 - W \Phi L)(1 - R \Phi V) - (W \Phi V)(R \Phi L)]}{1 - \delta R \Phi L}
\]

There are two generating systems involved in the above relationship:

\[
\Theta_1(s) = \begin{bmatrix} W \Phi V & (1 - W \Phi L)(1 - R \Phi V) \\ 1 & 0 \end{bmatrix},
\]

\[
\Theta_2(s) = \begin{bmatrix} W \Phi V & (W \Phi V)(R \Phi L) - (1 - W \Phi L)(1 - R \Phi V) \\ 1 & R \Phi L \end{bmatrix},
\]

which are both rational. Notice that the determinant of these generating systems has zeros at the interpolating points \( \lambda_i \) and \( \mu_j \). Clearly

\[
\Theta_2(s) = \Theta_1(s) \begin{bmatrix} 1 & R \Phi L \\ 0 & -1 \end{bmatrix}.
\]

This implies that the quotient of the entries of the second column of \( \Theta_2 \) yields an interpolant as well, namely

\[
\frac{(W \Phi V)(R \Phi L) - (1 - W \Phi L)(1 - R \Phi V)}{R \Phi L}.
\]

It follows that this interpolant has the same degree as \( W \Phi V \).

Finally a parametrization of all interpolants \( p/q \) of the given degree is given by

\[
\begin{bmatrix} p \\ q \end{bmatrix} = \Theta_2(s) \begin{bmatrix} 1 \\ \delta \end{bmatrix} = \begin{bmatrix} W \Phi V + \delta[(W \Phi V)(R \Phi L) - (1 - W \Phi L)(1 - R \Phi V)] \\ 1 + \delta R \Phi L \end{bmatrix}.
\]

The above results are illustrated next.

8. Examples

The results of the preceding sections will now be illustrated by means of a few simple examples.

8.1. Example A

\[
A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.
\]

Thus the transfer function is

\[
H(s) = C(sE - A)^{-1}B = \begin{bmatrix} s & -1 \\ -1 & \frac{1}{s} \end{bmatrix}.
\]

We will now recover this system by sampling its transfer function directionally. The chosen interpolation data are: \( A = \text{diag}([1, 2, 3]) \), \( M = \text{diag}([-1, -2, -3, -4]) \), while

\[
L = \begin{bmatrix} 1 & -1 \\ 1 & -2 \\ 1 & -3 \\ 1 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}.
\]
These imply
\[
V = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -4 & -1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.
\]

It follows that the tangential controllability and observability matrices associated with this data are:

\[
X = [(E - A)^{-1}BR(:, 1), \quad (2E - A)^{-1}BR(:, 2), \quad (3E - A)^{-1}BR(:, 3)]
\]
\[
= \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 1 & \frac{1}{3} \end{bmatrix},
\]

\[
Y = \begin{bmatrix} L(1, :)C(-1E - A)^{-1} \\ L(2, :)C(-2E - A)^{-1} \\ L(3, :)C(-3E - A)^{-1} \\ L(4, :)C(-4E - A)^{-1} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & -4 & 0 \end{bmatrix}.
\]

We notice that \( \text{rank } X = \text{rank } Y = 2 \). Thus the Loewner and shifted Loewner matrices are

\[
\mathbb{L} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -\frac{1}{3} \\ 0 & 0 & -\frac{1}{3} \\ 1 & 1 & 0 \end{bmatrix}, \quad \sigma\mathbb{L} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \\ -4 & -4 & -1 \end{bmatrix}.
\]

We check assumption (23) for \( \delta = 0 \):

