SIAM J. MATRIX ANAL. APPL.
Vol. 29, No. 2, pp. 377–395

A SUBSPACE-BASED METHOD FOR SOLVING LAGRANGE–SYLVESTER INTERPOLATION PROBLEMS*

HÜSEYIN AKÇAY† AND SEMIHA TÜRKAY†

Abstract. In this paper, we study the 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. The proposed algorithm is numerically efficient and delivers a minimal interpolant in state-space form. The solvability condition for the subspace-based algorithm is particularly simple and depends only on the total multiplicity of the interpolation nodes. As an application, we consider subspace-based system identification with interpolation constraints, which arises, for example, in the identification of continuous-time systems with a given relative degree.

Key words. rational interpolation, Lagrange–Sylvester, identification, subspace-based

AMS subject classifications. 93A30, 30E05, 65D05

DOI. 10.1137/050622171

1. Introduction.

Many problems in control, circuit theory, and signal processing can be reduced to the solution of matrix rational interpolation problems which have been widely studied (see, for example, [14, 15, 2, 19, 21, 3, 4, 1, 5, 6, 7, 31, 30, 11, 10] and the references therein). Applications arise, for example, in robust controller synthesis [19, 21], in the Q-parameterization of stabilizing controllers for unstable plants [18], in the problem of model validation [32], in circuit theory [34], in spectral estimation [12], and in adaptive filtering and control [31, 26].

In the simplest form, given complex numbers z_k and w_k for $k = 1, \ldots, N$, an interpolation problem asks for scalar rational functions $G(z)$ which meet the interpolation conditions

$$G(z_k) = w_k, \qquad k = 1, \ldots, N.$$

The interpolants can further be required to have minimal complexity in terms of their McMillan degree. Let $\mathbf{R}$ and $\mathbf{C}$ denote the fields of the real and complex numbers, respectively. An extension of this problem to the matrix case is as follows.

Given: a subset $\vartheta \subset \mathbf{C}$, points $z_1, \ldots, z_L$ in ϑ, rational $1 \times p$ row vector functions $v_1(z), \ldots, v_L(z)$ with $v_k(z_k) \neq 0$ for all k, and rational $1 \times m$ row vectors $w_1(z), \ldots, w_L(z)$.

Find: (at least one or all) $p \times m$ rational matrix functions $G(z)$ with no poles in ϑ which satisfy the *tangential* interpolation conditions

$$\frac{d^j}{dz^j}\{v_k(z)G(z)\}\bigg|_{z=z_k} = \frac{d^j}{dz^j} w_k(z)\bigg|_{z=z_k} \tag{1.1}$$

for $0 \leq j \leq N_k$, $1 \leq k \leq L$.

This problem is known as the tangential *Lagrange–Sylvester* rational interpolation problem. One approach to finding a solution is to reduce the problem to a system of

*Received by the editors January 7, 2005; accepted for publication (in revised form) by P. Benner October 26, 2006; published electronically March 19, 2007.
http://www.siam.org/journals/simax/29-2/62217.html

†Department of Electrical and Electronics Engineering, Anadolu University, 26470 Eskişehir, Turkey (huakcay@anadolu.edu.tr, semihaturkay@anadolu.edu.tr).

independent scalar problems, which is not interesting from the viewpoint of matrix interpolation theory. In addition, a minimal realization can be obtained only after the elimination of unobservable or/and uncontrollable modes. The contour integral version of this problem is treated in the comprehensive work [6]. The *bitangential* or *bidirectional* version is studied, for example, in [14, 15, 7, 4]. Related problems are the nonhomogenous interpolation problem with metric constraints, as in the various types of Nevanlinna–Pick interpolation and its generalizations [10, 20], 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 [17, 1, 6, 27, 28]. Further applications of interpolation theory to control and systems theory and estimation are presented in [6, 13, 29].

Prior work on the unconstrained tangential interpolation problem has been largely carried out by Ball, Gohberg, and Rodman [6, 7]. The solvability issues of the interpolation problem, i.e., the existence and the uniqueness of the solutions, have been analyzed in [8] by using a *residual interpolation* framework. A more direct algebraic approach in [11] shows that solving a tangential interpolation problem is equivalent to solving a matrix Padé approximation problem with Taylor coefficients obeying a set of linear constraints. In [1, 2, 3, 4], the tangential interpolation problem above was studied using a tool called the Löwner matrix. In [4], the problem of finding admissible degrees of complexity of the solutions to the above interpolation problem, that is, finding all positive integers n for which there exits an interpolant with McMillan degree n, and the problem of parameterizing all solutions for a given admissible degree of complexity were investigated. Clearly, the solutions of minimal complexity are of special interest.

The main result in [7] states that the family of rational matrix functions satisfying (1.1) can be parameterized in terms of a certain *linear fractional map*. First, the interpolation data is translated into a so-called *left null pair* that describes the zero structure of a $(p + m) \times (p + m)$ *resolvent matrix*. The computation of the resolvent matrix requires that the solution of a particular Sylvester equation be invertible. The details can be found in [6]. In [11], a recursive method for computing the resolvent matrix as a product of elementary first-order rational matrix functions is presented. This scheme allows recursive updating of the resolvent matrix whenever a new interpolation point is added to the input data. In the special case when the resolvent matrix is in *column-reduced* form, it is possible to extract the admissible degrees of complexity as well as the minimal degree of complexity from the linear fractional parameterization formula. The resolvent matrix obtained by an unconstrained algorithm can be transformed into column-reduced form via a sequence of elementary unimodular transformations [16]. A detailed algorithm for the construction of a column-reduced rational matrix function from a given *null-pole triple* is given in [9]. This algorithm is not recursive, whereas in [11] a column-reduced transfer function is recursively obtained.

In this paper, we present a numerically efficient algorithm for solving the unconstrained tangential interpolation problem formulated above. This algorithm is inspired by the recent work on the frequency domain subspace-based identification [23, 24, 25, 33]. The solvability conditions for the proposed algorithm are simple, and depend only on the total multiplicities of the interpolation points. The resulting interpolating function is in the minimal state-space form. To this date, interpolation properties of the subspace-based methods have not been investigated in the generality of this paper. Only in [24] was an interpolation result obtained for uniformly spaced data on the unit circle of the complex plane. The problem of curve fitting is also closely related to the interpolation problem. The use of the frequency domain subspace-based methods for curve fitting is briefly described in [22].

Let us reformulate the tangential interpolation problem described above in terms of system properties. More precisely, let us consider a multi-input/multi-output, linear-time invariant, discrete-time system represented by the state-space equations

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k), \\ y(k) &= Cx(k) + Du(k), \end{aligned} \tag{1.2}$$

where $x(k) \in \mathbf{R}^n$ is the state and $u(k) \in \mathbf{R}^m$ and $y(k) \in \mathbf{R}^p$ are, respectively, the input and the output of the system. The transfer function of the system (1.2) denoted by $G(z)$ is computed as

$$G(z) = D + C(zI_n - A)^{-1}B, \tag{1.3}$$

where I_n is the $n \times n$ identity matrix. We assume that that the system (1.2) is stable and the pairs (A, B) and (C, A) are *controllable* and *observable*, respectively. The stability of (1.2) means that $G(z)$ is a *proper rational* matrix that is analytic and bounded in the region $\vartheta = \{z \in \mathbf{C} : |z| \geq 1\}$, and both the controllability and the observability of the pairs (A, B) and (C, A) mean 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 $z_k \in \vartheta$,

$$\left.\frac{d^j}{dz^j} G(z)\right|_{z=z_k} = w_{kj}, \qquad j = 0, 1, \ldots, N_k, \quad k = 1, 2, \ldots, L. \tag{1.4}$$

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

Clearly, (1.4) is a special case of (1.1) with suitably selected left vectors $v_k(z)$ and nodes z_k. A subspace-based algorithm handling the tangential-type constraints (1.1) as well can be derived along the same lines of the proposed algorithm. The minimality and the uniqueness of the interpolant are the parts of the problem formulation. What is left unanswered is a condition on the number of the interpolation nodes, counting multiplicities. It is also clear that, if it exists, the subspace-based solution is a minimal interpolating function in the set of all possible solutions.

