INRIA

Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm.

Bernard Hanzon — Martine Olivi — Ralf L.M. Peeters

N° ????

Février 2004

THÈME 4

Rapport de recherche

Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm.

Bernard Hanzon* , Martine Olivi , Ralf L.M. Peeters [†]

Abstract: In this report, the connections are investigated between two different approaches towards the parametrization of multivariable stable all-pass systems in discrete-time. The first approach involves the tangential Schur algorithm, which employs linear fractional transformations. It stems from the theory of reproducing kernel Hilbert spaces and enables the direct construction of overlapping local parametrizations using Schur parameters and interpolation points. The second approach proceeds in terms of state-space realizations. In the scalar case, a balanced canonical form exists that can also be parametrized by Schur parameters. This canonical form can be constructed recursively, using unitary matrix operations. Here, this procedure is generalized to the multivariable case by establishing the connections with the first approach. It gives rise to balanced realizations and overlapping canonical forms directly in terms of the parameters used in the tangential Schur algorithm.

* Centre for Mathematics and Computer Science, Dept.MAS, P.O.Box 94079, 1090 GB Amsterdam, The Netherlands, Tel: +31-20-5924176, Email: Bernard.Hanzon@cwi.nl
† Dept. Mathematics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, Tel.: +31-43-3883365, Fax: +31-43-3884910, Email: ralf.peeters@math.unimaas.nl

Réalisations équilibrées des systèmes passe-tout stables en temps discret et l'algorithme de Schur tangentiel.

Résumé : Dans ce rapport, deux approches différentes du problème de paramétrage des systèmes multivariables passe-tout stables en temps discret sont reliées entre elles. La première approche basée sur une version matricielle de l'algorithme de Schur (dit algorithme de Schur tangentiel) utilise des applications linéaires fractionnelles et permet la construction d'un atlas de cartes. La deuxième approche s'appuie sur la théorie des réalisations. Dans le cas scalaire, une forme canonique existe, qui peut être paramétrée par les paramètres de Schur. Cette forme canonique peut-être obtenue récursivement à l'aide de multiplications par des matrices unitaires. Ici, en établissant les liens avec la première approche, cette procédure a pu être généralisée au cas multivariable. Cela permet d'obtenir des réalisations équilibrées et un atlas de formes canoniques directement en termes de paramètres de Schur.

1   Introduction

Stable all-pass systems of finite order have several applications in linear systems theory. Within the fields of system identification, approximation and model reduction, they have been used in connection with the Douglas-Shapiro-Shields factorization, see e.g., [7, 4, 17, 10], to obtain effective algorithms for various purposes. They are one-to-one related to rational inner functions, of which the differential structure has been studied in [1]. There, a parametrization for the multivariable case has been obtained by means of a recursive procedure, the *tangential Schur algorithm*, which involves Schur parameter vectors, interpolation points and normalized direction vectors. In the scalar case, a single coordinate chart suffices to entirely describe the manifold of stable all-pass (or lossless) systems of a fixed finite order. In the multivariable case, the approach leads to atlases of generic charts covering these manifolds.

In another line of research, balanced state-space canonical forms have been constructed for various classes of linear systems, with special properties of these classes (such as stability) built in, see e.g., [19, 15]. Balanced realizations are well-known to have numerical advantages and are useful for model reduction purposes in conjunction with balance-and-truncate type procedures. In the constructions of [15], the case of stable all-pass systems in *continuous-time* plays a central role. In the scalar case, the resulting canonical form is balanced with a positive upper triangular reachability matrix. In the multivariable case, Kronecker indices and nice selections are used to arrive at block-balanced overlapping canonical forms. For *discrete-time* stable all-pass systems, canonical forms can be obtained from the results in continuous-time by application of a bilinear transformation. However, this destroys certain nice properties of the canonical form; e.g., truncation of state components no longer leads to reduced order systems that are balanced and in canonical form. Therefore, the ideas of [15] are applied in [14] to the scalar discrete-time stable all-pass case directly. This leads to a balanced canonical form with the desired properties, for which it turns out that it can in fact be parametrized using Schur parameters.

In this paper, the connections between these two approaches are investigated and the results of [14] are generalized to the multivariable case. Moreover, the atlases developed in [1] are supplied with balanced state-space realizations in terms of the parameters used in the tangential Schur algorithm. In Section 3, we introduce a mapping which acts on proper rational matrix functions and which has a particularly simple expression on the level of realizations, see formula (34). When specialized to the case of lossless functions, it allows for a recursive construction of balanced realizations that can be implemented as a product of unitary matrices. This construction has several advantages from a numerical point of view. In Section 4, we recall some basic results on linear fractional transformations of lossless functions, which are at the heart of the tangential Schur algorithm, described in Section 5. The results of these sections are mostly well-known, except for the nice factorization of J-inner functions in Proposition 5.2, which we could not find in the literature and which enables an elegant presentation of these results. The main technical results are presented in Section 6 and they describe the complete connection between the construction of Section 3 and the tangential Schur algorithm in its more general form. In Section 7 the results are used to construct atlases for the manifold of lossless systems of a fixed McMillan degree and for the manifold of stable linear systems of a fixed McMillan degree.

2   Preliminaries

In this section a brief overview of results from the literature is presented to provide the necessary background for the main constructions of this paper. They concern realization theory, balancing and the basic theory of lossless and J-inner functions.

Realization theory and balancing.

Consider a linear time-invariant state-space system in discrete time with m inputs and p outputs:

$$x_{t+1} = Ax_t + Bu_t, \tag{1}$$
$$y_t = Cx_t + Du_t, \tag{2}$$

with $t \in \mathbb{Z}$, $x_t \in \mathbb{C}^n$ for some nonnegative integer n, $u_t \in \mathbb{C}^m$, $y_t \in \mathbb{C}^p$. Furthermore, the matrices A, B, C and D with complex-valued entries are of compatible sizes: $n \times n$, $n \times m$, $p \times n$ and $p \times m$, respectively. The corresponding transfer matrix of this system is given by $G(z) = D + C(zI_n - A)^{-1}B$, which is a $p \times m$ matrix with rational functions as its entries. Two state-space systems are called input-output equivalent if they have the same transfer matrix. Here, two rational (matrix) functions are identified if they agree almost everywhere, i.e., common factors are always canceled. Note that a transfer matrix in this set-up is always proper (i.e., analytic at infinity), D being the value at infinity.

>From realization theory it follows that, conversely, any $p \times m$ rational matrix function $G(z)$ analytic at infinity can be written in the form

$$G(z) = D + C(zI_n - A)^{-1}B \tag{3}$$

where (A, B, C, D) is an appropriate quadruple of matrices and n a suitable state space dimension. Such a quadruple with the associated expression (3) is called a *state-space realization* of $G(z)$. To such a realization we associate the block-partitioned matrix

$$R = \left[\begin{array}{cc} D & C \\ B & A \end{array} \right] \tag{4}$$

which we call the *realization matrix*. It will play an important role in the sequel.

If all possible realizations of $G(z)$ have state-space dimension at least as big as the state-space dimension n of (A, B, C, D) then the realization is said to be minimal, and n is called the order or the McMillan degree of the transfer function. Two minimal realizations (A, B, C, D) and (A', B', C', D') of a given function $G(z)$ are always similar: there exists a unique invertible matrix T such that

$$\left[\begin{array}{cc} D' & C' \\ B' & A' \end{array} \right] = \left[\begin{array}{cc} I_p & 0 \\ 0 & T \end{array} \right] \left[\begin{array}{cc} D & C \\ B & A \end{array} \right] \left[\begin{array}{cc} I_m & 0 \\ 0 & T^{-1} \end{array} \right]. \tag{5}$$

As is well-known, an output pair (C, A) is observable if the observability matrix

$$\left[\begin{array}{c} C \\ CA \\ \vdots \\ CA^{n-1} \end{array} \right] \tag{6}$$

has full column rank. An input pair (A, B) is reachable if the associated reachability matrix

$$\begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix} \tag{7}$$

has full row rank. It is well-known that a realization (A, B, C, D) is minimal if and only if (C, A) is observable and (A, B) is reachable; cf., e.g., [16]. In this case, the poles of the function $G(z)$ are the eigenvalues of A, and its McMillan degree is equal to the sum of the degrees (see [16, Sect. 6.5]) of all the poles of $G(z)$. This provides an alternative definition for the McMillan degree, which generalizes to the case of non-proper rational functions by including the pole at infinity.

Let (A, B, C, D) be some realization of a transfer function. If the eigenvalues of A all belong to the open unit disk, then the matrix A is called (discrete-time) asymptotically stable, and (A, B, C, D) an asymptotically stable realization. If (A, B, C, D) is an asymptotically stable realization, then the controllability Gramian W_c and the observability Gramian W_o are well defined as the exponentially convergent series

$$W_c = \sum_{k=0}^{\infty} A^k BB^*(A^*)^k, \quad W_o = \sum_{k=0}^{\infty} (A^*)^k C^* C A^k, \tag{8}$$

where the notation * is used to denote Hermitian transposition of a matrix (i.e., the joint action of matrix transposition and complex conjugation of the matrix entries). The Gramians are characterized as the unique (and positive semi-definite) solutions of the respective Lyapunov-Stein equations

$$W_c - AW_cA^* = BB^*, \tag{9}$$
$$W_o - A^*W_oA = C^*C. \tag{10}$$

Moreover, under asymptotic stability of A it holds that W_c is positive definite if and only if the pair (A, B) is reachable, and W_o is positive definite if and only if the pair (C, A) is observable. A minimal and asymptotically stable realization (A, B, C, D) of a transfer function is called *balanced* if its observability and controllability Gramians are both diagonal and equal. Any minimal and asymptotically stable realization (A, B, C, D) is similar to a balanced realization. The concept of balanced realizations was first introduced in [18] in the continuous time case and used for model reduction. In [22] the same was done for the discrete time case. Balanced realizations are now a well-established tool which often exhibit good numerical properties.

