Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

A Subspace-Based Approach to Lagrange-Sylvester Interpolation of Rational Matrix Functions

Hüseyin Akçay and Semiha Türkay

***Abstract*— In this paper, we study Lagrange-Sylvester interpolation of rational matrix functions which are analytic at infinity and propose a new interpolation algorithm based on the recent subspace-based identification methods. As an application, we consider the problem of system identification with interpolation constraints.**

I. Introduction

Let us consider a multi-input/multi-output, linear-time invariant, discrete-time system represented by the state-space equations:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{1}$$

where $x(k) \in \mathbf{R}^n$ is the state, $u(k) \in \mathbf{R}^m$ and $y(k) \in \mathbf{R}^p$ are, respectively, the input and the output of the system. Here, $\mathbf{R}$ is the set of the real numbers, and the set of the complex numbers is denoted by $\mathbf{C}$. The transfer function of the system (1) denoted by $G(z)$ is computed as

$$G(z) = D + C(zI_n - A)^{-1}B \tag{2}$$

where I_n is the n by n identity matrix. We assume that the system (1) is stable and the pairs (A, B) and (C, A) are *controllable* and *observable*, respectively. The stability of (1) means that $G(z)$ is a *proper rational* matrix that is analytic and bounded in the region $\mathcal{D} = \{z \in \mathbf{C} : |z| \geq 1\}$; and both the controllability and the observability of the pairs (A, B) and (C, A) means that the quadruplet (A, B, C, D) is a *minimal realization* of $G(z)$.

The interpolation problem studied in this paper can be stated as follows.

Given: noise-free samples of $G(z)$ and its derivatives at L distinct points $z_k \in \mathcal{D}$

$$\frac{d^j G(z_k)}{dz^j} = w_{kj}, \quad j = 0, 1, \cdots, N_k; \quad k = 1, 2, \cdots, L. \tag{3}$$

Find: a quadruplet $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ that is a minimal realization of $G(z)$.

This problem is known as the *Lagrange-Sylvester* rational interpolation problem. One approach to finding a solution is to reduce the problem to a system of independent scalar problems, which is not interesting from the point of matrix interpolation theory. Besides, a minimal realization is obtained only after the elimination of unobservable or/and uncontrollable modes. The tangential and contour integral versions of this problem are treated in the comprehensive work [1]. Related problems are the nonhomogenous interpolation problem with metric constraints as in the various types of Nevanlinna-Pick interpolation and its generalizations, and the partial realization problem, *that is*, finding a rational matrix function analytic at infinity of the smallest possible McMillan degree with prescribed values of itself and a few of its derivatives at infinity. The interested reader is referred to [1], [4], [5]. Applications of interpolation theory to control and systems theory and estimation are presented in [1], [2], [6].

In this paper, we present a subpace-based method to solve the Lagrange-Sylvester interpolation problem formulated above. This method is particularly useful when the samples of $G(z)$ and its derivatives are corrupted by noise and the number of data is large. In the noisy case, most interpolation schemes deliver state-space realizations with McMillan degrees tending to infinity as the number of data grows unboundedly; thus sensitive to inaccuracies in the interpolation data. The proposed algorithm is based on the subspace identification algorithm presented in [3]; and hence it has the benefits of the latter. For example, there is no need for explicit model parameterization, and it is computationaly efficient since it uses numerically robust QR and the singular value decomposition. In the paper, we also consider subspace-based system identification with interpolation constraints.

II. Subspace-Based Interpolation Algorithm

We begin by taking the z-transform of (1):

$$\begin{aligned} zX(z) &= AX(z) + BU(z) \\ Y(z) &= CX(z) + DU(z) \end{aligned} \tag{4}$$

where $X(z)$, $Y(z)$, and $U(z)$ denote respectively the z-transforms of $x(k)$, $y(k)$, and $u(k)$ defined by

$$U(z) \triangleq \sum_{k=0}^{\infty} u(k)\, z^{-k}. \tag{5}$$

Let $X_j(z)$ be the resulting state z-transform when $u(k) = e_j$, the unit vector in R^m with one on the jth position and zero elsewhere. By defining the compound state z-transform matrix:

$$X_{\mathrm{C}}(z) \triangleq [X_1(z) \;\; X_2(z) \;\; \cdots \;\; X_m(z)], \tag{6}$$

$G(z)$ can implicitly be described as

$$G(z) = CX_{\mathrm{C}}(z) + D \tag{7}$$

Department of Electrical and Electronics Engineering, Anadolu University, 26470 Eskişehir, Turkey.
E-mail: huakcay@anadolu.edu.tr; Tel: + (90) 222 335 0580 -X 6459; Fax: +(90) 222 323 9501.