- \( \text{rank } (\mathbb{L} - \sigma\mathbb{L}) = 2 \), \( \text{rank } (-2 \cdot \mathbb{L} - \sigma\mathbb{L}) = 2 \),
- \( \text{rank } (2 \cdot \mathbb{L} - \sigma\mathbb{L}) = 2 \), \( \text{rank } (-3 \cdot \mathbb{L} - \sigma\mathbb{L}) = 2 \),
- \( \text{rank } (3 \cdot \mathbb{L} - \sigma\mathbb{L}) = 1 \), \( \text{rank } (-4 \cdot \mathbb{L} - \sigma\mathbb{L}) = 1 \),
- \( \text{rank } (-1 \cdot \mathbb{L} - \sigma\mathbb{L}) = 2 \), \( \text{rank } ([\mathbb{L}; \sigma\mathbb{L}]) = 2 \),
- \( \text{rank } (-1 \cdot \mathbb{L} - \sigma\mathbb{L}) = 2 \), \( \text{rank } ([\mathbb{L}; \sigma\mathbb{L}]) = 2 \).

And for \( \delta = 1 \):

- \( \text{rank } (\mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \), \( \text{rank } (-2 \cdot \mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \),
- \( \text{rank } (2 \cdot \mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \), \( \text{rank } (-3 \cdot \mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \),
- \( \text{rank } (3 \cdot \mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \), \( \text{rank } (-4 \cdot \mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \),
- \( \text{rank } (-1 \cdot \mathbb{L} - \sigma\mathbb{L} + L \cdot R) = 3 \), \( \text{rank } ([\mathbb{L}; \sigma\mathbb{L} - L \cdot R]) = 3 \),
- \( \text{rank } ([\mathbb{L}; \sigma\mathbb{L} - L \cdot R]) = 3 \).

Since assumption (23) is violated for \( D = \delta = 0 \), but is not violated for \( \delta = 1 \), we will compute interpolants with \( \delta \neq 0 \). From \( \mathbb{L} - \sigma\mathbb{L} + LR\delta \), we obtain the matrices

\[
\pi_l = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \end{bmatrix}, \quad \pi_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

such that the column span of \( \pi_l^* \) spans that of \( [\sigma\mathbb{L} - LR \mathbb{L}] \), and the row span of \( \pi_r^* \) spans that of \( [\mathbb{L}; \sigma\mathbb{L} - L \cdot R] \). The resulting transfer function is
\[
\hat{H}(s) = (W - \delta R)\pi_r[\pi_l(\sigma L + \delta LR - s L)\pi_r]^{-1}\pi_l(V - \delta L) + \delta I_2
\]

Notice that the original rational function is obtained for \(\delta \to \infty\). It can be readily checked that all seven interpolation conditions are satisfied. A realization of this interpolant is as follows:

\[
s\hat{E} - \hat{A} = \pi_l \cdot (\sigma L - \delta LR - s L) \cdot \pi_r = \begin{bmatrix}
-2 \cdot \delta & -4 \cdot \delta & -2 \cdot \delta \\
-5 \cdot \delta & -3 + s + 8 \cdot \delta \\
5 \cdot \delta & 13 \cdot \delta & -4 \cdot \delta
\end{bmatrix}
\]

and thus \(\hat{E} = -\pi_l \cdot L \cdot \pi_r, \hat{A} = -\pi_l \cdot \sigma L \cdot \pi_r\), turn out to be

\[
\begin{align*}
\hat{E} &= \begin{bmatrix} 0 & 0 & 0 \\
0 & 0 & 1 \end{bmatrix}, \\
\hat{A} &= \begin{bmatrix} 2 \cdot \delta & 4 \cdot \delta & 2 \cdot \delta \\
4 + \delta & 4 + \delta & 1 \\
-5 \cdot \delta & -13 \cdot \delta & 3 - 8 \cdot \delta \end{bmatrix}.
\end{align*}
\]