J-inner, J-unitary and lossless functions.

For any matrix function $R(z)$, we define the matrix functions $R^*(z)$ and $R^\sharp(z)$ by

$$R^*(z) := R(\overline{z})^*, \quad \text{and} \quad R^\sharp(z) := R^*(z^{-1}). \tag{11}$$

Note that if z lies on the unit circle, then $R^\sharp(z) = R(z)^*$. For (square) Hermitian matrices P and Q, either of the notations $P \leq Q$ and $Q \geq P$ will be used to express that $Q - P$ is a positive semi-definite matrix. As is well-known this introduces a partial ordering on the set of square Hermitian matrices.

Now let Σ be a constant $k \times k$ matrix which is both Hermitian and unitary. Note that Σ is

unitarily similar to a signature matrix, i.e., there exists a unitary $k \times k$ matrix U for which $U\Sigma U^*$ attains the form $\begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$ for some non-negative integers p and q with $p + q = k$.

A square rational matrix function $\Theta(z)$ of size $k \times k$ is called Σ-inner (in the unit disk), if at every point of analyticity z of $\Theta(z)$ it satisfies

$$\Theta(z)^*\Sigma\Theta(z) \;\leq\; \Sigma, \quad |z| < 1, \tag{12}$$

$$\Theta(z)^*\Sigma\Theta(z) \;=\; \Sigma, \quad |z| = 1, \tag{13}$$

$$\Theta(z)^*\Sigma\Theta(z) \;\geq\; \Sigma, \quad |z| > 1. \tag{14}$$

It can be shown that this definition contains redundancy in the sense that either one of these three defining properties (12)-(14) is implied by the other two; cf., e.g., [20, 8, 12].

Any rational matrix function $\Theta(z)$ which satisfies the property (13) is called Σ-unitary. For such functions, replacing $\overline{z}$ by z^{-1} on the unit circle, the identity (13) extends almost everywhere by analytic continuation, so that a rational Σ-unitary function is invertible and its inverse is given by

$$\Theta(z)^{-1} = \Sigma\Theta^\sharp(z)\Sigma. \tag{15}$$

This applies *a fortiori* to Σ-inner functions. A function is called Σ-lossless if it is $(-\Sigma)$-inner. If $\Theta(z)$ is Σ-inner, then $\Theta^*(z)$ is also Σ-inner whereas both $\Theta^\sharp(z)$ and $\Theta(z)^{-1}$ are Σ-lossless. The class of Σ-inner functions is closed under multiplication.

As usual, an I_k-lossless function is just called *(discrete-time) lossless*, or *(discrete-time) stable all-pass* and an I_k-inner function is simply called *inner*. Throughout this paper we will be much concerned with lossless and inner functions of size $p \times p$, and with J-unitary and J-inner functions of size $2p \times 2p$, where J denotes the following signature matrix:

$$J = \begin{bmatrix} I_p & 0 \\ 0 & -I_p \end{bmatrix}. \tag{16}$$

Note that J-inner and J-lossless functions in general may have poles everywhere in the complex plane, but inner functions are analytic inside the unit disk and lossless functions are analytic outside the unit disk (including at the point at infinity). Therefore a rational lossless function $G(z)$ is proper, having an inverse which is inner and given by

$$G(z)^{-1} = G^\sharp(z). \tag{17}$$

A scalar rational function is a *Blaschke product* if it maps the unit circle into the unit circle. It can be shown that such a function can be written as the product of a finite number of *Blaschke factors*, which are rational functions of first order that map the unit circle into the unit circle.

Lemma 2.1 *Let $G(z)$ be a $p \times p$ lossless function of McMillan degree n. Then its determinant $\det G(z)$ is a Blaschke product. If one writes in the form $\det G(z) = p(z)/q(z)$ with p and q co-prime polynomials, then the degree of $q(z)$ is equal to n.*

Proof. See e.g. [5]. □

In the next sections we will go into some more detail about the construction of balanced state-space realizations of lossless functions.

We finally describe some particular J-inner functions that will be intensively used in the sequel, namely the constant J-unitary matrices and the J-inner functions of McMillan degree one. Details and proofs can be found in [8].

Every *constant* J-unitary matrix can be represented in a unique way (see [8, Thm. 1.2]) as follows:

$$M = H(E) \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}, \tag{18}$$

where P and Q are $p \times p$ unitary matrices and $H(E)$ denotes the Halmos extension of a strictly contractive $p \times p$ matrix E (i.e., such that $I - E^*E > 0$). This Halmos extension $H(E)$ is defined by

$$H(E) = \begin{bmatrix} (I - EE^*)^{-1/2} & E(I - E^*E)^{-1/2} \\ E^*(I - EE^*)^{-1/2} & (I - E^*E)^{-1/2} \end{bmatrix} = \begin{bmatrix} (I - EE^*)^{-1/2} & (I - EE^*)^{-1/2}E \\ (I - E^*E)^{-1/2}E^* & (I - E^*E)^{-1/2} \end{bmatrix}. \tag{19}$$

It holds that $H(E)$ is Hermitian, J-unitary and invertible with inverse $H(E)^{-1} = H(-E)$. Also, $KH(E)K = H(E^*)$ for $K = \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix}$.

Two other important forms of J-unitary matrices are described in the following lemma (see [8, Thm.1.3]).

Lemma 2.2 *If x is a $2p$-vector such that $x^*Jx \neq 0$, then the set of $2p \times 2p$ matrices of the form*

$$X_x(\alpha) = I_{2p} + (\alpha - 1)x(x^*Jx)^{-1}x^*J, \quad \alpha \in \mathbb{C}, \tag{20}$$

is closed under multiplication. We have that

$$X_x(\alpha)X_x(\beta) = X_x(\alpha\beta), \quad \det X_x(\alpha) = \alpha, \quad X_x(\alpha)^{-1} = X_x(\alpha^{-1}), \quad \text{for } \alpha \neq 0, \tag{21}$$

$$J - X_x(\alpha)JX_x(\beta)^* = (1 - \alpha\beta^*)x(x^*Jx)^{-1}x^*J, \tag{22}$$

and $X_x(\alpha)$ is J-unitary if and only if $|\alpha| = 1$.

If x is a $2p$-vector such that $x^*Jx = 0$, then the set of matrices of the form

$$Y_x(\alpha) = I_{2p} + \alpha xx^*J, \quad \alpha \in \mathbb{C}, \tag{23}$$

is closed under multiplication. We have that

$$Y_x(\alpha)Y_x(\beta) = Y_x(\alpha + \beta), \quad \det Y_x(\alpha) = 1, \quad Y_x(\alpha)^{-1} = Y_x(-\alpha), \tag{24}$$

$$J - Y_x(\alpha)JY_x(\beta)^* = (\alpha + \beta^*)xx^*, \tag{25}$$

and $Y_x(\alpha)$ is J-unitary if and only if α is pure imaginary.

We now come to the description of J-inner matrix functions of McMillan degree 1. We shall call these functions 'elementary J-inner functions' or 'elementary J-inner factors', since it has been shown by Potapov (see [20]) that every J-inner matrix function of McMillan degree n can be decomposed into a product of n such elementary J-inner factors. Elementary J-inner functions are obtained by applying the functions X_x and Y_x to appropriate scalar functions of degree one, namely the Blaschke factor $b_w(z)$ defined by

$$b_w(z) := \frac{z-w}{1-\overline{w}z}, \quad w \notin \mathbb{T} \tag{26}$$

and the Carathéodory function,

$$c_w(z) := \frac{z+w}{z-w}, \quad w \in \mathbb{T}, \tag{27}$$

where $\mathbb{T}$ denotes the unit circle.

Theorem 2.3 *Let $\Phi(z)$ be an elementary J-inner factor with a pole at $z = 1/\overline{w}$. Then, apart from a constant J-unitary multiplier on the right, $\Phi(z)$ must be in one of the following three forms:*

(1) If $w \in \mathbb{C}$, $w \notin \mathbb{T}$,
$$\Phi(z) = X_x(b_w(z)) = I_{2p} + (b_w(z) - 1)\frac{xx^*J}{x^*Jx}, \tag{28}$$
(see Eqn. (20)) for some $x \in \mathbb{C}^{2p}$ such that :

$$\frac{x^*Jx}{1-|w|^2} > 0. \tag{29}$$

(2) If $w \in \mathbb{T}$,
$$\Phi(z) = Y_x(-\delta \ c_w(z)) = I_{2p} - \delta \ c_w(z) \ xx^*J, \tag{30}$$
(see Eqn. (23)) for some nonzero $x \in \mathbb{C}^{2p}$ such that $x^*Jx = 0$, and for some real $\delta > 0$.

(3) If $w = \infty$, then
$$\Phi(z) = X_x\left(\frac{1}{z}\right) = I_{2p} + \left(\frac{1}{z} - 1\right)\frac{xx^*J}{x^*Jx}, \tag{31}$$
for some $x \in \mathbb{C}^{2p}$ such that $x^*Jx < 0$.