The proposed interpolation scheme is particularly useful when the samples of $G(z)$ and its derivatives are corrupted by noise and the amount of data is large with respect to n. In the noisy case, most interpolation schemes deliver state-space realizations with McMillan degrees tending to infinity as the amount of data grows unboundedly; thus such schemes are sensitive to inaccuracies in the interpolation data. Since our algorithm is subspace-based, it inherits robustness properties of the subspace-based identification algorithms. In particular, there is no need for explicit model parameterization, and this algorithm is computationally efficient since it uses numerically robust QR factorization and the singular value decomposition. In the paper, we also consider subspace-based system identification with interpolation constraints.

Note that a given interpolation problem on the right half complex plane can be converted to an interpolation problem on the unit disk by using the Möbius transformation:

$$s = \psi(z) \triangleq \lambda \frac{z-1}{z+1} \qquad (\lambda > 0). \tag{1.5}$$

We omit the details.

2. Subspace-based interpolation algorithm. We begin by taking the z-transform of (1.2),

$$\begin{aligned} zX(z) &= AX(z) + BU(z), \\ Y(z) &= CX(z) + DU(z), \end{aligned} \tag{2.1}$$

assuming $x(0) = 0$, where $X(z)$, $Y(z)$, and $U(z)$ denote respectively the z-transforms of $x(k)$, $y(k)$, and $u(k)$ defined by

$$U(z) \triangleq \sum_{k=0}^{\infty} u(k)\, z^{-k}. \tag{2.2}$$

Let $X_j(x)$ be the resulting state z-transform when

$$u(k) = \begin{cases} e_j, & k = 0, \\ 0, & \text{otherwise,} \end{cases}$$

where e_j denotes the unit vector in $\mathbf{R}^m$ with 1 on the jth position and 0 elsewhere. By defining the compound state z-transform matrix,

$$X_{\mathrm{C}}(z) \triangleq [X_1(z) \;\; X_2(z) \;\; \cdots \;\; X_m(z)]\,, \tag{2.3}$$

$G(z)$ can implicitly be described as

$$G(z) = CX_{\mathrm{C}}(z) + D \tag{2.4}$$

with

$$zX_{\mathrm{C}}(z) = AX_{\mathrm{C}}(z) + B. \tag{2.5}$$

By recursive use of (2.5), we obtain the relation

$$z^k X_{\mathrm{C}}(z) = A^k X_{\mathrm{C}}(z) + \sum_{j=0}^{k-1} A^{k-1-j} B z^j, \qquad k \geq 1. \tag{2.6}$$

Multiplying both sides of (2.6) with C and using (2.4), we get

$$z^k G(z) = CA^k X_{\mathrm{C}}(z) + Dz^k + \sum_{j=0}^{k-1} CA^{k-1-j} B z^j, \qquad k \geq 1. \tag{2.7}$$

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

$$g_k = \begin{cases} D, & k = 0, \\ CA^{k-1}B, & k \geq 1. \end{cases} \tag{2.8}$$

Thus, from (2.4), (2.7), and (2.8),

$$z^k G(z) = CA^k X_{\mathrm{C}}(z) + \sum_{j=0}^{k} g_{k-j}\, z^j, \qquad k \geq 0. \tag{2.9}$$

Hence from (2.9),

$$\begin{bmatrix} G(z) \\ zG(z) \\ \vdots \\ z^{q-1}G(z) \end{bmatrix} = \mathcal{O}_q X_{\mathrm{C}}(z) + \Gamma_q \begin{bmatrix} I_m \\ zI_m \\ \vdots \\ z^{q-1}I_m \end{bmatrix}, \tag{2.10}$$

where

$$\mathcal{O}_q \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{q-1} \end{bmatrix}, \tag{2.11}$$

$$\Gamma_q \triangleq \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{q-1} & g_{q-2} & \cdots & g_0 \end{bmatrix}. \tag{2.12}$$

For later use, let us write (2.10) in a compact form. The matrix $\mathcal{O}_q$ is known as the *extended observability matrix* and has full rank n if (A, C) 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

$$E \otimes F \triangleq \begin{bmatrix} E_{11}F & E_{12}F & \cdots & E_{1n}B \\ E_{21}F & E_{22}F & \cdots & E_{2n}F \\ \vdots & \vdots & \ddots & \vdots \\ E_{m1}F & E_{m2}F & \cdots & E_{mn}F \end{bmatrix} \in \mathbf{C}^{mp\times nq}. \tag{2.13}$$

Let

$$\mathcal{Z}_q(z) \triangleq \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{q-1} \end{bmatrix}, \tag{2.14}$$

$$\mathcal{J}_{q,2} \triangleq \begin{bmatrix} 0 & & \cdots & & 0 \\ 1 & 0 & & & \\ 0 & 1 & 0 & & \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & 1 & 0 \end{bmatrix} \in \mathrm{R}^{q\times q}. \tag{2.15}$$

By a slight abuse of notation, let $\mathcal{J}_{q,1}$ denote the $q \times q$ identity matrix I_q. 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 $\mathcal{J}_{q,1}$ and $J^0_{q,2} = I_q$. Note the following relations:

$$\mathcal{J}_{q,j} = \begin{cases} \mathcal{J}_{q,2}^{j-1}, & j \leq q \\ 0, & j > q. \end{cases} \tag{2.16}$$

Thus, the lower triangular block Toeplitz matrix in (2.12) can be written as

$$\Gamma_q = \sum_{j=0}^{q-1} \mathcal{J}_{q,1+j} \otimes g_j. \tag{2.17}$$

Hence, from (2.11)–(2.17) we arrive at the following compact expression for (2.10):

$$\mathcal{Z}_q(z) \otimes G(z) = \mathcal{O}_q X_{\mathrm{C}}(z) + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j] \, [\mathcal{Z}_q(z) \otimes I_m] \, . \tag{2.18}$$

This equation forms the basis of the frequency domain subspace-based identification algorithms [24, 23]. In subspace-based identification algorithms, $\mathcal{Z}_q(z) \otimes G(z)$ and the right-hand side of (2.18) 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. We will adapt the same strategy.