0-7803-9568-9/05/$20.00 ©2005 IEEE

with

$$zX_{\mathrm{C}}(z) = AX_{\mathrm{C}}(z) + B. \tag{8}$$

By recursive use of (8), we obtain the relation

$$z^k X_{\mathrm{C}}(z) = A^k X_{\mathrm{C}}(z) + \sum_{j=0}^{k-1} A^{k-1-j} B z^j, \quad k \geq 1. \tag{9}$$

Multiplying both sides of (9) with C and using (7) we get

$$\begin{aligned} z^k G(z) &= CA^k X_{\mathrm{C}}(z) + Dz^k \\ &\quad + \sum_{j=0}^{k-1} CA^{k-1-j} B z^j, \quad k \geq 1. \end{aligned} \tag{10}$$

Now, recall that the impulse response coefficients of $G(z)$ are given by

$$g_k = \begin{cases} D, & k = 0; \\ CA^{k-1}B, & k \geq 1. \end{cases} \tag{11}$$

Thus, from (7), (10), and (11)

$$z^k G(z) = CA^k X_{\mathrm{C}}(z) + \sum_{j=0}^{k} g_{k-j}\, z^j, \qquad k \geq 0. \tag{12}$$

Hence from (12),

$$\begin{bmatrix} G(z) \\ zG(z) \\ \vdots \\ z^{q-1}G(z) \end{bmatrix} = \mathcal{O}_q X_{\mathrm{C}}(z) + \Gamma_q \begin{bmatrix} I_m \\ zI_m \\ \vdots \\ z^{q-1} I_m \end{bmatrix} \tag{13}$$

where

$$\begin{aligned} \mathcal{O}_q &\triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix}, \\ \Gamma_q &\triangleq \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{q-1} & g_{q-2} & \cdots & g_0 \end{bmatrix}. \end{aligned} \tag{14}$$

For later use, let us write (13) in a compact form. The matrix $\mathcal{O}_q$ is known as the *extended observability matrix* and has full rank n if (C, A) is an observable pair and $q \geq n$. We define the *Kronecker product* of two matrices $E \in \mathbf{C}^{m\times n}$ and $F \in \mathbf{C}^{p\times q}$ by

$$E \otimes F \triangleq \begin{bmatrix} E_{11}F & E_{12}F & \cdots & E_{1n}B \\ E_{21}F & E_{22}F & \cdots & E_{2n}F \\ \vdots & \vdots & \ddots & \vdots \\ E_{m1}F & E_{m2}F & \cdots & E_{mn}F \end{bmatrix} \in \mathbf{C}^{mp\times nq}.$$

Let

$$\begin{aligned} \mathcal{Z}_q(z) &\triangleq \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{q-1} \end{bmatrix}, \\ \mathcal{J}_{q,2} &\triangleq \begin{bmatrix} 0 & & \cdots & & 0 \\ 1 & 0 & & & \\ 0 & 1 & 0 & & \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & 1 & 0 \end{bmatrix} \in \mathrm{R}^{q\times q}. \end{aligned} \tag{15}$$

By a slight abuse of notation, let $\mathcal{J}_{q,1}$ denote the q by q identity matrix. Observe that $\mathcal{J}_{q,2}$ is obtained by shifting the elements of $\mathcal{J}_{q,1}$ one row down and filling its first row with zeros. Let $\mathcal{J}_{q,j}$ denote the matrix obtained by $j-1$ repeated applications of this process to $\mathcal{J}_{q,1}$. Note the following relations

$$\mathcal{J}_{q,j} = \begin{cases} \mathcal{J}_{q,2}^{j-1}, & j \leq q \\ 0, & j > q. \end{cases} \tag{16}$$