Furthermore \(\hat{B} = \pi_l \cdot (V - \delta L)\) and \(\hat{C} = (W - \delta R) \cdot \pi_r\):

\[
\begin{align*}
\hat{B} &= \begin{bmatrix} 0 & -2 \cdot \delta \\
-4 - \delta & -1 \\
-3 \cdot \delta & 8 \cdot \delta \end{bmatrix}, \\
\hat{C} &= \begin{bmatrix} -\delta & -\delta & -1 \\
-\delta & -2 \cdot \delta & 1/3 - \delta \\
-\delta & -\delta & 1/3 - \delta \end{bmatrix}.
\end{align*}
\]

Finally we need to check that the characteristic polynomial of the system is non zero at all interpolation points:

\[
\chi(s) = \det(s\hat{E} - \hat{A}) = -2\delta(s^2 + (\delta + 1)s - 12)
\]

Therefore

\[
\begin{align*}
\chi(1) &= -2\delta(\delta - 10), & \chi(-1) &= 2\delta(\delta + 12), \\
\chi(2) &= -4\delta(\delta - 3), & \chi(-2) &= 4\delta(\delta + 5), \\
\chi(3) &= -6\delta^2, & \chi(-3) &= 6\delta(\delta + 2), \\
\chi(-4) &= 8\delta^2
\end{align*}
\]

Thus \(\delta\) must be different from \(-12, -5, -2, 0, 3, 10\).

8.2. Example B

The next example shows that assumption (20) may fail and hence no interpolant can be constructed using the proposed method. Here we sample again the rational matrix function

\[
H(s) = \begin{bmatrix} s & -1 \\
-1 & \frac{1}{s} \end{bmatrix}
\]

as follows. As before \(A = \text{diag}([1 \ 2 \ 3]), M = \text{diag}([-1 \ -2 \ -3 \ -4])\).

\[
L = \begin{bmatrix} 1 & -1 \\
1 & -2 \\
1 & -3 \\
1 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 1 & 1 \\
1 & 2 & 3 \end{bmatrix},
\]

which imply the left and right values
\[ V = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -4 & -1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \]

The resulting tangential controllability and observability matrices are
\[ X = \begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 \\ 1 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & -4 & 0 \end{bmatrix}, \]
and therefore rank \( X = 1 \) while rank \( Y = 2 \). Finally
\[ L = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \sigma L = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -4 & -4 & -4 \end{bmatrix}. \]
Checking assumption (20) we get
\[
\begin{align*}
\text{rank}(L\sigma L) &= 1, \\
\text{rank}(L; \sigma L) &= 1, \\
\text{rank}(L - \sigma L) &= 1, \\
\text{rank}(2L - \sigma L) &= 1, \\
\text{rank}(3L - \sigma L) &= 1, \\
\text{rank}(4L - \sigma L) &= 0,
\end{align*}
\]
while rank \((L\sigma L - LR)\) = 3 and rank \((L; \sigma L - LR)\) = 2; the last equality holds for any \( D \neq 0 \).
Thus no interpolant can be constructed using the proposed method. In particular the interpolant provided by this method, namely \( H(s) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \), satisfies all interpolation conditions, but one.

8.3. Example C

This example incorporates derivative constraints. We are given the following data:
\[ \mathbf{r}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{r}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{w}^0_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{w}^1_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{w}^0_2 = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}, \quad \mathbf{w}^1_2 = \begin{bmatrix} 2/4 \\ 1 \end{bmatrix}, \]
on the right and
\[
\begin{align*}
\ell^*_1 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\
\ell^*_2 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\
[v^0_1]^* &= \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \\
[v^0_2]^* &= \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \\
[v^1_1]^* &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad [v^1_2]^* &= \begin{bmatrix} 0 \\ 1 \end{bmatrix},
\end{align*}
\]
on the left. We want to find a rational matrix function \( H(s) \) such that \( H^{j-1}(i) \mathbf{r}_i = \mathbf{w}^{j-1}_i \), \( i, j \in \{1, 2\} \) and \( \ell_i H^{j-1}(-i) = v^{j-1}_i \), \( i, j \in \{1, 2\} \). The data is then arranged as follows:
\[ A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} \mathbf{w}^0_1 & \mathbf{w}^1_1 & \mathbf{w}^0_2 & \mathbf{w}^1_2 \end{bmatrix}, \]
\[
M = \begin{bmatrix}
-1 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 \\
0 & 0 & -2 & 0 \\
0 & 0 & 1 & -2
\end{bmatrix}, \quad
L = \begin{bmatrix}
1 & 0 \\
0 & 0 \\
1 & 0 \\
0 & 0
\end{bmatrix}, \quad
V = \begin{bmatrix}
v_1^0 \\
v_1^1 \\
v_2^0 \\
v_2^1
\end{bmatrix}.
\]