First, we differentiate both sides of (2.18) l times with respect to z:

$$\begin{aligned} \frac{d^l}{dz^l} H_q(z) &= \sum_{j=0}^{l} \begin{pmatrix} l \\ j \end{pmatrix} \frac{d^j}{dz^j} \mathcal{Z}_q(z) \otimes \frac{d^{l-j}}{dz^{l-j}} G(z) \\ &= \mathcal{O}_q \frac{d^l}{dz^l} X_{\mathrm{C}}(z) + \sum_{j=0}^{q-1} [\mathcal{J}_{q,2}^j \otimes g_j] \left[\frac{d^l}{dz^l} \mathcal{Z}_q(z) \otimes I_m \right], \qquad l \geq 0, \end{aligned} \tag{2.19}$$

where

$$H_q(z) \triangleq \mathcal{Z}_q(z) \otimes G(z). \tag{2.20}$$

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

$$\mathcal{H}_k \triangleq \left[H_q(z) \;\; \frac{d}{dz} H_q(z) \;\; \cdots \;\; \frac{d^{N_k}}{dz^{N_k}} H_q(z) \right]_{z=z_k}, \qquad k = 1, \ldots, L. \tag{2.21}$$

Using the right-hand side of the first equality in (2.19), let us derive a compact expression for $\mathcal{H}_k$ in terms of the elementary matrices

$$\mathcal{D}_k \triangleq \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & & \\ & & 0 & \cdots & \\ \vdots & & & \ddots & N_k \\ 0 & & & \cdots & 0 \end{bmatrix} \in \mathbf{R}^{(N_k+1) \times (N_k+1)} \tag{2.22}$$

and

$$\mathcal{W}_k \triangleq \left[\mathcal{Z}_q(z) \;\; \frac{d}{dz} \mathcal{Z}_q(z) \;\; \cdots \;\; \frac{d^{N_k}}{dz^{N_k}} \mathcal{Z}_q(z) \right]_{z=z_k}, \qquad k = 1, \ldots, L, \tag{2.23}$$

as follows:

$$\begin{aligned}
\mathcal{H}_k = &\left[\mathcal{Z}_q(z)\ \ \frac{\mathrm{d}}{\mathrm{d}z}\mathcal{Z}_q(z)\ \ \frac{d^2}{dz^2}\mathcal{Z}_q(z)\ \ \cdots\ \ \frac{d^{N_k}}{dz^{N_k}}\mathcal{Z}_q(z)\right]_{z=z_k} \otimes G(z_k) \\
&+\left[0\ \ \mathcal{Z}_q(z)\ \ 2\frac{d}{\mathrm{d}z}\mathcal{Z}_q(z)\ \ \cdots\ \ \begin{pmatrix} N_k \\ 1 \end{pmatrix} \frac{d^{N_k-1}}{dz^{N_k-1}}\mathcal{Z}_q(z)\right]_{z=z_k} \otimes \frac{d}{dz}G(z_k) \\
&+\left[0\ \ 0\ \ \mathcal{Z}_q(z)\ \ \cdots\ \ \begin{pmatrix} N_k \\ 2 \end{pmatrix} \frac{d^{N_k-2}}{dz^{N_k-2}}\mathcal{Z}_q(z)\right]_{z=z_k} \otimes \frac{d^2}{dz^2}G(z_k) + \cdots \\
&+ [0\ \ 0\ \ 0\ \ \cdots\ \ \mathcal{Z}_q(z)]_{z=z_k} \otimes \frac{d^{N_k}}{dz^{N_k}}G(z_k) \\
= &\ \mathcal{W}_k \otimes G(z_k) + [\mathcal{W}_k\mathcal{D}_k] \otimes \frac{d}{dz}G(z_k) + \frac{1}{2!}[\mathcal{W}_k\mathcal{D}_k^2] \otimes \frac{d^2}{dz^2}G(z_k) \\
&+ \frac{1}{N_k!}[\mathcal{W}_k\mathcal{D}_k^{N_k}] \otimes \frac{d^{N_k}}{dz^{N_k}}G(z_k).
\end{aligned}$$

Note that $\mathcal{D}_k^j = 0$ for all $j > N_k$. Hence,

$$\mathcal{H}_k = \sum_{j=0}^{N_k} \frac{1}{j!}[\mathcal{W}_k\,\mathcal{D}_k^j] \otimes w_{kj}, \qquad k = 1, \ldots, L. \tag{2.24}$$

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

$$\mathcal{T}_q \triangleq \begin{bmatrix} 0! & 0 & \cdots & 0 \\ 0 & 1! & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (q-1)! \end{bmatrix} \in \mathbf{R}^{q\times q}. \tag{2.25}$$

Then, it is easy to verify that

$$\frac{d^l}{dz^l}\mathcal{Z}_q(z) = \mathcal{T}_q\mathcal{J}_{q,2}^l\mathcal{T}_q^{-1}\mathcal{Z}_q(z), \qquad l \geq 0. \tag{2.26}$$

Hence from (2.23) and (2.26),

$$\mathcal{W}_k = \mathcal{T}_q\left[I_q\ \ \mathcal{J}_{q,2}\ \ \cdots\ \ \mathcal{J}_{q,2}^{N_k}\right]\left[I_{N_k+1} \otimes \mathcal{T}_q^{-1}\mathcal{Z}_q(z_k)\right], \qquad k = 1, \ldots, L. \tag{2.27}$$

An alternative compact expression for $\mathcal{H}_k$ is obtained by evaluating the right-hand side of the second equality in (2.19) for $l = 0, \ldots, N_k$, $k = 1, \ldots, L$, and augmenting the similar terms in compound matrices as follows:

$$\mathcal{H}_k = \mathcal{O}_q\,\mathcal{X}_k + \sum_{j=0}^{q-1}[\mathcal{J}_{q,2}^j \otimes g_j]\,[\mathcal{W}_k \otimes I_m], \qquad k = 1, \ldots, L, \tag{2.28}$$

where

$$\mathcal{X}_k \triangleq \left[X_{\mathrm{C}}(z)\ \ \frac{d}{dz}X_{\mathrm{C}}(z)\ \ \cdots\ \ \frac{d^{N_k}}{dz^{N_k}}X_{\mathrm{C}}(z)\right]_{z=z_k}, \qquad k = 1, \ldots, L. \tag{2.29}$$

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

$$\mathcal{H} \triangleq [\mathcal{H}_1\ \mathcal{H}_2\ \cdots\ \mathcal{H}_L], \tag{2.30}$$

$$\mathcal{X} \triangleq [\mathcal{X}_1\ \mathcal{X}_2\ \cdots\ \mathcal{X}_L], \tag{2.31}$$

$$\mathcal{W} \triangleq [\mathcal{W}_1\ \mathcal{W}_2\ \cdots\ \mathcal{W}_L]. \tag{2.32}$$

Hence,

$$\mathcal{H} = \mathcal{O}_q\,\mathcal{X} + \sum_{j=0}^{q-1}[\mathcal{J}_{q,2}^{j} \otimes g_j]\,[\mathcal{W} \otimes I_m], \tag{2.33}$$

where $\mathcal{H}$ and $\mathcal{W}$ are computed from the problem data $\{z_k, \{w_{kj}\}_{j=0}^{N_k}\}_{k=1}^{L}$ by the formulae (2.30), (2.32), (2.27), (2.24), (2.22), (2.14), (2.15), (2.25). 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 (2.33) into a relation involving only real valued matrices:

$$\widehat{\mathcal{H}} = \mathcal{O}_q\,\widehat{\mathcal{X}} + \sum_{j=0}^{q-1}[\mathcal{J}_{q,2}^{j} \otimes g_j]\,\mathcal{F}, \tag{2.34}$$

where

$$\widehat{\mathcal{H}} \triangleq [\mathrm{Re}\mathcal{H}\;\; \mathrm{Im}\mathcal{H}]\,, \tag{2.35}$$

$$\mathcal{F} \triangleq [\mathrm{Re}\mathcal{W}\;\; \mathrm{Im}\mathcal{W}] \otimes I_m, \tag{2.36}$$

$$\widehat{\mathcal{X}} \triangleq [\mathrm{Re}\mathcal{X}\;\; \mathrm{Im}\mathcal{X}]\,. \tag{2.37}$$