Thus, the lower triangular block Toeplitz matrix in (14) can be written as

$$\Gamma_q = \sum_{j=0}^{q-1} \mathcal{J}_{q,1+j} \otimes g_j \tag{17}$$

Hence, from (14)–(17) we arrive the following compact expression for (13)

$$\begin{aligned} \mathcal{Z}_q(z) \otimes G(z) &= \mathcal{O}_q X_{\mathrm{C}}(z) \\ &\quad + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j]\ [\mathcal{Z}_q(z) \otimes I_m]. \end{aligned} \tag{18}$$

This equation forms the basis of the frequency domain subspace-based identification algorithms [3]. In subspace-based identification algorithms, $\mathcal{Z}_q(z) \otimes G(z)$ and the right hand side of (18) are evaluated at a set of distinct points on the unit circle, and then stacked into columns of long matrices. This procedure yields a matrix equation *affine* in $\mathcal{O}_q$. From this equation, the range space of $\mathcal{O}_q$ is recovered by a projection. Once the observability range space is recovered, a realization of $G(z)$ is derived in a routine manner. This will also be our strategy.

First, we differentiate $\mathcal{Z}_q(z) \otimes G(z)$ in (18) l times with respect to z:

$$\begin{aligned} H_q^{(l)}(z) &= \sum_{j=0}^{l} \begin{pmatrix} l \\ j \end{pmatrix} \left[\mathcal{Z}_q^{(j)}(z) \otimes G^{(l-j)}(z)\right] \\ &= \mathcal{O}_q \frac{\mathrm{d}^l X_{\mathrm{C}}(z)}{\mathrm{d}z^k} \\ &\quad + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j] \left[\mathcal{Z}_q^{(l)}(z) \otimes I_m\right], \quad l \geq 0 \end{aligned} \tag{19}$$

where

$$H_q(z) \triangleq \mathcal{Z}_q(z) \otimes G(z). \tag{20}$$

Then, we augment $H_q(z_k)$ and the first N_k derivatives of $H_q(z)$ at z_k in a data matrix:

$$\mathcal{H}_k \triangleq \left[H_q(z_k)\ \ H_q'(z_k)\ \ \cdots\ \ H_q^{(N_k)}(z_k)\right], \quad k \leq L. \tag{21}$$

Using the right hand side of the first equality in (19), let us derive a compact expression for $\mathcal{H}_k$ in terms of the elementary matrices:

$$\mathcal{D}_{N_k+1} \triangleq \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & & \\ & & 0 & \cdots & \\ \vdots & & & \ddots & N_k \\ 0 & & & \cdots & 0 \end{bmatrix} \tag{22}$$

and

$$\mathcal{W}_k \triangleq \left[\mathcal{Z}_q(z_k)\ \ \mathcal{Z}_q'(z_k)\ \ \cdots\ \ \mathcal{Z}_q^{(N_k)}(z_k)\right], \quad k \leq L \tag{23}$$

as follows

$$\begin{aligned} \mathcal{H}_k \ =\ & \left[\mathcal{Z}_q(z_k)\ \ \mathcal{Z}_q'(z_k)\ \ \cdots\ \ \mathcal{Z}_q^{(N_k)}(z_k)\right] \otimes G(z_k) \\ & + \left[0\ \ \mathcal{Z}_q(z_k)\ \ \cdots\ \ \begin{pmatrix} N_k \\ 1 \end{pmatrix} \mathcal{Z}_q^{(N_k-1)}(z_k)\right] \otimes G'(z_k) \\ & + \left[0\ \ 0\ \ \cdots\ \ \begin{pmatrix} N_k \\ 2 \end{pmatrix} \mathcal{Z}_q^{(N_k-2)}(z_k)\right] \otimes G''(z_k) \\ & + \cdots + [0\quad 0\quad \cdots\quad \mathcal{Z}_q(z_k)] \otimes G^{(N_k)}(z_k) \\ =\ & \mathcal{W}_k \otimes G(z_k) + [\mathcal{W}_k \mathcal{D}_{N_k+1}] \otimes G'(z_k) \\ & + \frac{1}{2!}\,[\mathcal{W}_k \mathcal{D}_{N_k+1}^2] \otimes G''(z_k) \\ & + \frac{1}{N_k!}\,[\mathcal{W}_k \mathcal{D}_{N_k+1}^{N_k}] \otimes G^{(N_k)}(z_k) \\ =\ & \sum_{j=0}^{N_k} \frac{1}{j!}\,[\mathcal{W}_k\, \mathcal{D}_{N_k+1}^j] \otimes w_{kj}, \quad k \leq L. \end{aligned} \tag{24}$$

Note that $\mathcal{D}_{N_k+1}^j = 0$ for all $j > N_k$.