In the literature, an alternative description of elementary J-inner factors is in terms of reproducing kernel Hilbert spaces associated with J-inner functions. It holds that elementary J-inner functions are precisely those functions for which the reproducing kernel has dimension one. The theory of reproducing kernel Hilbert spaces offers a very nice unifying framework to study various closely related questions from interpolation theory. However, we shall not go into details here but refer the reader to [8] and [3] instead.

3   Recursive construction of balanced realizations of lossless systems

With each pair (U, V) of $(p + 1) \times (p + 1)$ matrices we associate a mapping $\mathcal{F}_{U,V}$ which is defined to act on proper rational $p \times p$ matrix functions $G(z)$ as follows.

$$\mathcal{F}_{U,V} : G(z) \mapsto F_1(z) + \frac{F_2(z)F_3(z)}{z - F_4(z)}, \tag{32}$$

with $F_1(z)$ of size $p \times p$, $F_2(z)$ of size $p \times 1$, $F_3(z)$ of size $1 \times p$ and $F_4(z)$ scalar, defined by:

$$F(z) = \left[\begin{array}{cc} F_1(z) & F_2(z) \\ F_3(z) & F_4(z) \end{array} \right] = V \left[\begin{array}{cc} 1 & 0 \\ 0 & G(z) \end{array} \right] U^*. \tag{33}$$

A state-space realization of $\widetilde{G}(z) = \mathcal{F}_{U,V}(G(z))$ can be obtained by working directly on any realization matrix R of $G(z)$. This is the content of the following proposition.

Proposition 3.1 *Let $G(z)$ be a $p \times p$ proper rational transfer function and (U, V) a pair of $(p + 1) \times (p + 1)$ matrices, then $\widetilde{G}(z) = \mathcal{F}_{U,V}(G(z))$ is well-defined. Let (A, B, C, D) be a state-space realization of $G(z)$ with n-dimensional state-space. Then a state-space realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ of $\widetilde{G}(z)$ with $(n + 1)$-dimensional state-space is given by:*

$$\left[\begin{array}{cc} \widetilde{D} & \widetilde{C} \\ \widetilde{B} & \widetilde{A} \end{array} \right] = \left[\begin{array}{cc} V & 0 \\ 0 & I_n \end{array} \right] \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & D & C \\ 0 & B & A \end{array} \right] \left[\begin{array}{cc} U^* & 0 \\ 0 & I_n \end{array} \right]. \tag{34}$$

Proof. Since $G(z)$ is proper, $F_4(z)$ is proper too and therefore $z - F_4(z)$ does not vanish identically so that $\mathcal{F}_{U,V}(G(z))$ is well-defined.

Now observe that the right-hand side of Eqn. (34) provides a realization matrix with n-dimensional state-space for the transfer function $F(z)$ given by Eqn. (33):

$$\left[\begin{array}{cc} D_F & C_F \\ B_F & A_F \end{array} \right] = \left[\begin{array}{cc} V & 0 \\ 0 & I_n \end{array} \right] \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & D & C \\ 0 & B & A \end{array} \right] \left[\begin{array}{cc} U^* & 0 \\ 0 & I_n \end{array} \right].$$

Here D_F is $(p + 1) \times (p + 1)$, C_F is $(p + 1) \times n$, B_F is $n \times (p + 1)$ and $A_F = A$ is $n \times n$. We shall prove that changing the partitioning of this matrix to the one in (34)

$$\left[\begin{array}{cc} D_F & C_F \\ B_F & A_F \end{array} \right] = \left[\begin{array}{cc} \widetilde{D} & \widetilde{C} \\ \widetilde{B} & \widetilde{A} \end{array} \right],$$

where $\widetilde{D}$ is $p \times p$, $\widetilde{C}$ is $p \times (n + 1)$, $\widetilde{B}$ is $(n + 1) \times p$ and $\widetilde{A}$ is $(n + 1) \times (n + 1)$, gives a realization of $\widetilde{G}(z)$ with $n + 1$-dimensional state-space. To see this, let the matrices $\widetilde{C}$, $\widetilde{B}$ and $\widetilde{A}$ be partitioned

as follows:

$$\left[\begin{array}{cc} \widetilde{D} & \widetilde{C} \\ \widetilde{B} & \widetilde{A} \end{array}\right] = \left[\begin{array}{c|cc} \widetilde{D} & \widetilde{C}_1 & \widetilde{C}_2 \\ \hline \widetilde{B}_1 & \widetilde{A}_1 & \widetilde{A}_2 \\ \widetilde{B}_2 & \widetilde{A}_3 & A_F \end{array}\right] = \left[\begin{array}{cc|c} \widetilde{D} & \widetilde{C}_1 & \widetilde{C}_2 \\ \widetilde{B}_1 & \widetilde{A}_1 & \widetilde{A}_2 \\ \hline \widetilde{B}_2 & \widetilde{A}_3 & A_F \end{array}\right] = \left[\begin{array}{cc} D_F & C_F \\ B_F & A_F \end{array}\right],$$

where $\widetilde{C}_1$ is $p \times 1$, $\widetilde{C}_2$ is $p \times n$, $\widetilde{B}_1$ is $1 \times p$, $\widetilde{B}_2$ is $n \times p$, $\widetilde{A}_2$ is $1 \times n$, $\widetilde{A}_3$ is $n \times 1$ and $\widetilde{A}_1$ is scalar. Note that it follows that $F_4(z)$ has the realization $(A_F, \widetilde{A}_3, \widetilde{A}_2, \widetilde{A}_1)$ and therefore

$$z - \widetilde{A}_1 - \widetilde{A}_2(zI_n - A_F)^{-1}\widetilde{A}_3 = z - F_4(z).$$

Using this together with the well-known formula for the inverse of a partitioned matrix (see [8] or [16]) we compute

$$(zI_{n+1} - \widetilde{A})^{-1} = \left[\begin{array}{cc} (z - F_4(z))^{-1} & -(z - F_4(z))^{-1}\widetilde{A}_2(zI_n - A_F)^{-1} \\ -(zI_n - A_F)^{-1}\widetilde{A}_3(z - F_4(z))^{-1} & \Delta(z) \end{array}\right]$$

with

$$\Delta(z) = (zI_n - A_F)^{-1} + (zI_n - A_F)^{-1}\widetilde{A}_3(z - F_4(z))^{-1}\widetilde{A}_2(zI_n - A_F)^{-1}.$$

We then have

$$\begin{aligned} \widetilde{D} + \widetilde{C}(zI_{n+1} - \widetilde{A})^{-1}\widetilde{B} &= \widetilde{D} + \left[\begin{array}{cc} \widetilde{C}_1 & \widetilde{C}_2 \end{array}\right](zI_{n+1} - \widetilde{A})^{-1}\left[\begin{array}{c} \widetilde{B}_1 \\ \widetilde{B}_2 \end{array}\right] \\ &= \widetilde{D} + \widetilde{C}_1(z - F_4(z))^{-1}\widetilde{B}_1 + \widetilde{C}_1(z - F_4(z))^{-1}\widetilde{A}_2(zI_n - A_F)^{-1}\widetilde{B}_2 + \\ &\quad + \widetilde{C}_2(zI_n - A_F)^{-1}\widetilde{A}_3(z - F_4(z))^{-1}\widetilde{B}_1 + \widetilde{C}_2(zI_n - A_F)^{-1}\widetilde{B}_2 + \\ &\quad + \widetilde{C}_2(zI_n - A_F)^{-1}\widetilde{A}_3(z - F_4(z))^{-1}\widetilde{A}_2(zI_n - A_F)^{-1}\widetilde{B}_2 \\ &= F_1(z) + F_2(z)(z - F_4(z))^{-1}F_3(z), \end{aligned}$$

which finally proves that $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ is a state-space realization of $\widetilde{G}(z)$ with $(n+1)$-dimensional state-space. $\qquad\square$

The following proposition characterizes the nature of minimal balanced state-space realizations, in discrete time, of rational lossless functions. The fact that a lossless function admits a minimal balanced realization for which the realization matrix is unitary is well-known from the literature; see, e.g., [12]. However, the converse result, which asserts that unitary realization matrices correspond to possibly non-minimal realizations of lossless functions, appears to be novel and generalizes [14, Prop. 2.4] to the multivariable case.

Proposition 3.2 *(i) For any minimal balanced realization of a $p \times p$ rational lossless function the observability and controllability Gramians are both equal to the identity matrix and the associated realization matrix is unitary.*
(ii) Conversely, if the realization matrix associated with a realization (A, B, C, D) of order n of some $p \times p$ rational function G is unitary, then G is lossless of McMillan degree $\leq n$. The realization is minimal if and only if A is asymptotically stable and then it is balanced.

Proof. (i) As is well known, any minimal realization (A, B, C, D) of a lossless function G is asymptotically stable and satisfies (see [12])

$$
\begin{bmatrix} D & C \\ B & A \end{bmatrix} \begin{bmatrix} I_p & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} D^* & B^* \\ C^* & A^* \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & P \end{bmatrix}, \tag{35}
$$

for some unique positive definite $n \times n$ matrix P which is precisely the controllability Gramian. Since P is positive definite, it can be reduced to identity by using any state space transformation matrix T such that $P = TT^*$, producing a new minimal state space realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}) = (T^{-1}AT, T^{-1}B, CT, D)$. Then Eqn. (35) asserts that the associated new realization matrix is unitary. Upon changing the order of multiplication of this unitary realization matrix and its Hermitian transpose it follows that the new observability Gramian is also equal to identity. Hence, $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ is balanced. Any other minimal balanced realization of $G(z)$ can be obtained from $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ by means of a unitary change of basis of the state space, which leaves the associated realization matrix unitary.