Let z^* denote the complex conjugate of z. When $z_k \in \mathbf{R}$, from (2.14) we have $\mathcal{Z}_q(z_k) \in \mathbf{R}^q$. This, by (2.27), implies that $\mathcal{W}_k \in \mathbf{R}^{q\times(N_k+1)}$. From (2.5),

$$X_{\mathrm{C}}(z) = (zI_n - A)^{-1}B. \tag{2.38}$$

Then, from (2.4), (2.38), and (2.29), it follows that $\mathcal{X}_k \in \mathbf{R}^{n\times m(N_k+1)}$ and, for all $j = 0, \ldots, N_k$, $w_{kj} \in \mathbf{R}^{p\times m}$ whenever $z_k \in \mathbf{R}$. Thus, whenever $z_k \in \mathbf{R}$ from (2.24) we have $\mathcal{H}_k \in \mathbf{R}^{pq\times m(N_k+1)}$. Hence, the imaginary parts of $\mathcal{H}_k$, $\mathcal{F}$, and $\mathcal{X}_k$ are all zero, and they need not be included in (2.35)–(2.37) if $z_k \in \mathbf{R}$; without loss of generality, we will assume this in what follows. Let

$$N \triangleq \sum_{k:z_k\in\mathbf{R}}(N_k+1) + \sum_{k:z_k\in\mathbf{C}-\mathbf{R}} 2(N_k+1). \tag{2.39}$$

Then, $\widehat{\mathcal{H}} \in \mathbf{R}^{pq\times mN}$, $\mathcal{F} \in \mathbf{R}^{mq\times mN}$, and $\widehat{\mathcal{X}} \in \mathbf{R}^{n\times mN}$.

2.1. Projection onto the observability range space. Let $\mathcal{F}^{\perp}$ be the projection matrix onto the null space of $\mathcal{F}$ given by

$$\mathcal{F}^{\perp} \triangleq I_{mN} - \mathcal{F}^T(\mathcal{F}\mathcal{F}^T)^{-1}\mathcal{F}, \tag{2.40}$$

where $\mathcal{F}^T$ denotes the transpose of $\mathcal{F}$. The summand in (2.34) is cancelled for all j when multiplied from right by $\mathcal{F}^{\perp}$. Thus,

$$\widehat{\mathcal{H}}\mathcal{F}^{\perp} = \mathcal{O}_q\,\widehat{\mathcal{X}}\mathcal{F}^{\perp}. \tag{2.41}$$

A numerically efficient way of forming $\widehat{\mathcal{H}}\mathcal{F}^{\perp}$ is to use the QR-factorization

$$\begin{bmatrix} \mathcal{F} \\ \widehat{\mathcal{H}} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix}. \tag{2.42}$$

A simple derivation yields

$$\widehat{\mathcal{H}}\mathcal{F}^{\perp} = R_{22}Q_2^T, \tag{2.43}$$

and it suffices to use $R_{22} \in \mathbf{R}^{pq\times m(N-q)}$ in the extraction of the observability range space since Q_2^T is a matrix of full rank.

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 of the row spaces of $\mathcal{F}$ and $\widehat{\mathcal{X}}$ be empty. In the following, we present sufficient conditions in terms of the data and the system.

LEMMA 2.1. *Let $\widehat{\mathcal{X}}$, $\mathcal{F}$, and N be as in* (2.37), (2.36), *and* (2.39), *respectively. Suppose that $N \geq q+n$ and the eigenvalues of A do not coincide with the distinct complex numbers z_k. Then,*

$$\operatorname{rank}\begin{bmatrix}\mathcal{F}\\ \widehat{\mathcal{X}}\end{bmatrix} = qm+n \qquad \Longleftrightarrow \qquad (A,B) \text{ controllable pair.} \tag{2.44}$$

Proof. 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

$$[\alpha_0 \ \ \cdots \ \ \alpha_{q-1} \ \ \beta] \neq 0 \tag{2.45}$$

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

$$[\alpha_0 \ \ \cdots \ \ \alpha_{q-1} \ \ \beta]\begin{bmatrix}\mathcal{W}\otimes I_m\\ \mathcal{X}\end{bmatrix} = 0. \tag{2.46}$$

From (2.32), (2.23), and (2.31), (2.29), equation (2.46) holds if and only if

$$[\alpha_0 \ \ \cdots \ \ \alpha_{q-1} \ \ \beta]\frac{d^j}{dz^j}\begin{bmatrix}\mathcal{Z}_q(z)\otimes I_m\\ \mathcal{X}_{\mathrm{C}}(z)\end{bmatrix}_{z=z_k} = 0, \quad 0 \leq j \leq N_k,\ k=1,\dots,L,$$

$$\Updownarrow$$

$$\left.\frac{d^j}{dz^j}E(z)\right|_{z=z_k} = 0, \quad 0 \leq j \leq N_k,\ k=1,\dots,L, \tag{2.47}$$

where

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

Equation (2.47) implies that for each 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 of $E(z)$ if and only if z_k^* is also a zero of $E(z)$. Therefore, $E(z)$ happens to have a total number of N zeros counting multiplicities. However, the elements of $E(z)$ have numerator degrees not exceeding $n+q-1$. Hence, any element of $E(z)$ cannot have N zeros. 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 result follows from the fact that $\beta(zI_n-A)^{-1}B$ is analytic and has a zero at $z=\infty$; hence it is 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 and real vector x,

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

Since all the eigenvalues of A are inside the unit circle, none of them coincide with any of z_k. Thus, by applying Lemma 2.1, 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 $\widehat{\mathcal{H}}\mathcal{F}^{\perp}$,

$$\begin{aligned} \widehat{\mathcal{H}}\mathcal{F}^{\perp} &= \widehat{U}\widehat{\Sigma}\widehat{V}^T \\ &= \begin{bmatrix} \widehat{U}_s & \widehat{U}_o \end{bmatrix} \begin{bmatrix} \widehat{\Sigma}_s & 0 \\ 0 & \widehat{\Sigma}_o \end{bmatrix} \begin{bmatrix} \widehat{V}_s^T \\ \widehat{V}_o^T \end{bmatrix}, \end{aligned} \tag{2.48}$$

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

$$\begin{aligned} \widehat{A} &= (J_1\widehat{U}_s)^{\dagger} J_2 \widehat{U}_s, \\ \widehat{C} &= J_3 \widehat{U}_s, \end{aligned} \tag{2.49}$$

where

$$J_1 = \begin{bmatrix} I_{(q-1)p} & 0_{(q-1)p\times p} \end{bmatrix}, \tag{2.50}$$

$$J_2 = \begin{bmatrix} 0_{(q-1)p\times p} & I_{(q-1)p} \end{bmatrix}, \tag{2.51}$$

$$J_3 = \begin{bmatrix} I_p & 0_{p\times(q-1)p} \end{bmatrix}, \tag{2.52}$$

$0_{i\times j}$ is the $i \times j$ zero matrix, and $X^{\dagger} = (X^TX)^{-1}X^T$ is the Moore–Penrose pseudoinverse of the full column rank matrix X. Provided that (C, A) is an observable pair, the pseudoinverse in (2.49) exists if and only if $q > n$. Therefore, in order to apply the lemma it suffices to let $q = n + 1$. In this case, we have the sole requirement $N > 2n$ with N defined by (2.39). From Lemma 2.1, it follows that $\widehat{A}$ and $\widehat{C}$ defined in (2.49) are related to A and C in (1.2) by

$$\begin{aligned} \widehat{A} &= T^{-1}AT, \\ \widehat{C} &= CT \end{aligned} \tag{2.53}$$

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

As noted before, in (2.48) $\widehat{\mathcal{H}}\mathcal{F}^{\perp}$ can be replaced with R_{22}.