An alternative compact expression for $\mathcal{H}_k$ is obtained by evaluating the right hand side of the second equality in (19) for $r = 0, \cdots, N_k$; $k = 1, \cdots, L$, and augmenting the similar terms in compound matrices as follows

$$\mathcal{H}_k = \mathcal{O}_q\, \mathcal{X}_k + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j]\, [\mathcal{W}_k \otimes I_m], \quad k \leq L \tag{25}$$

where

$$\mathcal{X}_k \triangleq \left[X_{\mathrm{C}}(z_k)\ \ X_{\mathrm{C}}'(z_k)\ \ \cdots\ \ X_{\mathrm{C}}^{(N_k)}(z_k)\right], \quad k \leq L. \tag{26}$$

It remains to compute the derivatives of $\mathcal{Z}_q(z)$. To this end, let

$$\mathcal{T}_q \triangleq \begin{bmatrix} 0! & 0 & \cdots & 0 \\ 0 & 1! & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (q-1)! \end{bmatrix} \in \mathbf{R}^{q \times q}. \tag{27}$$

Then, it is easy to verify that

$$\frac{\mathrm{d}^l \mathcal{Z}_q(z)}{\mathrm{d}z^l} = \mathcal{T}_q \mathcal{J}_{q,2}^l \mathcal{T}_q^{-1} \mathcal{Z}_q(z), \qquad l \geq 0. \tag{28}$$

Now, we collect $\mathcal{H}_k$, $\mathcal{X}_k$, and $\mathcal{W}_k$, $k = 1, \cdots, L$ in the compound matrices:

$$\mathcal{H} \ \triangleq\ [\mathcal{H}_1\ \mathcal{H}_2\ \cdots\ \mathcal{H}_L]\,, \tag{29}$$

$$\mathcal{X} \ \triangleq\ [\mathcal{X}_1\ \mathcal{X}_2\ \cdots\ \mathcal{X}_L]\,, \tag{30}$$

$$\mathcal{W} \ \triangleq\ [\mathcal{W}_1\ \mathcal{W}_2\ \cdots\ \mathcal{W}_L]\,. \tag{31}$$

Hence,

$$\mathcal{H} = \mathcal{O}_q\, \mathcal{X} + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j]\, [\mathcal{W} \otimes I_m] \tag{32}$$

where $\mathcal{H}$ and $\mathcal{W}$ are computed from the formulae (15), (22)–(24), (27)–(29), and (31). This completes the first stage of our subspace-based interpolation algorithm. Observe that $\mathcal{H}$ is affine in $\mathcal{O}_q$ as advertised.

Since $\mathcal{O}_q$ is a real matrix and we are interested in the real range space, we can convert (32) into a relation involving only real valued matrices:

$$\widehat{\mathcal{H}} = \mathcal{O}_q\, \widehat{\mathcal{X}} + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j]\, \mathcal{F} \tag{33}$$

where

$$\widehat{\mathcal{H}} \ \triangleq\ [\mathrm{Re}\mathcal{H}\ \ \mathrm{Im}\mathcal{H}]\,, \tag{34}$$

$$\widehat{\mathcal{X}} \ \triangleq\ [\mathrm{Re}\mathcal{X}\ \ \mathrm{Im}\mathcal{X}]\,, \tag{35}$$

$$\mathcal{F} \ \triangleq\ [\mathrm{Re}\mathcal{W}\ \ \mathrm{Im}\mathcal{W}] \otimes I_m. \tag{36}$$

A. Projection onto the observability range space

Let $\mathcal{F}^{\perp}$ be the projection matrix onto the null space of $\mathcal{F}$ given by

$$\mathcal{F}^{\perp} \triangleq I_{2m\sum_{k=1}^{L}(N_k+1)} - \mathcal{F}^T(\mathcal{F}\mathcal{F}^T)^{-1}\mathcal{F} \tag{37}$$

where $\mathcal{F}^T$ denotes the transpose of $\mathcal{F}$. The summand in (33) is cancelled for all j when multiplied from right by $\mathcal{F}^{\perp}$. Thus,

$$\widehat{\mathcal{H}}\mathcal{F}^{\perp} = \mathcal{O}_q\, \widehat{\mathcal{X}}\mathcal{F}^{\perp}. \tag{38}$$