(ii) Conversely, if a realization (A, B, C, D) of some proper rational function G has a unitary realization matrix, then it holds that

$$
\begin{cases} DD^* + CC^* &=& I_p, \\ BD^* + AC^* &=& 0, \\ DB^* + CA^* &=& 0, \\ BB^* + AA^* &=& I_n. \end{cases} \tag{36}
$$

For $z, w \in \mathbb{C}$ outside the spectrum of A, one may compute

$$
\begin{aligned}
I_p - G(z)G(w)^* &=& I_p - (D + C(zI_n - A)^{-1}B)(D^* + B^*(\overline{w}I_n - A^*)^{-1}C^*) = \\
&=& CC^* + C(zI_n - A)^{-1}AC^* + CA^*(\overline{w}I_n - A^*)^{-1}C^* + \\
&& -C(zI_n - A)^{-1}(I_n - AA^*)(\overline{w}I_n - A^*)^{-1}C^* = \\
&=& (z\overline{w} - 1)C(zI_n - A)^{-1}(\overline{w}I_n - A^*)^{-1}C^*.
\end{aligned}
$$

Setting $w = z$, it follows that G is lossless. Therefore, G has all its poles inside the open unit disk, which makes asymptotic stability of A a necessary condition for minimality of (A, B, C, D). On the other hand asymptotic stability of A implies that minimality of (A, B, C, D) is equivalent to the existence of positive definite solutions to the Lyapunov-Stein equations (9)-(10). Unitarity of the realization matrix, however, implies that I_n is always a positive definite solution to both these equations, showing asymptotic stability of A to be a sufficient condition for minimality of (A, B, C, D) as well. Clearly in such a case (A, B, C, D) is also balanced.   □

Remark. Let G_1 and G_2 be two lossless functions of McMillan degree n_1 and n_2, respectively, with minimal realizations (A_1, B_1, C_1, D_1) and (A_2, B_2, C_2, D_2) both having unitary associated realization matrices. Then the cascade realization (A, B, C, D) of the lossless function $G = G_1 G_2$ is obtained as

$$
\left[\begin{array}{c|c} D & C \\ \hline B & A \end{array} \right] = \left[\begin{array}{c|cc} D_1 & 0 & C_1 \\ \hline 0 & I_{n_2} & 0 \\ B_1 & 0 & A_1 \end{array} \right] \left[\begin{array}{c|cc} D_2 & C_2 & 0 \\ \hline B_2 & A_2 & 0 \\ 0 & 0 & I_{n_1} \end{array} \right] = \left[\begin{array}{c|cc} D_1 D_2 & D_1 C_2 & C_1 \\ \hline B_2 & A_2 & 0 \\ B_1 D_2 & B_1 A_2 & A_1 \end{array} \right]
$$

which yields again a unitary realization matrix, but which may fail to be minimal.

Three properties of the class of mappings $\mathcal{F}_{U,V}$ are collected in the following lemma for later use. These properties follow straightforwardly from the definition of the mapping $\mathcal{F}_{U,V}$ with U, V arbitrary $(p+1) \times (p+1)$ matrices.

Lemma 3.3 *Let U and V be arbitrary $(p+1)\times(p+1)$ matrices and P and Q both $p \times p$ matrices. Then we have:*

(i) $\mathcal{F}_{U\begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix},V\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}}(G(z)) = \mathcal{F}_{U,V}(QG(z)P^*),$

(ii) $\widetilde{G}(z) = \mathcal{F}_{U,V}(G(z)) \Leftrightarrow \widetilde{G}^*(z) = \mathcal{F}_{V,U}(G^*(z)),$

(iii) $\mathcal{F}_{\begin{bmatrix} I_p & 0 \\ 0 & \xi \end{bmatrix}U,\begin{bmatrix} I_p & 0 \\ 0 & \xi \end{bmatrix}V}(G(z)) = \mathcal{F}_{U,V}(G(z)),$ *for all $\xi \in \mathbb{C}$ with $|\xi| = 1$.*

The next property shows some of the importance of the $\mathcal{F}_{U,V}$ mappings in relation to lossless systems: if U and V are unitary matrices then lossless systems are mapped into lossless systems. The McMillan degree increases by at most one under this mapping.

Proposition 3.4 *Let U and V be unitary $(p+1) \times (p+1)$ matrices.*
The mapping $\mathcal{F}_{U,V}$ sends a lossless function of order n to a lossless function of order $\leq n+1$.

Proof. Since $G(z)$ is lossless of order n it has a minimal realization with a unitary realization matrix of size $n + p$. Formula (34) gives a realization of $\widetilde{G} = \mathcal{F}_{U,V}(G)$ which is unitary of size $n + p + 1$. Therefore, $\widetilde{G}(z)$ is a lossless function of order at most $n + 1$. □

Starting with a $p \times p$ unitary matrix G_0, interpreted as a lossless transfer matrix of order zero, and applying the previous proposition repeatedly, one obtains a class of lossless transfer matrices and corresponding realizations. The class that is obtained after n such recursion steps is parametrized by the finite sequence of unitary matrices $U_1, V_1, U_2, V_2, \ldots, U_n, V_n$, together with the unitary matrix G_0. It consists of lossless systems of order $\leq n + 1$. In Section 6 the question which lossless transfer matrix functions can be obtained in this way will be treated. This generalizes the results of [14] for the SISO case, where $U_k = I_2$ and V_k is a 2×2 Householder reflection matrix, for $k = 1, \ldots, n$.

4   Linear fractional transformations

Linear fractional transformations occur extensively in representation formulas for the solution of various interpolation problems [3]. They are at the heart of the parametrization of lossless functions through the tangential Schur algorithm. The properties of a linear fractional transformation of matrices were studied in a form adapted to the needs of J-theory in [21] and in an

algebraic setting in [23]. For more details and certain proofs on the basic results presented in this section, we refer to these papers.

Let $\Theta \in \mathbb{C}^{2p \times 2p}(z)$ be an invertible rational matrix in the variable z, block partitioned as $\Theta = \begin{bmatrix} \Theta_1 & \Theta_2 \\ \Theta_3 & \Theta_4 \end{bmatrix}$ with blocks Θ_i $(i = 1, \ldots, 4)$ of size $p \times p$.

Associated with Θ, let the linear fractional transformations $\mathcal{T}_\Theta$ and $\widehat{\mathcal{T}}_\Theta$ be defined to act on rational matrices $G \in \mathbb{C}^{p \times p}(z)$ in the following way:

$$\mathcal{T}_\Theta : G \mapsto (\Theta_4 G + \Theta_3)(\Theta_2 G + \Theta_1)^{-1}, \tag{37}$$

$$\widehat{\mathcal{T}}_\Theta : G \mapsto (G \Theta_2 + \Theta_4)^{-1}(G \Theta_1 + \Theta_3). \tag{38}$$

The domains of these mappings are denoted by $\mathcal{M}_\Theta$ and $\widehat{\mathcal{M}}_\Theta$, respectively. They consist of those G for which the expressions $\Theta_2 G + \Theta_1$ and $G \Theta_2 + \Theta_4$, respectively, have full rank p. Due to the invertibility assumption on Θ it follows that these domains are generic subsets of $\mathbb{C}^{p \times p}(z)$.

It is easily established that the following group properties hold for the LFTs associated with two invertible matrices Φ and Ψ:

$$\mathcal{T}_\Phi \circ \mathcal{T}_\Psi = \mathcal{T}_{\Phi\Psi}, \tag{39}$$

$$\widehat{\mathcal{T}}_\Phi \circ \widehat{\mathcal{T}}_\Psi = \widehat{\mathcal{T}}_{\Psi\Phi}. \tag{40}$$

Here it may be noted that the compositions of mappings on the left hand sides of these equations may formally have domains which are strict subsets of the domains of the mappings on the right hand sides. However, since the domains of the compositions of mappings on the left-hand sides are generic too, this issue is resolved by identifying (compositions of) LFTs if they agree on the intersection of their domains, which will be our policy throughout this paper.

It then holds that the LFTs $\mathcal{T}_\Theta$ and $\widehat{\mathcal{T}}_\Theta$ are bijections from their domains $\mathcal{M}_\Theta$ and $\widehat{\mathcal{M}}_\Theta$ to their co-domains $\mathcal{M}_{\Theta^{-1}}$ and $\widehat{\mathcal{M}}_{\Theta^{-1}}$, and that every LFT can be represented in each of the two forms described above. We have the following lemma.

Lemma 4.1 *Let $\Theta \in \mathbb{C}^{2p \times 2p}(z)$ be an invertible rational matrix. Then:*
(i) $\mathcal{T}_\Theta$ and $\widehat{\mathcal{T}}_\Theta$ are bijections. Their inverses are given by

$$\mathcal{T}_\Theta^{-1} = \mathcal{T}_{\Theta^{-1}} = \widehat{\mathcal{T}}_{J\Theta J},$$

$$\widehat{\mathcal{T}}_\Theta^{-1} = \widehat{\mathcal{T}}_{\Theta^{-1}} = \mathcal{T}_{J\Theta J}.$$

(ii) $\mathcal{T}_\Theta = \widehat{\mathcal{T}}_{J\Theta^{-1}J}$ and $\widehat{\mathcal{T}}_\Theta = \mathcal{T}_{J\Theta^{-1}J}$.

We now proceed to study the conditions under which two LFTs coincide. From the literature, the following result is well known.