2.2. 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

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

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

$$\frac{d^j}{dz^j}G(z) = \delta_{0j}\,D + (-1)^j j!\,C(zI_n - A)^{-j-1}B, \qquad j \geq 0, \tag{2.54}$$

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

$$\mathcal{G}_k \triangleq \begin{bmatrix} w_{k0} \\ w_{k1} \\ \vdots \\ w_{kN_k} \end{bmatrix}, \qquad k = 1, \ldots, L, \tag{2.55}$$

and

$$\mathcal{G} \triangleq \begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \vdots \\ \mathcal{G}_L \end{bmatrix}. \tag{2.56}$$

Observe from (2.54) 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

$$\widehat{B}, \widehat{D} = \arg\min_{B,D} \left\| \widehat{\mathcal{G}} - \widehat{\mathcal{Y}} \begin{bmatrix} B \\ D \end{bmatrix} \right\|_F^2, \tag{2.57}$$

where

$$\|X\|_F \triangleq \left[\sum_k \sum_s |x_{ks}|^2 \right]^{1/2}$$

is the *Frobenius norm*,

$$\widehat{\mathcal{G}} \triangleq \begin{bmatrix} \mathrm{Re}\mathcal{G} \\ \mathrm{Im}\mathcal{G} \end{bmatrix} \in \mathbf{R}^{pN \times m}, \tag{2.58}$$

$$\widehat{\mathcal{Y}} \triangleq \begin{bmatrix} \mathrm{Re}\mathcal{Y} \\ \mathrm{Im}\mathcal{Y} \end{bmatrix} \in \mathbf{R}^{pN \times (n+p)}, \tag{2.59}$$

and

$$\mathcal{Y}_k \triangleq \begin{bmatrix} C(z_k I_n - A)^{-1} & I_p \\ -C(z_k I_n - A)^{-2} & 0 \\ \vdots & \\ (-1)^{N_k} N_k!\, C(z_k I_n - A)^{-N_k - 1} & 0 \end{bmatrix}, \tag{2.60}$$

$$\mathcal{Y} \triangleq \begin{bmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \\ \vdots \\ \mathcal{Y}_L \end{bmatrix}, \tag{2.61}$$

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

LEMMA 2.2. *Let N and $\widehat{\mathcal{Y}}$ be as in* (2.39) *and* (2.59), *respectively. Suppose that $N > n$ and the eigenvalues of A do not coincide with the distinct complex numbers z_k. Then,*

$$\mathrm{rank}\widehat{\mathcal{Y}} = p + n \qquad \Longleftrightarrow \qquad (C, A) \text{ observable pair.} \tag{2.62}$$

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

$$\mathcal{Y}\begin{bmatrix} B \\ D \end{bmatrix} = 0 \iff \frac{d^j}{dz^j}G(z)\bigg|_{z=z_k} = 0, \quad 0 \le j \le N_k,\ k = 1, \dots, L.$$

As in the proof of Lemma 2.1, 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 $N \le n$. $\square$

Thus, from (2.53) and Lemma 2.2, if $N \ge q + n$ and $q > n$, we have

$$\begin{aligned} \widehat{B} &= T^{-1}B, \\ \widehat{D} &= D. \end{aligned} \tag{2.63}$$

Moreover,

$$\widehat{G}(z) \triangleq \widehat{C}(zI_n - \widehat{A})^{-1}\widehat{B} + \widehat{D} = G(z). \tag{2.64}$$

2.3. Solvability conditions. By picking $q = n + 1$ in the subspace-based algorithm developed above, we obtain a sufficient condition for the interpolation of $G(z)$ from its noise-free samples and derivatives evaluated at L distinct points in ϑ as $N \ge 2n+1$, where N is defined by (2.39). This condition turns out to be a necessary condition for the interpolation of $G(z)$, as demonstrated next by a simple example.

Consider an nth-order stable single-input/single-output system represented by the transfer function

$$G(z) = \frac{b_0 z^n + b_1 z + \cdots + b_n}{z^n + a_1 z + \cdots + a_n}. \tag{2.65}$$

We are to determine $2n + 1$ unknown real coefficients $a_1, \dots, a_n, b_0, \dots, b_n$ from the evaluations of $G(z)$ and its derivatives at a given set of distinct frequencies $z_k \in \vartheta$. Let N be as in (2.39).

Let us first assume in (1.4) that $N_k = 0$ and $z_k \in \mathbf{C} - \mathbf{R}$ for all k; i.e., the interpolation nodes are simple and purely complex numbers. Then, $N = 2L$. With $q = n + 1$, the subspace-based algorithm delivers a minimal realization of $G(z)$, provided that $2L \ge 2n + 1$. This condition is satisfied by choosing $L = n + 1$. Clearly, this is the least amount of data one could use to interpolate an arbitrary nth-order system, as can directly be verified by writing $2L$-linear equations down from (1.4) and (2.65) to determine the unknowns $a_1, \dots, a_n, b_0, \dots, b_n$. Notice that if some interpolation nodes have multiplicities, then the resulting equations become nonlinear in $a_1, \dots, a_n, b_0, \dots, b_n$.

Now, as a special case, let us consider the situation that all z_k are on the unit circle excluding the points ± 1. Thus, Algorithm 2.1 recovers nth-order stable systems from $n + 1$ noise-free frequency response measurements, excluding the frequencies 0 and π. If the frequencies contain 0, from (2.39) we then have $N = 2L - 1$. Hence, with $q = n + 1$ selected, we must have $2L - 1 \ge 2n + 1$, which is fulfilled by letting $L = n + 1$. If, in addition, the frequencies contain π as well, we end up with the interpolation condition $L = n + 2$. The last conclusion extends an interpolation result in [24] derived for the uniformly spaced frequencies case to the nonuniformly spaced frequencies case.

It is easy to see, for example, by the partial fraction expansion or similar techniques, that these results hold for multi-input/multi-output systems with multiple interpolation nodes as well. Therefore, Algorithm 2.1 is capable of using a minimum amount of the frequency domain data for the Lagrange–Sylvester interpolation of stable systems.

2.4. Summary of the subspace-based interpolation algorithm. Let us summarize the interpolation algorithm in the following.

ALGORITHM 2.1. *Subspace-based interpolation algorithm.*

1. *Given the data* (1.4)*, compute the matrices* $\widehat{\mathcal{H}}$ *and* $\mathcal{F}$ *defined by* (2.35) *and* (2.36) *through* (2.30)*,* (2.32)*,* (2.24)*,* (2.27)*,* (2.22)*,* (2.25)*,* (2.15)*, and* (2.14)*.*
2. *Perform the QR-factorization in* (2.42)*.*
3. *Calculate the singular value decomposition in* (2.48) *with* $\widehat{\mathcal{H}}\mathcal{F}^{\perp}$ *replaced by* R_{22} *defined in* (2.42)*.*
4. *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.*
5. *With* J_1*,* J_2*, and* J_3 *defined by* (2.50)–(2.52)*, calculate* $\widehat{A}$ *and* $\widehat{C}$ *from* (2.49)*.*
6. *Solve the least-squares problem* (2.57) *for* $\widehat{B}$ *and* $\widehat{D}$*, where* $\widehat{\mathcal{G}}$ *and* $\widehat{\mathcal{Y}}$ *are defined by* (2.58) *and* (2.59) *through* (2.60)–(2.61) *and* (2.55)–(2.56)*.*

Clearly, $\widehat{\Sigma}_o = 0$ in (2.48) when the data are not corrupted by noise, the system that has generated the data is of McMillan degree n, $N \geq q + n$, and $q > n$. As we stated earlier, Algorithm 2.1 produces a minimal stable realization of the interpolant, given that the latter exits. In most interpolation problems, the existence and the uniqueness questions are easily settled, and the construction of a solution (or all solutions) with certain properties such as the McMillan degree constraints, in particular minimality, remains a difficult one. The algorithm outlined above is straightforward to implement. In the implementation of the algorithm, it suffices to let $q = n + 1$ and $N = 2n + 1$, where N is defined by (2.39). The system order, if unknown a priori, can be determined in step 4 of Algorithm 2.1 from the inspection of the singular values. This process also reveals redundancies in the data. Numerically, the most expensive step in the algorithm is the singular value decomposition of R_{22}. Notice with $q = n + 1$ and $N = 2n + 1$ selected, that $R_{22} \in \mathbf{R}^{p(n+1)\times mn}$.