The Loewner and shifted Loewner matrices \( L, \sigma L \) turn out to be
\[
L = \begin{bmatrix}
1 & 2 & 0 \\
0 & 0 & 0 \\
1 & 2 & -1 \\
2 & 4 & -1
\end{bmatrix}, \quad
\sigma L = \begin{bmatrix}
-1 & 1 & 1 & 2 \\
1 & 0 & 2 & 0 \\
-1 & 0 & -2 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}.
\]

One can easily verify that (20) is satisfied with \( k = 3 \) and so proceed as in the normal case. The following is a representation of an interpolant:
\[
E = \begin{bmatrix}
-5 & -8 & -1 \\
-\frac{5}{2} & -\frac{185}{16} & -\frac{5}{2} \\
1 & -\frac{37}{8} & -1
\end{bmatrix}, \quad
A = \begin{bmatrix}
-1 & 8 & -7 \\
10 & 16 & 2 \\
-10 & -16 & -2
\end{bmatrix},
\]
\[
B = \begin{bmatrix}
-1 & -1 \\
-2 & -\frac{5}{4} \\
2 & \frac{1}{2}
\end{bmatrix}, \quad
C = \begin{bmatrix}
6 & 0 & 8 \\
-3 & \frac{5}{4} & -3
\end{bmatrix}.
\]

The resulting transfer function is \( H(s) = \begin{bmatrix} s & -1 \\ -1 & \frac{1}{s} \end{bmatrix} \).

8.4. Example D

The purpose of this example is to illustrate the parametrization result of Section 7. We consider the polynomial \( H(s) = s^3 \). The points are \( \Lambda = \text{diag}[1, 2, 3] \) and \( M = \text{diag}[-1, -2, -3] \); furthermore \( R = L^* = [1, 1, 1] \). This implies \( W = -V^* = [1, 8, 27] \). The Loewner and shifted Loewner matrices are
\[
L = \begin{bmatrix}
1 & 3 & 7 \\
3 & 4 & 7 \\
7 & 7 & 9
\end{bmatrix}, \quad
\sigma L = \begin{bmatrix}
0 & 5 & 20 \\
-5 & 0 & 13 \\
-20 & -13 & 0
\end{bmatrix}.
\]

Since \( \det(sL - \sigma L) = 4s(s^2 - 14) \), assumption (20) is satisfied and therefore the procedure of Lemma 5.1 yields the minimal realization \( E = -L, A = -\sigma L, B = V, C = W \); furthermore the parametrization of all interpolants of McMillan degree three is given in terms of the (scalar) parameter \( \delta = D \) as follows:
\[
\hat{H}(s) = (W - \delta R)(sL - \delta LR - sL)^{-1}(V - \delta L) + \delta = \frac{(\delta s^3 - 49s^2 + 36)}{s^3 - 14s + \delta}.
\]

The generating system approach of [5] yields the following matrix:
\[
\Theta = \begin{bmatrix}
s^3 & -1 \\
49s^2 - 36 & s^3 - 14s
\end{bmatrix} \Rightarrow \det \Theta(s) = (s - 1)(s - 2)(s - 3)(s + 3)(s + 2)(s + 1),
\]
which implies that all interpolants are obtained by taking linear combinations of the two rows and forming the quotient of the first over the second entry. This is exactly the expression obtained above. Therefore we obtain a parametrization of all interpolants of McMillan degree three. The polynomial \( s^3 \) that we started out with is obtained for \( \delta \to \infty \).