The range space of $\widehat{\mathcal{H}}\mathcal{F}^{\perp}$ equals the range space of $\mathcal{O}_q$ unless rank cancellations occur. A sufficient condition for the range spaces to be equal is that the intersection between the row spaces of $\mathcal{F}$ and $\widehat{\mathcal{X}}$ is empty. In the following, we present sufficient conditions in terms of the data and the system. Let z^* denote the complex conjugate of z. In (3), without loss of generally we assume that $z_k \neq z_j^*$ for all nonreal numbers z_k and z_j since $G(z) = [G(z^*)]^*$.

Lemma 1: Let $\widehat{\mathcal{X}}$ and $\mathcal{F}$ be as in (35) and (36), respectively. Let

$$N \triangleq \sum_{k:z_k \in \mathbf{R}} (N_k + 1) + \sum_{k:z_k \in \mathbf{C}-\mathbf{R}} 2(N_k + 1). \tag{39}$$

Suppose that $N \geq q + n$ and the eigenvalues of A do not coincide with the distinct complex numbers z_k. Then,

$$\mathrm{rank} \begin{bmatrix} \mathcal{F} \\ \widehat{\mathcal{X}} \end{bmatrix} = qm + n \iff (A, B) \text{ controllable.} \tag{40}$$

Proof: Equation (8) implies that

$$X_{\mathrm{C}}(z) = (zI_n - A)^{-1}B.$$

The matrix $\begin{bmatrix} \mathcal{W} \otimes I_m \\ \mathcal{X} \end{bmatrix}$ is rank deficient if and only if there exists a row vector

$$[\alpha_0 \ \cdots \ \alpha_{q-1} \ \beta] \neq 0 \tag{41}$$

with $\alpha_k^T \in \mathrm{R}^m, \ k = 0, \cdots, q-1$ and $\beta^T \in \mathrm{R}^n$ such that

$$[\alpha_0 \ \cdots \ \alpha_{q-1} \ \beta] \begin{bmatrix} \mathcal{W} \otimes I_m \\ \mathcal{X} \end{bmatrix} = 0. \tag{42}$$

From (31), (23), (15) and (30), (26), equation (42) holds if and only if

$$[\alpha_0 \ \cdots \ \alpha_{q-1} \ \beta] \frac{\mathrm{d}^{r_k}}{\mathrm{d}z^{r_k}} \begin{bmatrix} \mathcal{Z}_q(z) \otimes I_m \\ \mathcal{X}_{\mathrm{C}}(z) \end{bmatrix}_{z=z_k} = 0,$$

$$0 \leq r_k \leq N_k; \ k = 1, \cdots, L$$

$$\Updownarrow$$

$$\frac{\mathrm{d}^{r_k} E(z_k)}{\mathrm{d}z^{r_k}} = 0, \quad 0 \leq r_k \leq N_k; \ k = 1, \cdots, L \tag{43}$$

where

$$E(z) \triangleq \sum_{k=0}^{q-1} \alpha_k z^k + \beta(zI_n - A)^{-1}B.$$

Equation (43) implies that for eack k, the elements of the rational vector $E(z)$ have common zeros at z_k with multiplicity N_k+1. Since $E(z)$ is real-rational, z_k is a zero if and only if z_k^* is also a zero. Therefore, $E(z)$ happens to have a total number of N zeros counting multiplicities. But, the elements of $E(z)$ have numerator degrees not exceeding $n+q-1$. Hence, it can not have N zeros unless it vanishes on the entire complex plane. Thus, $E(z) \equiv 0$. This implies that $\alpha_k = 0$ for all k and $\beta(zI_n - A)^{-1}B \equiv 0$. The latter, follows from the fact that $\beta(zI_n - A)^{-1}B$ is analytic at $z = \infty$; hence orthogonal to $\sum_{k=0}^{q-1} \alpha_k z^k$. Recall that (A, B) is an uncontrollable pair if and only if it is possible to find a vector $\beta \neq 0$ such that $\beta(zI_n - A)^{-1}B \equiv 0$. Finally, note that $\begin{bmatrix} \mathcal{F} \\ \widehat{\mathcal{X}} \end{bmatrix}$ is rank deficient if and only if $\begin{bmatrix} \mathcal{W} \otimes I_m \\ \mathcal{X} \end{bmatrix}$ is rank deficient. The last assertion is due to the fact that for any complex matrix Z,