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

THEOREM 2.3. *Consider Algorithm* 2.1 *with the data in* (1.4) *originating from a discrete-time stable system of order* n*. Let* N *be as in* (2.39)*. If* $N \geq q + n$ *and* $q > n$*, then the quadruplet* $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ *is a minimal realization of* $G(z)$*.*

2.5. Discussion. In the rest of this section, we will briefly comment on the similarities and the differences between Algorithm 2.1 and the Löwner matrix–based approach [1].

The most striking difference between the methods appears to be the formation of data matrices. In [1], elements of a Löwner matrix are computed by taking partial derivatives of the divided differences $[G(z) - G(s)]/(z - s)$ evaluated at $z = z_k$ and $s = z_l$, where the number of the derivatives is determined by the particular choice of the (block) row and column sets and the multiplicities of the nodes. If z_k equals z_l, a limiting process has to be used to define that particular element. It is required that the numbers of the chosen block rows and columns add up to N. The elements of $\mathcal{H}$ in the proposed algorithm, on the other hand, consist of linear combinations of the derivatives of the products $z^l G(z)$ evaluated at $z = z_k$, where for each k, l satisfies $0 \leq l \leq N_k$. A simple transformation that relates $\mathcal{H}$ to a Löwner matrix does not seem possible unless all the z_k are the same, in which case the problem solved reduces to a conventional partial realization problem. In the latter case, notice that this link is provided by the bilinear map (1.5).

Both algorithms rely on the factorization of the data matrices discussed above as a product of two matrices which are directly related to the observability and controllability concepts. In [1], the Löwner matrix is expressed as a product of the so-called generalized observability and the controllability matrices, whereas in the proposed algorithm this relation is recovered after some projections. In fact, the proofs of Lemmas 2.1, 2.2, and 3.1R in [1] use the same ideas.

The most striking similarity between the algorithms is the condition $N > 2n$. It should be noted that the stability assumption is not essential in the formulation of the interpolation problem, since the data are already assumed to originate from a finite-dimensional dynamical system with a complexity bounded above and the number of the nodes is finite. This assumption is necessary in an identification setup. Without the knowledge that the data have originated from a dynamical system with a complexity bounded above, the condition $N > 2n$ is precisely one of the requirements for the existence of a unique minimal-order interpolating rational matrix [1]. In addition to this requirement, there is also a more stringent rank condition captured in Assumption 4.1 in [1]. Thus, both algorithms operate under the same conditions which assure the existence of a unique minimal interpolating rational matrix. We have not addressed the properness issue in this paper due to our standing assumption on the origins of the data. Again, without the knowledge of the origins of the data, one has to secure that the solution of the interpolation problem is a proper transfer function. The properness is guaranteed by Assumption 4.2 in [1]. It is also noted there that this assumption can be eliminated by means of a suitably chosen bilinear transformation.

The Löwner matrix–based and proposed algorithms cannot be directly applied when there does not exist a unique minimal interpolating function and the data are not scalar. This may happen either in the presence of noise which corrupts transfer function evaluations or when the true dynamics is of higher dimension. The problem is then to find the admissible degrees of complexity, i.e., those positive integers n for which there exist solutions $G(z)$ to the interpolation problem (1.4) with $deg\, G = n$, and to construct all corresponding solutions for a given admissible degree n. This problem is known as the partial realization problem. If the original data do not satisfy the criterion for the existence of a unique minimal interpolating function, one needs to add interpolation data until the criterion becomes satisfied. The fact that the data can be found so that the increase in degree is finite is nontrivial. The added data will necessarily drive up the degree of the interpolating transfer function. In the scalar case, dealt with in [2], the way this can be done is set out and is rather complicated. The multivariable case is studied in [4] using the generating system approach. While [4] gives the theory behind the determination of the minimal McMillan degree and all admissible degrees, the current paper and [1] provide the theory behind the construction in state-space terms of the solution of admissible degrees.

A departure of Algorithm 2.1 from the Löwner matrix–based approach is the determination of the minimal order. Under the stated conditions, in Algorithm 2.1 the minimal order and the observability range space are extracted by a singular value decomposition, while in the Löwner matrix–based approach the minimal order is determined by checking ranks of several (generalized) Löwner matrices. The singular value decomposition is not sensitive to random inaccuracies in data; that is, the true singular values and the observability range space are consistently estimated as N increases unboundedly, provided that n is finite or increases more slowly than N [24, 25]. To our best knowledge, an asymptotic error analysis for randomly corrupted transfer function evaluations has not been performed for any of the interpolation algorithms in the literature.

Deficiencies of the proposed interpolation algorithm and the Löwner matrix–based approach are the same. As pointed out in [1], a parameterization of solutions when the original data have to be added and derivation of recursive formulae for allowing update of a realization when one or more interpolation data become available are absent. It would be interesting to develop connections between the constrained interpolation problems such as the Nevanlinna–Pick and the positive-real interpolation and Algorithm 2.1. It is worth mentioning that the Nevanlinna–Pick interpolation can be transformed into an interpolation problem without norm constraint by adding the mirror image interpolation points to the original data [3].

3. Subspace-based identification with interpolation constraints. In this section, we will consider identification of an nth-order stable system with transfer function $G(z)$ from noisy samples of the frequency response,

$$w_l = G(e^{i\theta_l}) + \eta_l, \qquad l = 1, \dots, M, \tag{3.1}$$

with the interpolation constraints

$$\left.\frac{d^j}{dz^j} G(z)\right|_{z=z_k} = E_{kj}, \qquad j = 0, 1, \dots, N_k, \;\; k = 1, \dots, L, \tag{3.2}$$

where $0 \leq \theta_l \leq \pi$, $l = 1, \dots, M$, denote the discrete-time frequencies and η_l is a sequence of independent zero-mean complex random variables with a known covariance function that is uniformly bounded. The number of the constraints defined in (2.39) satisfies $N < n$. The interpolation constraints (3.2) 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 $\{w_l, \theta_l\}_{l=1}^{M}$ to an nth-order model $\widehat{G}_M(z)$ that satisfies the interpolation constraints in (3.2) such that, with probability one,

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

where

$$\|X\|_\infty \triangleq \sup_\omega \sigma_1(X(e^{i\omega}))$$

and σ_1 denotes the largest singular value. Algorithms with this property are called *strongly consistent*. This identification setup except for the constraints in (3.2) can be found, for example, in [24].