8.5. Example E

We are given the rational matrix

\[
H(s) = \begin{bmatrix} s^2 & 1 \\ s+1 & s+1 \end{bmatrix},
\]

which is sampled with \( r = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \) at \( \lambda = 0 \). Thus the following samples are obtained:

\[
\begin{align*}
H(0)(0) &= [01], & h_0 &= H(0)(0)r = 1, & h_4 &= H(4)(0)r = 2, \\
H(1)(0) &= [0 -1], & h_1 &= H(1)(0)r = -1, & h_5 &= H(5)(0)r = -2, \\
H(2)(0) &= [11], & h_2 &= H(2)(0)r = 2, & h_6 &= H(6)(0)r = 2, \\
H(3)(0) &= [-1 -1], & h_3 &= H(3)(0)r = -2, & h_7 &= H(7)(0)r = -2,
\end{align*}
\]

where \( H^{(k)}(0) = \frac{d^k H}{ds^k} \big|_{s=0} \), \( k = 0, 1, \ldots, 7 \). Following (28) and (29), the Loewner and shifted Loewner matrices are

\[
\mathbb{L} = \begin{bmatrix} h_1 & h_2 & h_3 & h_4 \\
h_2 & h_3 & h_4 & h_5 \\
h_3 & h_4 & h_5 & h_6 \\
h_4 & h_5 & h_6 & h_7 \end{bmatrix}, \quad \mathbb{L} = \begin{bmatrix} -1 & 2 & -2 & 2 \\
2 & -2 & 2 & -2 \\
-2 & 2 & -2 & 2 \\
2 & -2 & 2 & -2 \end{bmatrix},
\]

\[
\sigma \mathbb{L} = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 \\
h_1 & h_2 & h_3 & h_4 \\
h_2 & h_3 & h_4 & h_5 \\
h_3 & h_4 & h_5 & h_6 \end{bmatrix}, \quad \mathbb{L} = \begin{bmatrix} 1 & -1 & 2 & -2 \\
-1 & 2 & -2 & 2 \\
2 & -2 & 2 & -2 \\
-2 & 2 & -2 & 2 \end{bmatrix}.
\]

Then, the quadruplet \((E, A, B, C)\), where \( E = -\mathbb{L}_3 = -\mathbb{L}(1:3, 1:3) \), \( A = -\sigma \mathbb{L}_3 = -\sigma \mathbb{L}(1:3, 1:3) \),

\[
B = \begin{bmatrix} H(0)(0) \\
H(1)(0) \\
H(2)(0) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\
0 & -1 \\
1 & 1 \end{bmatrix}, \quad C = [h_0 \ h_1 \ h_2] = [1 \ -1 \ 2],
\]

is a minimal realization of the data:

\[
C(sE - A)^{-1}B = \begin{bmatrix} s^2 & 1 \\
\end{bmatrix} = H(s).
\]

9. Conclusions

In this paper we present a framework for the construction of matrix rational interpolants \( H \), in generalized state space form, given tangential interpolation data. We thus construct generalized state space realizations \((E, A, B, C, D)\), where \( E \) may be singular, such that

\[
H(s) = C(sE - A)^{-1}B + D,
\]
from tangential interpolation data. We call this the generalized realization problem. Central objects for this construction are the Loewner matrix $\mathcal{L}$, and the shifted Loewner matrix $\sigma \mathcal{L}$, associated with the data; $\sigma$ is introduced here for the first time. The approach parallels the solution of the realization problem in which state space representations $(A, B, C)$ are constructed given the Hankel matrix of Markov parameters, that is given interpolation data at infinity.

References