Lemma 4.2 *Let $\Phi, \Psi \in \mathbb{C}^{2p \times 2p}(z)$ be two invertible rational matrices. It holds that $\mathcal{T}_\Phi = \mathcal{T}_\Psi$ if and only if there exists a scalar function $\lambda \in \mathbb{C}(z)$ for which $\Phi = \lambda \Psi$.*

For our purposes, however, it will be convenient to develop a slightly specialized version of this lemma, which states that the same result remains to hold true if two LFTs $\mathcal{T}_\Phi$ and $\mathcal{T}_\Psi$ are merely required to coincide on the subset of $p \times p$ lossless functions.

Lemma 4.3 *Let* $\Phi, \Psi \in \mathbb{C}^{2p \times 2p}(z)$ *be two invertible rational matrices. If the* $p \times p$ *lossless functions are all contained in* $\mathcal{M}_\Phi \cap \mathcal{M}_\Psi$, *and if for all lossless* G *it holds that* $\mathcal{T}_\Phi(G) = \mathcal{T}_\Psi(G)$, *then there exists a scalar function* $\lambda \in \mathbb{C}(z)$ *for which* $\Phi = \lambda \Psi$.

Proof. Note that upon application of the bijection $\mathcal{T}_{\Psi^{-1}}$, we have that the identity $\mathcal{T}_\Phi(G) = \mathcal{T}_\Psi(G)$ becomes equivalent to $\mathcal{T}_{\Psi^{-1}\Phi}(G) = G$. Denoting $\Theta = \Psi^{-1}\Phi$, this can be rewritten equivalently as $\Theta_4 G + \Theta_3 = G\Theta_2 G + G\Theta_1$, since $(\Theta_2 G + \Theta_1)^{-1}$ exists (for all lossless G). If G is lossless, then also $-G$ is lossless. Comparing the two expressions which follow from substitution of these values for G, it is found that for all lossless G

$$\Theta_4 G = G\Theta_1,$$
$$\Theta_3 = G\Theta_2 G.$$

Substitution of $G = I_p$ yields $\Theta_1 = \Theta_4$ and $\Theta_2 = \Theta_3$. Setting $G = z^{-1}I_p$, which is again lossless, yields $\Theta_3 = z^{-2}\Theta_2$. In combination with $\Theta_2 = \Theta_3$ it follows that $\Theta_2 = \Theta_3 = 0$.

We are left with the identity $\Theta_1 G = G\Theta_1$ for all lossless G. Note that every unitary matrix is lossless, so that G can be varied over all the products of signature matrices and permutation matrices. From this it easily follows that Θ_1 is a scalar multiple of the identity matrix. So there exists a scalar function λ such that $\Theta_1 = \Theta_4 = \lambda I_p$. (Rationality of λ is obvious from rationality of Θ, which is a consequence of rationality of Φ and Ψ.) Hence $\Theta = \lambda I_{2p}$, and therefore $\Phi = \lambda \Psi$. $\square$

Remark. The proof of Lemma 4.3 is such that it also applies to a setting in which all the rational matrix functions are restricted to be *real*. In the *complex* case, one could replace the choice $G = z^{-1}I_p$ (which is lossless of McMillan degree p) by the choice $G = iI_p$, so that the whole proof involves values for G from the set of unitary matrices only (i.e., G lossless of McMillan degree 0). In the real case with $p > 1$ a more careful analysis shows that a proof can also be designed using orthogonal matrices only. However, in the real case with $p = 1$, the only lossless functions of McMillan degree 0 are $G = \pm 1$. Here, lossless functions of McMillan degree ≥ 1 are needed to obtain the desired result.

For our purposes, it is important to study in some more detail the situation where the $2p \times 2p$ rational matrix Θ is J-inner and the $p \times p$ rational matrices G on which the LFT $\mathcal{T}_\Theta$ acts are lossless.

Proposition 4.4 *Let* Θ *be* J*-inner of size* $2p \times 2p$ *and of McMillan degree* m. *If* G *is* $p \times p$ *lossless and of McMillan degree* n, *then* $G \in \mathcal{M}_\Theta$ *and the matrix function* $\widehat{G} = \mathcal{T}_\Theta(G)$ *is also lossless and of McMillan degree* $\leq n + m$.

Proof. The function Θ being J-inner, condition (14) implies for any fixed point z of analyticity with $|z| > 1$ that

$$\Theta_1(z)\Theta_1(z)^* \geq I_p + \Theta_2(z)\Theta_2(z)^*,$$

so that $\Theta_1(z)\Theta_1(z)^*$ is positive definite and $\Theta_1(z)$ invertible. Multiplying the last inequality by $\Theta_1(z)^{-1}$ on the left and by $\Theta_1(z)^{-*}$ on the right, we can see that

$$\|\Theta_1(z)^{-1}\Theta_2(z)\| < 1, \quad |z| > 1,$$

where $\|\cdot\|$ denotes the usual operator norm, which is computed as the largest singular value of the matrix. If G is lossless it follows that $\|G(z)\| \leq 1$ for $|z| > 1$, and therefore

$$\|\Theta_1(z)^{-1}\Theta_2(z)G(z)\| < 1, \quad |z| > 1.$$

This implies that $\Theta_2(z)G(z) + \Theta_1(z)$ is invertible. (Indeed, if x would be a nontrivial vector in the kernel of $\Theta_2(z)G(z) + \Theta_1(z)$, then $\Theta_1(z)^{-1}\Theta_2(z)G(z)x = -x$ would imply $\|\Theta_1(z)^{-1}\Theta_2(z)G(z)\| \geq 1$.) Since at most m values of z are excluded from the domain $|z| > 1$ (the poles of $\Theta(z)$), it follows that $\Theta_2 G + \Theta_1$ is invertible as a rational matrix and $\mathcal{T}_\Theta(G)$ is well-defined.

Now $\widehat{G}$ satisfies the relation

$$\begin{bmatrix} I_p \\ \widehat{G} \end{bmatrix} = \Theta \begin{bmatrix} I_p \\ G \end{bmatrix} (\Theta_2 G + \Theta_1)^{-1}. \tag{41}$$

We thus have

$$I_p - \widehat{G}^*\widehat{G} = \begin{bmatrix} I_p & \widehat{G}^* \end{bmatrix} J \begin{bmatrix} I_p \\ \widehat{G} \end{bmatrix} = (\Theta_2 G + \Theta_1)^{-*} \begin{bmatrix} I_p & G^* \end{bmatrix} \Theta^* J \Theta \begin{bmatrix} I_p \\ G \end{bmatrix} (\Theta_2 G + \Theta_1)^{-1},$$

which proves that Θ being J-inner and G lossless, $\widehat{G}$ is also lossless (see Sect. 2).

The fact that the McMillan degree of $\widehat{G}$ is at most $n + m$ is not shown here, but constitutes a special case of Proposition A.1; see Appendix A. $\square$

Also, if the LFT corresponds to a J-inner matrix, we have the following elementary results.

Lemma 4.5 *Let $\Theta(z)$ be J-inner of size $2p \times 2p$.*
(a) The mappings $\mathcal{T}_\Theta$ and $\widehat{\mathcal{T}}_{\Theta^\sharp}$ coincide.
(b) If P and Q are $p \times p$ unitary matrices, then $\Theta_{P,Q}(z) := \Theta(z) \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$ is again J-inner, and for all $G \in \mathcal{M}_{\Theta_{P,Q}}$:

$$\mathcal{T}_{\Theta_{P,Q}(z)}(G(z)) = \mathcal{T}_{\Theta(z)}(QG(z)P^*).$$

(c) Let $K = \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix}$. Then the function defined by

$$\Theta^o(z) := K\Theta(z^{-1})K \tag{42}$$

is again J-inner, and for all $G \in \mathcal{M}_\Theta$ it holds that $\mathcal{T}_\Theta(G)^* = \mathcal{T}_{\Theta^o}(G^*)$.

Proof. (a) Since Θ is J-inner, it holds that $\widehat{\mathcal{T}}_{\Theta^\sharp} = \widehat{\mathcal{T}}_{J\Theta^{-1}J} = \mathcal{T}_\Theta$.

(b) The proof that $\Theta_{P,Q}(z)$ is again J-inner is elementary. The rest of this statement follows from the group property of LFTs and from the simple action of an LFT associated with a block-diagonal matrix.

(c) Note that $\mathcal{T}_\Theta(G)^* = (G^*\Theta_2^* + \Theta_1^*)^{-1}(G^*\Theta_4^* + \Theta_3^*) = \widehat{\mathcal{T}}_{K\Theta^*K}(G^*) = \mathcal{T}_{JK(\Theta^*)^{-1}KJ}(G^*) = \mathcal{T}_{K(\Theta^*)^\sharp K}(G^*) = \mathcal{T}_{\Theta^o}(G^*)$. The proof that $\Theta^o(z)$ is again J-inner is elementary upon observing that $KJ = -JK$. $\square$

A particular situation of interest occurs when the J-inner matrix function $\Theta(z)$ happens to be of McMillan degree 0. In this case $\Theta(z) = M$ is a constant J-unitary matrix and the associated class of LFTs is the one of *generalized Möbius transformations* $\mathcal{T}_M$. The following theorem indicates how the action of such transformations $\mathcal{T}_M$ on *lossless* functions $G(z)$ can be represented entirely in terms of *balanced* state-space realizations.