$$x^T Z = 0 \Longleftrightarrow x\,[\mathrm{Re}Z \ \mathrm{Im}Z] = 0.$$

■

Since all the eigenvalues of A are inside the unit circle, none of them coincide with any of z_k. Thus, by applying Lemma 1 we conclude that the two row spaces of $\widehat{\mathcal{X}}$ and $\mathcal{F}$ do not intersect and the range space of $\widehat{\mathcal{H}}\mathcal{F}^\perp$ coincides with the range space of $\mathcal{O}_q$. Then, using the singular value factorization of $\widehat{\mathcal{H}}\mathcal{F}^\perp$:

$$\begin{aligned} \widehat{\mathcal{H}}\mathcal{F}^\perp &= \widehat{U}\widehat{\Sigma}\widehat{V}^T \\ &= \begin{bmatrix} \widehat{U}_s & \widehat{U}_o \end{bmatrix} \begin{bmatrix} \widehat{\Sigma}_s & 0 \\ 0 & \widehat{\Sigma}_o \end{bmatrix} \begin{bmatrix} \widehat{V}_s^T \\ \widehat{V}_o^T \end{bmatrix} \end{aligned} \tag{44}$$

where $\widehat{\Sigma}_s \in \mathbf{R}^{n \times n}$, we determine the system matrices $\widehat{A}$ and $\widehat{C}$ as

$$\widehat{A} = (J_1\widehat{U}_s)^\dagger J_2\widehat{U}_s, \qquad \widehat{C} = J_3\widehat{U}_s \tag{45}$$

where

$$J_1 = \begin{bmatrix} I_{(q-1)p} & 0_{(q-1)p \times p} \end{bmatrix}, \tag{46}$$

$$J_2 = \begin{bmatrix} 0_{(q-1)p \times p} & I_{(q-1)p} \end{bmatrix}, \tag{47}$$

$$J_3 = \begin{bmatrix} I_p & 0_{p \times (q-1)p} \end{bmatrix}, \tag{48}$$

$0_{i \times j}$ is the $i \times j$ zero matrix, and $X^\dagger = (X^T X)^{-1} X^T$ is the Moore-Penrose pseudo inverse of the full column rank matrix X. Provided that (C, A) is an observable pair, the pseudo inverse in (45) exists if and only if $q > n$. Thus, N defined in Lemma 1 is forced to satisfy $N > 2n$. In this case, $\widehat{A}$ and $\widehat{C}$ are related to A and C in (1) by

$$\widehat{A} = T^{-1}AT, \qquad \widehat{C} = CT \tag{49}$$

for some $T \in \mathbf{R}^{n \times n}$.

B. Extracting B and D from the data

We will now determine B and D matrices in the realization using the given frequency domain data. Repeated application of the differentiation formula

$$\frac{dX^{-1}}{dz} = -X^{-1}\frac{dX}{dz}X^{-1}$$

to $X_{\mathrm{C}}(z) = (zI_n - A)^{-1}B$ yields the derivatives of $G(z)$ as follows

$$G^{(j)}(z) = \delta_{0j}\, D + (-1)^j j!\, C(zI_n - A)^{-j-1}B, \ j \geq 0 \tag{50}$$

where δ_{ks} is the Kronecker delta. Now, let

$$\begin{aligned} \mathcal{G}_k &\triangleq \begin{bmatrix} w_{k0} \\ w_{k1} \\ \vdots \\ w_{kN_k} \end{bmatrix}, \\ \mathcal{G} &\triangleq \begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \vdots \\ \mathcal{G}_L \end{bmatrix}. \end{aligned} \tag{51}$$