A motivating example for the constraints in (3.2) is as follows. Suppose that the system to be identified is nth order stable single-input/single-output continuous-time system represented by the transfer function

$$G^c(s) = \frac{b_0 s^m + b_1 s + \cdots + b_m}{s^n + a_1 s + \cdots + a_n}, \tag{3.3}$$

where the denominator degree n is greater than the numerator degree m, and we are given M noise corrupted frequency response measurements

$$w_l = G(i\omega_l) + \eta_l, \qquad l = 1, \dots, M. \tag{3.4}$$

Assuming $b_0 \neq 0$, the *relative degree* of $G^c(s)$ is defined as $\tau \triangleq n - m$.

A direct use of the Möbius transform technique (1.5) targets identifying the discrete-time equivalent of $G^c(s)$ defined by

$$G^d(z) \triangleq G^c\left(\psi(z)\right), \tag{3.5}$$

using $w_l,\ l = 1, \dots, M$, at the transformed discrete-time frequencies

$$\theta_k = 2\arctan\left(\frac{\omega_k}{\lambda}\right), \qquad k = 1, \dots, M. \tag{3.6}$$

Then, the continuous-time identified transfer function denoted by $\widehat{G}^c_M(s)$ is obtained from the discrete-time identified transfer function denoted by $\widehat{G}^d_M(z)$ by using the inverse Möbius map $z = \psi^{-1}(s)$; i.e., $\widehat{G}^c_M(s) = \widehat{G}^d_M(\psi^{-1}(s))$. Due to noise and unmodeled dynamics, the former is only a proper transfer function.

If maintaining the relative degree is a concern, we then high-pass filter $\widehat{G}^c_M(s)$ as follows:

$$\widehat{G}_M(s) = \frac{\widehat{G}^c_M(s)}{(s+\mu)^\tau},$$

where $\mu > 0$ is chosen sufficiently outside the bandwidth of $\widehat{G}^c_M(s)$. This filtering increases the order of the identified model by τ. This problem can be circumvented by including the constraints

$$\left.\frac{d^j}{dz^j} G^d(z)\right|_{z=-1} = 0, \qquad j = 1, \dots, \tau,$$

in the problem formulation. Observe that when applied to (3.3), the Möbius map (1.5) introduces a zero of $G^d(z)$ at $z = -1$ with multiplicity τ.

Now, the solution of the constrained identification problem (3.1)–(3.2) is particularly simple if one notes from (2.54) the following set of equations:

$$\delta_{0j}\, D + (-1)^j j!\, C(z_k I_n - A)^{-j-1} B = E_{kj}, \quad j = 0, \dots, N_k, \quad k = 1, \dots, L, \tag{3.7}$$

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 (2.57) with the linear constraints (3.7). With this modification, the frequency domain subspace-based identification algorithm presented in [24] is strongly consistent. The inclusion of the noise covariance information in the algorithm is straightforward and can be found in [24]. This extension can be viewed as the tangential version of the Lagrange–Sylvester interpolation problem (1.1).

4. Example. The purpose of this section is to illustrate Algorithm 2.1 with a step-by-step numerical example. Suppose that the system to be found by interpolation has the following state-space representation:

$$A = \begin{bmatrix} -0.5 & 0.5 & 0 & 0 \\ -0.5 & -0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & -0.25 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$
$$C = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Thus, $n = 4$, $p = 2$, and $m = 3$. This system has the transfer function

$$G(z) = \begin{bmatrix} \dfrac{z^2+3z+1.5}{z^2+z+0.5} & -\dfrac{z^3+0.5z^2+0.5z+0.75}{z^3+0.5z^2-0.25} & 0 \\ \dfrac{2z^2+1.25z+0.5}{z^3+1.25z^2+0.75z+0.125} & \dfrac{z^3+3.25z^2+2.5z+0.75}{z^3+1.25z^2+0.75z+0.125} & \dfrac{z+1.25}{z+0.25} \end{bmatrix}.$$

Let us assume that the interpolation data are as follows:

$$z_1 = 1+i,\ N_1 = 0,\ z_2 = 1-i,\ N_2 = 0,\ z_3 = 2,\ N_3 = 4$$

and

$$w_{10} = \begin{bmatrix} 1.9333 - 0.5333i & -0.8667 + 0.4000i & 0 \\ 0.8878 - 0.5236i & 1.9545 - 0.6569i & 1.4878 - 0.3902i \end{bmatrix},$$

$$w_{20} = \begin{bmatrix} 1.9333 + 0.5333i & -0.8667 - 0.4000i & 0 \\ 0.8878 + 0.5236i & 1.9545 + 0.6569i & 1.4878 + 0.3902i \end{bmatrix},$$

$$w_{30} = \begin{bmatrix} 1.7692 & -1.2051 & 0 \\ 0.7521 & 1.8291 & 1.4444 \end{bmatrix}, \quad w_{31} = \begin{bmatrix} -0.2840 & 0.2433 & 0 \\ -0.2804 & -0.3395 & -0.1975 \end{bmatrix},$$

$$w_{32} = \begin{bmatrix} 0.2003 & -0.4251 & 0 \\ 0.2084 & 0.2757 & 0.1756 \end{bmatrix}, \quad w_{33} = \begin{bmatrix} -0.2000 & 0.9844 & 0 \\ -0.2333 & -0.3341 & -0.2341 \end{bmatrix},$$

$$w_{34} = \begin{bmatrix} 0.2456 & -2.8518 & 0 \\ 0.3531 & 0.5390 & 0.4162 \end{bmatrix}.$$

Then we set $q = 5$ and compute $N = 9$. Therefore, the inequalities $N \geq q + n$ and $q > n$ are both satisfied. In step 1, we compute the matrices $\widehat{\mathcal{H}} \in \mathbf{R}^{10\times 27}$ and $\mathcal{F} \in \mathbf{R}^{15\times 27}$. The QR-factorization in step 2 results in $R_{22} \in \mathbf{R}^{10\times 12}$ given by

$$R_{22} = \begin{bmatrix} -0.4622 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0.0381 & -0.0518 & 0 & 0 & \vdots & \ddots & \vdots \\ -0.2544 & 0.0203 & -0.0240 & 0 & & & \\ -0.0176 & 0.0194 & -0.0075 & -0.0009 & & & \\ -0.1144 & -0.0070 & 0.0091 & -0.0089 & & & \\ 0.0094 & -0.0110 & 0.0094 & -0.0025 & & & \\ -0.0583 & 0.0035 & -0.0045 & 0.0045 & & & \\ -0.0033 & 0.0057 & -0.0061 & 0.0037 & & & \\ -0.0344 & 0.0033 & -0.0037 & -0.0022 & \vdots & \ddots & \vdots \\ -0.0007 & -0.0013 & 0.0015 & -0.0027 & 0 & \cdots & 0 \end{bmatrix},$$

which is not unexpected since $n = 4$. In step 3, we compute the nonzero singular values 0.5460, 0.0609, 0.0249, and 0.0098. The matrices $\widehat{A}$ and $\widehat{C}$ computed in step 5 are

$$\widehat{A} = \begin{bmatrix} 0.5204 & -0.1361 & 0.3199 & 0.5352 \\ 0.0882 & -0.4983 & 0.4848 & -0.1035 \\ 0.0052 & 0.0820 & -0.4810 & 0.7195 \\ -0.0295 & 0.1919 & -0.3546 & -0.2911 \end{bmatrix},$$

$$\widehat{C} = \begin{bmatrix} 0.8460 & 0.2123 & -0.2149 & -0.3233 \\ -0.0721 & 0.8069 & 0.5289 & 0.1046 \end{bmatrix}.$$

In step 6, we compute $\widehat{\mathcal{G}} \in \mathbf{R}^{18\times 3}$ and $\widehat{\mathcal{Y}} \in \mathbf{R}^{18\times 6}$ matrices, and the solution of the least-squares problem is

$$\widehat{B} = \begin{bmatrix} 1.0502 & -0.5390 & -0.0816 \\ 2.8626 & 1.8321 & 0.9041 \\ -0.1545 & 1.0984 & 0.4896 \\ -1.4555 & -0.9375 & 0.0547 \end{bmatrix},$$

$$\widehat{D} = \begin{bmatrix} 1.0000 & -1.0000 & -0.0000 \\ -0.0000 & 1.0000 & 1.0000 \end{bmatrix}.$$

The realization $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ is similar to (A, B, C, D). In fact, the estimates of the interpolation data computed from the former has a maximum error 5.9746×10^{-14}.