Theorem 4.6 *The linear fractional transformation $\mathcal{T}_M$ associated with a constant J-unitary matrix M is a bijection on the set of lossless functions which preserves the McMillan degree. Let $M = \begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix}$ be the block decomposition of M and let $G(z)$ be a lossless function with minimal state-space realization (A, B, C, D). Then, a minimal state-space realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ for $\widetilde{G}(z) = \mathcal{T}_M(G(z))$ is given by:*

$$\begin{cases} \widetilde{A} &= A - B(M_2 D + M_1)^{-1} M_2 C, \\ \widetilde{B} &= B(M_2 D + M_1)^{-1}, \\ \widetilde{C} &= [M_4 - (M_4 D + M_3)(M_2 D + M_1)^{-1} M_2]C, \\ \widetilde{D} &= (M_4 D + M_3)(M_2 D + M_1)^{-1}. \end{cases} \tag{43}$$

If in addition (A, B, C, D) is balanced, then this realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ is also balanced.

Proof. A J-unitary constant matrix M is in particular J-inner and its inverse is also J-inner, so that $\mathcal{T}_M$ is clearly, by Prop. 4.4, a bijection on the set of lossless functions.

The computation on the realization is just a particular case of Proposition A.1 (i), see Appendix A. This shows that the McMillan degree does not increase under the action of $\mathcal{T}_M$. It cannot decrease either, because the inverse mapping $\mathcal{T}_{M^{-1}} = \mathcal{T}_M^{-1}$ can be treated likewise, and also does not increase the McMillan degree. Hence, the McMillan degree is preserved by $\mathcal{T}_M$.

Finally, recall that a minimal realization of a lossless function is balanced if and only if the associated realization matrix is unitary. Starting from a minimal balanced realization (A, B, C, D) it now can be verified by direct computation that the realization matrix associated with $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ is also unitary and thus the realization is balanced. $\square$

5 The tangential Schur algorithm

In this section we outline the use of the tangential Schur algorithm for the recursive construction of a parametrization of the space of stable all-pass systems of fixed finite order. It is derived from

the method of [1, 10], where the tangential Schur algorithm is used to construct an infinite atlas of generic overlapping parametrizations for the space of $p \times p$ inner functions of McMillan degree n. The relationship between these two situations is constituted by the map $R(z) \mapsto R(z)^{-1}$, or equivalently $R(z) \mapsto R^{\sharp}(z)$ (see also Eqns. (11) and (17)), which is used to relate the space of $p \times p$ inner functions of McMillan degree n to the space of $p \times p$ stable all-pass systems of McMillan degree n.

In the context of inner functions, the tangential Schur algorithm consists of an iterative procedure by which a given $p \times p$ inner function of McMillan degree n is reduced in n iteration steps to a $p \times p$ inner function of McMillan degree 0 (i.e., to a constant unitary matrix). In each iteration step the McMillan degree of the inner function at hand is reduced by 1, by application of a suitable linear fractional transformation which is chosen to meet a particular interpolation condition, and which involves an associated J-inner matrix function of McMillan degree 1 (or elementary J-inner factor, see Section 2). The actual parametrization procedure consists of the reverse process, by which a chart of inner functions of McMillan degree n is constructed in n iteration steps, starting from an initial unitary matrix. The choice of interpolation points and (normalized) direction vectors may serve to index such a chart, while the local coordinates correspond to the n Schur vectors.

In the Schur algorithm, the elementary J-inner factors involved are of the form (28) and it suffices to consider those which are analytic on the closed unit disk (see the remark at the end of the section). Such an elementary J-inner factor $\Theta(z)$, having its pole outside the closed unit disk at $z = 1/\overline{w}$, can be represented as:

$$\Theta(u, v, w, \xi, H)(z) = \left(I_{2p} + \left(\frac{b_w(z)}{b_w(\xi)} - 1 \right) \frac{\begin{bmatrix} u \\ v \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}^* J}{(1 - \|v\|^2)} \right) H, \tag{44}$$

where
(1) $w \in \mathbb{C}$, $|w| < 1$,
(2) $u \in \mathbb{C}^{p \times 1}$, $\|u\| = 1$,
(3) $v \in \mathbb{C}^{p \times 1}$, $\|v\| < 1$,
(4) $\xi \in \mathbb{C}$, $|\xi| = 1$,
(5) H is a $2p \times 2p$ constant J-unitary matrix.

This representation is obtained from (28) in which some normalizing conditions have been imposed:

(i) the vector x has been rewritten $x = \sigma \begin{bmatrix} u \\ v \end{bmatrix}$, where σ is a nonzero scalar, $\|u\| = 1$ and thus $\|v\| < 1$, since (29) must be satisfied.

(ii) an arbitrary complex number ξ on the unit circle has been introduced which allows to write the right J-unitary factor H as $H = \Theta(\xi)$.

Since any nonzero scaling of the vector x does not affect the expression (28), we have that

$$\Theta(u, v, w, \xi, H)(z) = X_x \left(b_w(z) b_w(\xi)^{-1} \right) H = X_x \left(b_w(z) \right) X_x \left(b_w(\xi) \right)^{-1} H, \tag{45}$$

and $\Theta(u, v, w, \xi, H)(z)$ is of the forn (28) up to a right constant J-unitary factor. The constant matrix $X_x(b_w(\xi)^{-1}) = X_x(b_w(\xi))^{-1}$ is indeed J-unitary, since $|\xi| = 1$ yields $|b_w(\xi)| = 1$ (see Lemma 2.2).

Once ξ is chosen arbitrarily on the unit circle, the representation (44) is unique up to a unimodular complex number, which corresponds to the remaining freedom in the scaling of the vector x. It will turn out that for several of our purposes, the additional freedom to choose ξ and H is of rather limited use. In the literature on the parametrization of inner functions it therefore often happens that fixed choices are made, in particular $\xi = 1$ and $H = I_{2p}$ (see, e.g., [1, 10]). However, especially the role played by H cannot be ignored in establishing the precise connection between the tangential Schur algorithm and the state-space approach involving the mappings $\mathcal{F}_{U,V}$ in Section 6. Also, for the construction of atlases of overlapping parametrizations, the additional freedom to choose H leads to useful new alternatives, as will be discussed in Section 7. Therefore, we will not fix a choice yet and present our results for general ξ and H.

The following lemma establishes a useful new factorization result for the class of J-unitary matrices $X_x(\alpha)$ with $x^* J x > 0$.

Lemma 5.1 *Let $x \in \mathbb{C}^{2p \times 1}$ be such that $x^* J x > 0$. Let x be written as $x = \sigma \begin{bmatrix} u \\ v \end{bmatrix}$, where σ is any nonzero scalar such that $u, v \in \mathbb{C}^{p \times 1}$ satisfy $\|u\| = 1$ and $\|v\| < 1$. It then holds that*

$$X_x(\alpha) = H(uv^*) X_{\begin{bmatrix} u \\ 0 \end{bmatrix}}(\alpha) H(uv^*)^{-1}, \tag{46}$$

where $H(uv^*)$ denotes the Halmos extension of the strictly contractive matrix uv^*.

Proof. Note that σ, u and v with the required properties always exist, making uv^* into a strictly contractive matrix. The Halmos extension $H(uv^*)$ is not difficult to compute explicitly as the Hermitian J-unitary matrix

$$H(uv^*) = \begin{bmatrix} I_p - (1 - \frac{1}{\sqrt{1 - \|v\|^2}}) uu^* & \frac{1}{\sqrt{1 - \|v\|^2}} uv^* \\ \frac{1}{\sqrt{1 - \|v\|^2}} vu^* & I_p - (1 - \frac{1}{\sqrt{1 - \|v\|^2}}) \frac{vv^*}{\|v\|^2} \end{bmatrix}, \tag{47}$$

and it satisfies

$$H(uv^*) \begin{bmatrix} u \\ -v \end{bmatrix} = \sqrt{1 - \|v\|^2} \begin{bmatrix} u \\ 0 \end{bmatrix}.$$

It then follows that $X_x(\alpha) H(uv^*) = H(uv^*) + \dfrac{(\alpha - 1)}{(1 - \|v\|^2)} \begin{bmatrix} u \\ v \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}^* J H(uv^*) = H(uv^*) +$

$\dfrac{(\alpha - 1)}{\sqrt{1 - \|v\|^2}} \begin{bmatrix} u \\ v \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix}^* = H(uv^*) + (\alpha - 1) H(uv^*) \begin{bmatrix} u \\ 0 \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix}^* J = H(uv^*) X_{\begin{bmatrix} u \\ 0 \end{bmatrix}}(\alpha)$, which proves the lemma. $\square$

The following proposition plays a central role in our construction of balanced parametrizations for discrete-time lossless transfer functions.

Proposition 5.2 *The J-inner matrix function* $\Theta(u, v, w, \xi, H)(z)$ *can be factorized as:*

$$\Theta(u, v, w, \xi, H)(z) = H(uv^*)S_{u,w}(z)S_{u,w}(b_w(\xi))^{-1}H(uv^*)^{-1}H, \tag{48}$$

where $H(uv^*)$ *denotes the Halmos extension of the strictly contractive matrix* uv^*, *and where*

$$S_{u,w}(z) := X_{\left[\begin{array}{c} u \\ 0 \end{array}\right]}(b_w(z)) = \left[\begin{array}{cc} I_p - (1 - b_w(z))\, uu^* & 0 \\ 0 & I_p \end{array}\right]. \tag{49}$$

Proof. This is an immediate consequence of the expression (45) of $\Theta(u, v, w, \xi, H)(z)$, in combination with Lemma 5.1, in which α is replaced by the Blaschke factor $b_w(z)$.   □

The next proposition establishes that application of the linear fractional transformation associated with the elementary J-inner function $\Theta(u, v, w, \xi, H)(z)$ to a lossless function of McMillan degree n, results in a lossless function of McMillan degree $n + 1$ which satisfies an interpolation condition involving only u, v and w.