Observe that for fixed A and C, the matrices B and D appear linearly in $\mathcal{G}$. Hence, we can uniquely determine B and D by solving the following linear least-squares problem

$$\widehat{B}, \widehat{D} = \arg\min_{B,D} \left\| \mathcal{G} - \mathcal{Y} \begin{bmatrix} B \\ D \end{bmatrix} \right\|_F^2 \tag{52}$$

where $\|X\|_F = (\sum_k \sum_s |x_{ks}|^2)^{1/2}$ is the Frobenius norm and

$$\mathcal{Y}_k \triangleq \begin{bmatrix} C(z_kI_n - A)^{-1} & I_p \\ -C(z_kI_n - A)^{-2} & 0 \\ \vdots & \\ (-1)^{N_k} N_k!\, C(z_kI_n - A)^{-N_k-1} & 0 \end{bmatrix}, \tag{53}$$

$$\mathcal{Y} \triangleq \begin{bmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \\ \vdots \\ \mathcal{Y}_L \end{bmatrix}$$

provided that $\mathcal{Y}$ is not rank deficient. For the last requirement, a sufficient condition is presented next.

Lemma 2: Let N and $\mathcal{Y}$ be as in (39) and (53), respectively. Suppose that $N > n$ and the eigenvalues of A do not coincide with the distinct complex numbers z_k. Then,

$$\operatorname{rank}\mathcal{Y} = p + n \qquad \Longleftrightarrow \qquad (C, A) \text{ observable.} \tag{54}$$

Proof: The matrix $\mathcal{Y}$ is rank deficient if and only if there exists $\begin{bmatrix} B \\ D \end{bmatrix} \neq 0$ such that

$$\mathcal{Y}\begin{bmatrix} B \\ D \end{bmatrix} = 0 \quad \Longleftrightarrow \quad \frac{\mathrm{d}^{r_k} G(z_k)}{\mathrm{d}z^{r_k}} = 0,$$

$$0 \leq r_k \leq N_k;\ k = 1, \cdots, L.$$

As in the proof of Lemma 1, this equation implies that every element of $G(z)$, has a total number of N zeros counting multiplicities, a contradiction if $G(z)$ is not identically zero unless $N \leq n$. ■

Thus, from (49) and Lemma 2 if $N \geq q + n$ and $q > n$, we have

$$\widehat{B} = T^{-1}B, \qquad \widehat{D} = D.$$

Moreover,

$$\widehat{G}(z) \triangleq \widehat{C}(zI_n - \widehat{A})^{-1}\widehat{B} + \widehat{D} = G(z).$$

Let us summarize the interpolation algorithm in the following.

Algorithm 3: Subspace-Based Interpolation Algorithm:

1) Given the data (3), form matrices $\widehat{\mathcal{H}}$ and $\widehat{\mathcal{F}}^{\perp}$ defined through equations (15), (20)–(24), (27)–(29), (31), (34), (36)–(37).
2) Calculate the singular value decomposition in (44).
3) Determine the system order by inspecting the singular values and partition the singular value decomposition such that $\widehat{\Sigma}_s$ contains the n largest singular values.
4) With J_1, J_2, and J_3 defined by (46)–(48), calculate $\widehat{A}$ and $\widehat{C}$ from (45).
5) Solve the least-squares problem (52) for $\widehat{B}$ and $\widehat{D}$ where $\mathcal{Y}$ and $\mathcal{G}$ are defined by (53) and (51), respectively.

Clearly, $\widehat{\Sigma}_o = 0$ in (44) when the data are not corrupted by noise. Step 3 will be needed only if are the data noisy.

Example 4: Let $G(z)$ be a single-input/single-output system represented by

$$G(z) = \frac{b_0 z^n + b_1 z + \cdots + b_n}{z^n + a_1 z + \cdots + a_n}.$$

We are to determine $2n + 1$ unknown real coefficients $a_1, \cdots, a_n, b_0, \cdots, b_n$ from N evaluations of $G(z)$ and its derivatives at a given set of distinct frequencies z_k.

In (3), let us first assume that $N_k = 0$ and $z_k \in \mathbf{C} - \mathbf{R}$ for all k. Then, $N = 2L$. Now, pick q equal to $n + 1$. If $2L \geq 2n + 1$ and the poles of $G(z)$ do not coincide with any of z_k, then Algorithm 3 delivers a minimal realization of $G(z)$. The first condition is satisfied by choosing $L = n + 1$; and the second condition holds, for example, if all z_k are on the unit circle excluding the points ± 1. Thus, Algoritm 3 recovers nth order stable systems from $n + 1$ noise-free frequency response measurements, excluding the frequencies 0 and π. Clearly, this is the least amount of data one could use to interpolate an nth order system. If the frequencies contain 0 and π, from (39) we then have $N = 2L - 2$. Hence, with $q = n + 1$ selected we must have $2L - 2 \geq 2n + 1$, which is fulfilled by letting $L = n + 2$. The last conclusion extends an interpolation result in [3] derived for uniformly spaced frequencies to nonuniformly spaced frequencies case.