5. 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 frequency domain subspace-based identification methods and is not sensitive to inaccuracies in data. A necessary and sufficient condition for the existence and the uniqueness of a minimal interpolant was formulated in terms of the total multiplicity of the interpolation nodes. The purpose of this contribution was to pinpoint the kinship between the frequency domain subspace-based identification of stable linear systems and the minimal rational interpolation of stable systems.

Acknowledgment. The authors would like to thank the associate editor and the reviewers of the manuscript for helpful remarks and references.

REFERENCES

[1] B. D. O. Anderson and A. C Antoulas, *Rational interpolation and state-variable realizations*, Linear Algebra Appl., 137/138 (1990), pp. 479–509.

[2] A. C. Antoulas and B. D. O. Anderson, *On the scalar rational interpolation problem*, IMA J. Math. Control Inform., 3 (1986), pp. 61–88.

[3] A. C. Antoulas and B. D. O. Anderson, *On the problem of stable rational interpolation*, Linear Algebra Appl., 122/123/124 (1989), pp. 301–329.

[4] A. C Antoulas, J. A. Ball, J. Kang, and J. C. Willems, *On the solution of the minimal rational interpolation problem*, Linear Algebra Appl., 137/138 (1990), pp. 511–573.

[5] J. A. Ball and J. Kang, *Matrix polynomial solution of tangential Lagrange-Sylvester interpolation conditions of low McMillan degree*, Linear Algebra Appl., 137/138 (1990), pp. 699–746.

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

[7] J. A. Ball, I. Gohberg, and L. Rodman, *Two-sided Lagrange-Sylvester interpolation problem for rational matrix functions*, in Proc. Sympos. Pure Math. 51, AMS, Providence, RI, 1990, pp. 17–83.

[8] J. A. Ball, I. Gohberg, and L. Rodman, *Simultaneous residue interpolation for rational matrix functions*, Integral Equations Operator Theory, 13 (1990), pp. 611–637.

[9] J. A. Ball, M. A. Kaashoek, G. Groenewald, and J. Kim, *Column reduced rational matrix functions with given null-pole data in the complex plane*, Linear Algebra Appl., 203/204 (1994), pp. 111–138.

[10] A. Blomqvist, A. Lindquist, and R. Nagamune, *Matrix-valued Nevanlinna-Pick interpolation with complexity constraints: An optimization approach*, IEEE Trans. Automat. Control, 48 (2003), pp. 2172–2190.

[11] T. Boros, A. H. Sayed, and T. Kailath, *A recursive method for solving unconstrained tangential interpolation problems*, IEEE Trans. Automat. Control, 44 (1999), pp. 454–470.

[12] C. I. Byrnes, T. T. Georgiou, and A. Lindquist, *A new approach to spectral estimation: A tunable high-resolution spectral estimator*, IEEE Trans. Signal Process., 48 (2000), pp. 3189–3205.

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

[14] I. P. Fedcina, *A description of the solution of the Nevanlinna-Pick tangent problem*, Akad. Nauk Armyan. SSR. Dokl., 60 (1975), pp. 37–42 (in Russian).

[15] I. P. Fedcina, *The Nevanlinna-Pick tangent problem with multiple points*, Akad. Nauk Armyan. SSR. Dokl., 61 (1975), pp. 214–218 (in Russian).

[16] T. Kailath, *Linear Systems*, Prentice–Hall, Englewood Cliffs, NJ, 1980.

[17] R. E. Kalman, *On partial realization, transfer-functions and canonical forms*, Acta. Polytech. Scand. Math., 31 (1979), pp. 9–32.

[18] J. Kamali, T. Boros, T. Kailath, and G. Franklin, *Q-parametrization for unstable plants: A displacement structure approach*, in Proceedings of the American Control Conference, 1995, pp. 4401–4402.

[19] H. Kimura, *Directional interpolation approach to H^∞-optimization and robust stabilization*, IEEE Trans. Automat. Control, AC-32 (1987), pp. 1085–1093.

[20] H. Langer and A. Lasarow, *Solution of a multiple Nevanlinna-Pick problem via orthogonal rational functions*, J. Math. Anal. Appl., 293 (2004), pp. 605–632.

[21] D. J. N. Limebeer and B. D. O. Anderson, *An interpolation theory approach to H^∞-controller degree bounds*, Linear Algebra Appl., 98 (1988), pp. 347–386.

[22] L. Ljung, *Linear system identification as curve fitting*, in Directions in Mathematical Systems Theory and Optimization, A. Rantzer and C. I. Byrnes, eds., LNCIS 286, Springer-Verlag, Berlin, Heidelberg, 2003, pp. 203–215.

[23] K. Liu, R. N. Jacgues, and D. W. Miller, *Frequency domain structural system identification by observability range space extraction*, in Proceedings of the American Control Conference, Baltimore, MD, 1994, pp. 107–111.

[24] T. McKelvey, H. Akçay, and L. Ljung, *Subspace-based multivariable system identification from frequency response data*, IEEE Trans. Automat. Control, 41 (1996), pp. 960–979.

[25] T. McKelvey, H. Akçay, and L. Ljung, *Subspace-based identification of infinite-dimensional multivariable systems from from frequency response data*, Automatica, 32 (1996), pp. 885–902.

[26] C. Mosquera and F. Pérez, *Algebraic solution to the robust SPR problem for two polynomials*, Automatica, 37 (2001), pp. 757–762.

[27] Nikolai K. Nikolski, *Operators, Functions, and Systems: An Easy Reading, Vol.* 1*: Hardy, Hankel, and Toeplitz*, AMS, Providence, RI, 2002.

[28] Nikolai K. Nikolski, *Operators, Functions, and Systems: An Easy Reading, Vol.* 2*: Model Operators and Systems*, AMS, Providence, RI, 2002.

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

[30] A. H. Sayed, T. Kailath, H. Lev-Ari, and T. Constantinescu, *Recursive solutions of rational interpolation problems via fast matrix factorization*, Integral Equations Operator Theory, 20 (1994), pp. 84–118.

[31] A. H. Sayed, T. Constantinescu, and T. Kailath, *Time-variant displacement structure and interpolation problems*, IEEE Trans. Automat. Control, 39 (1994), pp. 960–976.

[32] R. S. Smith and J. C. Doyle, *Model validation: A connection between robust control and identification*, IEEE Trans. Automat. Control, 37 (1992), pp. 942–952.

[33] P. Van Overschee and B. De Moor, *Continuous-time frequency domain subspace identification*, Signal Processing, 52 (1996), pp. 179–194.

[34] D. C. Youla and M. Saito, *Interpolation with positive-real functions*, J. Franklin Inst., 284 (1967), pp. 77–108.