Proposition 5.3 *Let* $\Theta(z) = \Theta(u, v, w, \xi, H)(z)$ *be the elementary J-inner function defined in Eqn. (44), with u, v, w, ξ and H satisfying the accompanying properties. Let $G(z)$ be a lossless function of McMillan degree n. Then*

$$\widehat{G}(z) = \mathcal{T}_{\Theta(z)}(G(z)) = \widehat{\mathcal{T}}_{\Theta^\sharp(z)}(G(z)) \tag{50}$$

is lossless of McMillan degree $n + 1$ and satisfies the interpolation condition

$$\widehat{G}(1/\overline{w})u = v. \tag{51}$$

Proof. Proposition 4.4 asserts that $\widehat{G}(z)$ is lossless. From Proposition 5.2 we have that

$$\mathcal{T}_{\Theta(z)} = \mathcal{T}_{H(uv^*)} \circ \mathcal{T}_{S_{u,w}(z)} \circ \mathcal{T}_M,$$

where $M = S_{u,w}(\xi)^{-1}H(uv^*)^{-1}H$. The linear fractional transformations associated with the J-unitary matrices M and $H(uv^*)$ are both generalized Möbius transformations which leave the McMillan degree unchanged, see Theorem 4.6. Only the linear fractional transformation associated with the matrix $S_{u,w}(z)$ can change the McMillan degree, but it has a simple form: for the lossless function $R(z) := \mathcal{T}_M(G(z))$ of McMillan degree n it holds that

$$\widehat{R}(z) := \mathcal{T}_{S_{u,w}(z)}(R(z)) = R(z)(I_p - (1 - b_w^\sharp(z))uu^*).$$

Because $I_p - (1 - b_w^\sharp(z))uu^*$ is the transfer matrix of a lossless system and, by (21), $\det \widehat{R}(z) = b_w^\sharp(z) \det R(z)$, it follows from Lemma 2.1 and from the fact that stability ensures that no common

factors can occur, that $\widehat{R}(z)$ has McMillan degree $n + 1$. Then also $\widehat{G}(z) := \mathcal{T}_{H(uv^*)}(\widehat{R}(z))$ has McMillan degree $n + 1$.

For the matrix $X_x(\alpha)$ it holds that $x^* J X_x(\alpha) = \alpha x^* J$. Setting $x = \begin{bmatrix} u \\ v \end{bmatrix}$, $\alpha = b_w(z)$ and $z = w$ this results in the following identity, in view of Eqn. (45):

$$\begin{bmatrix} u^* & -v^* \end{bmatrix} \Theta(w) = 0. \tag{52}$$

Application of the operator $^\sharp$ yields the equivalent relationship

$$\Theta^\sharp(1/\overline{w}) \begin{bmatrix} u \\ -v \end{bmatrix} = 0. \tag{53}$$

Note that $\widehat{G}(z) = \widehat{\mathcal{T}}_{\Theta^\sharp(z)}(G(z))$ satisfies the relation

$$\begin{bmatrix} \widehat{G}(z) & I_p \end{bmatrix} = (G(z)\Theta_3^\sharp(z) + \Theta_4^\sharp(z))^{-1} \begin{bmatrix} G(z) & I_p \end{bmatrix} \Theta^\sharp(z). \tag{54}$$

The matrix $G(z)\Theta_3^\sharp(z) + \Theta_4^\sharp(z)$ is analytic and invertible everywhere outside the open unit disk, which follows in a manner entirely analogous to the proof of invertibility of $\Theta_2(z)G(z) + \Theta_1(z)$ at points of analyticity outside the unit disk, see the proof of Proposition 4.4. Setting $z = 1/\overline{w}$ the interpolation condition (51) now follows. $\square$

Remark. The relations (53) and (41) yield an alternative form of the interpolation condition (51):

$$u^* = v^* \widehat{G}(w) \tag{55}$$

provided that the matrix $\Theta_2(w)G(w) + \Theta_1(w)$ is invertible. Equivalently, this requires that w does not show up as a pole of $\widehat{G}$. Only then the form (55) is valid, while the form (51) applies always. This illustrates the usefulness of the identity $\widehat{\mathcal{T}}_{\Theta^\sharp} = \mathcal{T}_\Theta$ when dealing with analyticity. In a similar vein, note that for $w = 0$ the function $\Theta(u, v, w, \xi, H)$ is not proper and does not admit a state-space realization, so that part (i) of Proposition A.1 cannot be used to compute the action of the associated LFT entirely in terms of state-space realizations. However, the matrix function $\Theta(u, v, w, \xi, H)^\sharp$ does admit a state-space realization, so that part (ii) of Proposition A.1 can always be used instead to enable this nevertheless.

The following proposition addresses the reverse process and constructs from a lossless matrix function $G(z)$ of McMillan degree $n + 1$, a new lossless matrix function of reduced degree n. To achieve this, it proceeds from an interpolation condition of the form (51). Though this result is well-known (see, e.g., [1] where it follows from the existence of a certain isometry), we find it of interest to give a more constructive proof of it.

Proposition 5.4 *Let $\widehat{G}$ be a $p \times p$ lossless function of McMillan degree $n + 1$ which satisfies an interpolation condition of the form*

$$\widehat{G}(1/\overline{w})u = v, \tag{56}$$

in which $w \in \mathbb{C}$ is an interpolation point with $|w| < 1$, $u \in \mathbb{C}^{p \times 1}$ is a direction vector with $\|u\| = 1$ and $v \in \mathbb{C}^{p \times 1}$ is a Schur vector satisfying $\|v\| < 1$.

Let $\xi \in \mathbb{C}$ be an arbitrary number of modulus 1, and let H be an arbitrary constant J-unitary matrix. Then $\widehat{G}$ admits the representation

$$\widehat{G} = \mathcal{T}_{\Theta(u,v,w,\xi,H)}(G), \tag{57}$$

for some $p \times p$ lossless function G of McMillan degree n.

Proof. The interpolation condition (56) can be rewritten as follows:

$$
\begin{aligned}
\widehat{G}(1/\overline{w})u = v \quad &\Leftrightarrow \quad 0 = \left[\begin{array}{cc} \widehat{G}(1/\overline{w}) & I_p \end{array}\right]\left[\begin{array}{c} u \\ -v \end{array}\right] \\
&\Leftrightarrow \quad 0 = \left[\begin{array}{cc} \widehat{G}(1/\overline{w}) & I_p \end{array}\right] H(uv^*)^{-1} H(uv^*)\left[\begin{array}{c} u \\ -v \end{array}\right] \\
&\Leftrightarrow \quad 0 = \left[\begin{array}{cc} \widehat{G}(1/\overline{w}) & I_p \end{array}\right] H(uv^*)^{-1}\left[\begin{array}{c} u \\ 0 \end{array}\right] \\
&\Leftrightarrow \quad 0 = \widehat{\mathcal{T}}_{H(uv^*)^{-1}}(\widehat{G})(1/\overline{w})u.
\end{aligned}
$$

The lossless function $\widehat{R} := \widehat{\mathcal{T}}_{H(uv^*)^{-1}}(\widehat{G}) = \mathcal{T}_{H(uv^*)^{-1}}(\widehat{G})$ has the same McMillan degree $n+1$ as $\widehat{G}$ and satisfies the interpolation condition

$$\widehat{R}(1/\overline{w})u = 0.$$

This means that $\widehat{R}(z)$ has a zero at $z = 1/\overline{w}$, or equivalently a pole at $z = w$. Now let $B_w(z)$ be defined as

$$B_w(z) := I_p - (1 - b_w^\sharp(z))uu^* = (I_p - (1 - b_w(z))uu^*)^{-1}, \tag{58}$$

which is $p \times p$ lossless of McMillan degree 1, satisfying $\det B_w(z) = b_w^\sharp(z)$ and having its pole at $z = w$. Then $\widehat{R}(z)$ can be factored as

$$\widehat{R}(z) = R(z)B_w(z), \tag{59}$$

where by construction $R(z)$ is $p \times p$ lossless of McMillan degree n. (This is a matrix version of the Schwartz lemma on which the Potapov factorization of lossless and inner functions is based; see [20].) The identity (59) can be rewritten as

$$\widehat{R} = \mathcal{T}_{S_{u,w}}(R)$$

so that finally

$$\widehat{G} = \mathcal{T}_{\Theta(u,v,w,\xi,H)}(G),$$

where $G = \mathcal{T}_{M^{-1}}(R)$, with $M = S_{u,w}(\xi)^{-1}H(uv^*)^{-1}H$, is also lossless and of McMillan degree n.   □

Remark. In the standard case with $w = 0$ the value of $\widehat{G}(\infty)$ corresponds to the direct feedthrough term $\widehat{D}$ of any state-space realization $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ of $\widehat{G}(z)$, so that the interpolation condition then takes the form $\widehat{D}u = v$.

The tangential Schur algorithm can now be described. It is used in Section 7 to construct atlases of generic charts for the manifold of $p \times p$ lossless systems of degree n.