It is easy to see that these results hold for multi-input/multi-output systems and $N_k > 0$ case as well. Therefore, Algorithm 3 is capable of using minimum amount of frequency domain data for the Lagrange-Sylvester interpolation.

The main result of this paper is captured in the following.

Theorem 5: Consider Algorithm 3 with the noise-free frequency domain data of a discrete-time stable system of order n. Let N be as in (39). If $N \geq q + n$, $q > n$, then the quadruplet $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ is a minimal realization of $G(z)$.

C. Subspace-Based Identification with Interpolation Constraints

In the rest of this section, we will consider identification of $G(z)$ from noisy samples of the frequency response

$$G_k = G(e^{i\omega_k}) + \eta_k, \qquad k = 1, \cdots, M \tag{55}$$

with the interpolation constraints:

$$\frac{\mathrm{d}^j G(z_k)}{dz^j} = E_{kj}, \qquad j = 0, 1, \cdots, N_k;\ \ k = 1, \cdots, L \tag{56}$$

where η_k is a sequence of independent, zero-mean complex random variables with a known covariance function that is uniformly bounded. The number of constraints defined in (39) satisfies $N < n$. The interpolation constraints (56) reflect the prior knowledge on $G(z)$. For example, by taking $E_{kj} = 0$ for all $j \leq N_k$ we enforce a zero with multiplicity $N_k + 1$ at z_k. These constraints may also be used as design variables to focus on a frequency band of interest.

We would like to find an identification algorithm which maps the data G_k to an nth order model $\widehat{G}_M(z)$ which satisfies the constraints in (56) such that with probability one,

$$\lim_{M\to\infty} \|\widehat{G}_M - G\|_\infty = 0$$

where $\|X\|_\infty = \sup_\omega \sigma_1(X(e^{i\omega}))$ and σ_1 denotes the largest singular value. Algorithms with this property are called *strongly consistent*.

The solution of this constrained identification problem is simple if one notes from (50) the following set of equations

$$\delta_{0j}\, D + (-1)^j j!\, C(z_k I_n - A)^{-j-1} B = E_{kj},$$

for $j = 0, \cdots, N_k$; $k = 1, \cdots, L$ which describe N hyperplanes in the parameter space of B and D for fixed C and A. Hence, it suffices to solve the linear least-squares problem (52) with the linear constraints above. With this modification, the frequency domain subspace-based identification algorithm presented in [3] is strongly consistent. The inclusion of the noise covariance information in the algorithm is straightforward and can be found in [3].

III. Conclusions

In this paper, we presented a new algorithm for the Lagrange-Sylvester interpolation of rational matrix functions that are analytic at infinity. This algorithm is related to the recent subspace-based identification methods and is not sensitive to inaccuracies in data. The purpose of this contribution was to exploit the relationships between identification of stable linear systems and rational interpolation problem in the frequency domain.

References

[1] J. A. Ball, I. Gohberg, and L. Rodman. *Interpolation of Rational Matrix Functions*, Birkhäuser, Basel, 1990.

[2] J. Chen and G. Gu. *Control Oriented System Identification: An H_∞ Approach*, Wiley-Interscience, New York, 2000.

[3] T. McKelvey, H. Akçay, and L. Ljung. *Subspace-based multivariable system identification from frequency response data*, IEEE Transactions on Automatic Control, 41, 1996, pp. 960–979.

[4] Nikolai K. Nikolski. *Operators, Functions, and Systems: An Easy Reading: Volume 1: Hardy, Hankel, and Toeplitz*, AMS, 2002.

[5] Nikolai K. Nikolski. *Operators, Functions, and Systems: An Easy Reading: Volume 2: Model Operators and Systems*, AMS, 2002.

[6] K. Y. Osipenko. *Optimal Recovery of Analytic Functions*, Nova Science Publishers, New York, 2000.