Theorem 5.5 (Tangential Schur algorithm) *Let $G^{(n)}$ be a $p \times p$ lossless transfer matrix of McMillan degree n. For $k = n, \ldots, 1$, let interpolation points $w_k \in \mathbb{C}$ be given with $|w_k| < 1$, let constants $\xi_k \in \mathbb{C}$ be given with $|\xi_k| = 1$, and let mappings $H_k : (u, v, w, \xi) \mapsto H_k(u, v, w, \xi)$ be given which assign a constant $2p \times 2p$ J-unitary matrix $H_k(u, v, w, \xi)$ to each quadruple $(u, v, w, \xi) \in \mathbb{C}^p \times \mathbb{C}^p \times \mathbb{C} \times \mathbb{C}$ with $\|u\| = 1$, $\|v\| < 1$, $|w| < 1$ and $|\xi| = 1$.*

Then for $k = n, \ldots, 1$, there exist vectors $u_k \in \mathbb{C}^p$ with $\|u_k\| = 1$, such that the vectors $v_k \in \mathbb{C}^p$ constructed recursively by the following formulas, all have length $\|v_k\| < 1$:

$$v_k := G^{(k)}(1/\overline{w}_k)u_k, \tag{60}$$

$$\Theta_k := \Theta(u_k, v_k, w_k, \xi_k, H_k(u_k, v_k, w_k, \xi_k)), \tag{61}$$

$$G^{(k-1)} := \mathcal{T}_{\Theta_k}^{-1}(G^{(k)}). \tag{62}$$

With such a choice of the unit vectors u_k ($k = n, \ldots, 1$) each of the functions $G^{(k)}$ is lossless of McMillan degree k and one can write

$$G^{(n)} = \mathcal{T}_{\Theta_n}(\mathcal{T}_{\Theta_{n-1}}(\ldots (\mathcal{T}_{\Theta_1}(G^{(0)}))\ldots)) = \mathcal{T}_{\Theta_n \Theta_{n-1} \cdots \Theta_1}(G^{(0)}), \tag{63}$$

where $G^{(0)}$ is a constant unitary matrix.

Proof. In view of Proposition 5.4, which is applied repeatedly, it only remains to show that if G is a lossless function of degree $n > 0$, then for every choice of interpolation point w with $|w| < 1$, a direction vector u with $\|u\| = 1$ exists for which the Schur vector $v := G(1/\overline{w})u$ satisfies $\|v\| < 1$. A function theoretic proof of this fact is given in [10, Lemma 6]. An alternative state-space argument runs as follows.

For $|w| < 1$, the mapping $z \mapsto b_{-w}(z)$ constitutes an automorphism of the unit disk. Thus, the function $\widetilde{G}(z) := G(b_{-w}(z))$ is again lossless (and also of McMillan degree $n > 0$). This shows that $G(1/\overline{w}) = \widetilde{G}(\infty)$ appears as the direct feedthrough matrix $\widetilde{D}$ of any state-space realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ of the lossless function $\widetilde{G}$. If such a realization is chosen to be minimal and balanced, then $\widetilde{D}$ is a proper submatrix of the associated unitary realization matrix; see Proposition 3.2. Clearly, some column of $\widetilde{D}$ must have length < 1, since otherwise all columns of $\widetilde{B}$ would be zero, contradicting minimality with order $n > 0$. Thus, a suitable vector u can always be found among the set of standard basis vectors $\{e_1, \ldots, e_p\}$. As a matter of fact, any normalized basis will do. $\square$

Remark. Similar results can be proved in which the interpolation condition (51), which involves a relationship among the *columns* of $\widehat{G}$, is replaced by an interpolation condition involving its *rows*. This can be done by applying Proposition 5.4 to the lossless function $\widehat{G}^*(z)$ and using part (c) of Lemma 4.5. The interpolation condition then takes the form

$$u^*\widehat{G}(1/w) = v^*, \tag{64}$$

and should not be confused with the alternative form (55) for the interpolation condition (51). As a counterpart to the result of Proposition 5.4, one also obtains an alternative linear fractional representation

$$\widehat{G} = \mathcal{T}_{\Theta(u,v,w,\xi,H)^\circ}(\widetilde{G})$$

for some $p \times p$ lossless function $\widetilde{G}$ of degree n. Observe that the J-inner function $\Theta(u, v, w, \xi, H)^o$ is now analytic outside the open unit disk. Thus, the theory of this section admits a dual theory in which elementary J-inner factors analytic outside the disk and interpolation conditions on the rows are used.

Note that the tangential Schur algorithm described above proceeds from a set of n interpolation conditions on the columns of a sequence of intermediate lossless functions. Obviously, a set of n interpolation conditions on the rows of a sequence of intermediate lossless functions can also be imposed. But since each iteration step involves a single interpolation condition only, *mixed* situations can also be treated in which n interpolation conditions are prescribed involving columns and rows.

6   Connection between the classes of mappings $\mathcal{T}_{\Theta(z)}$ and $\mathcal{F}_{U,V}$

In the scalar case, mappings of the form $\mathcal{F}_{U,V}$ with $U = I_2$ have been used in [14] to recursively construct a balanced canonical form for the space of discrete-time lossless systems of finite McMillan degree. The parameters that occur in this recursion have the interpretation of Schur parameters, corresponding to the situation with interpolation points w at zero. With this connection in mind, it is the purpose of the present section to clarify the relationship between the two classes of mappings $\mathcal{T}_{\Theta(z)}$ and $\mathcal{F}_{U,V}$, with $\Theta(z)$ a J-inner matrix function of McMillan degree 1 and U and V unitary. We are interested in investigating the possibilities for representing a mapping $\mathcal{T}_{\Theta(z)}$ in terms of a corresponding mapping $\mathcal{F}_{U,V}$. This will give us balanced state-space parametrizations directly in terms of the set of parameters u, v, w, ξ and H used in the tangential Schur algorithm. Moreover, unitary matrices will be involved in the computation of these realizations, and these are known to be numerically well-conditioned. Let $\mathcal{L}_n^p$ denote the set of rational $p \times p$ lossless transfer matrices of McMillan degree n and let $\mathcal{L}^p = \bigcup_{n=0}^{\infty} \mathcal{L}_n^p$.

Theorem 6.1 *Let U and V be unitary $(p+1) \times (p+1)$ matrices, block-partitioned as*

$$U = \left[\begin{array}{cc} \alpha_U & M_U \\ k_U & \beta_U^* \end{array} \right], \quad V = \left[\begin{array}{cc} \alpha_V & M_V \\ k_V & \beta_V^* \end{array} \right], \tag{65}$$

with k_U and k_V scalar and the other blocks of compatible sizes.
The mapping $\mathcal{F}_{U,V} : \mathcal{L}^p \to \mathcal{L}^p$ can be expressed as a linear fractional transformation $\mathcal{T}_\Phi : \mathcal{L}^p \to \mathcal{L}^p$ for some J-inner function Φ if and only if $(k_U, k_V) \neq (0, 0)$.
(i) If $|k_U| = |k_V| = 1$, then M_U and M_V are unitary matrices and $\mathcal{F}_{U,V}$ is the generalized Möbius transform $\mathcal{F}_{U,V}(G) = M_V G M_U^*$ which can be represented as $\mathcal{T}_\Phi(G)$ with a constant J-unitary matrix Φ.
(ii) If $(k_U, k_V) \neq (0, 0)$ and $|k_U| < 1$ or $|k_V| < 1$, then $\mathcal{F}_{U,V}$ can be represented as $\mathcal{T}_\Phi$ with an elementary J-inner function Φ having a pole at k_V / k_U, located at infinity for $k_U = 0$.

Proof. Because U and V are unitary, according to Proposition 3.4 the mapping $\mathcal{F}_{U,V}$ maps $\mathcal{L}^p$ into $\mathcal{L}^p$. If Φ is J-inner, then the mapping $\mathcal{T}_\Phi$ also maps $\mathcal{L}^p$ into $\mathcal{L}^p$ but in addition it is bijective, see Lemma 4.1 and Proposition 4.4. Thus, a necessary requirement for $\mathcal{F}_{U,V}$ to be representable

as a linear fractional transformation $\mathcal{T}_\Phi$ with a J-inner function Φ, is injectivity.

First consider the scalar case $p = 1$ with $(k_U, k_V) = (0, 0)$. Then U and V are 2×2 diagonal matrices and all lossless G are mapped to the unimodular constant $\alpha_V \alpha_U^*$. Therefore the mapping $\mathcal{F}_{U,V}$ obviously is not injective.

In the multivariable case $p > 1$ with $(k_U, k_V) = (0, 0)$, it also holds that $\mathcal{F}_{U,V}$ is not injective and thus cannot be represented as an LFT associated with a J-inner function. Indeed, let $P(\xi) = I_p + (\xi - 1)\beta_U \beta_U^*$, and $Q(\xi) = I_p + (\xi - 1)\beta_V \beta_V^*$ where $\xi \neq 1$ is a unimodular number. Then $P(\xi)$ and $Q(\xi)$ are unitary and we have

$$U \begin{bmatrix} 1 & 0 \\ 0 & P(\xi) \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & \xi \end{bmatrix} U, \quad V \begin{bmatrix} 1 & 0 \\ 0 & Q(\xi) \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & \xi \end{bmatrix} V,$$

so that by Lemma 3.3 (i), (iii), for every lossless function G,

$$\mathcal{F}_{U,V}(Q(\xi)G(z)P(\xi)^*) = \mathcal{F}_{\begin{bmatrix} I_p & 0 \\ 0 & \xi \end{bmatrix} U, \begin{bmatrix} I_p & 0 \\ 0 & \xi \end{bmatrix} V}(G(z)) = \mathcal{F}_{U,V}(G(z)).$$

Therefore injectivity of $\mathcal{F}_{U,V}$ requires that $Q(\xi)GP(\xi)^* = G$ for every lossless function G. However if this is the case, then $\beta_V \beta_V^* G = G \beta_U \beta_U^*$ for all lossless G. Choosing $G = I_p$ this implies $\beta_V \beta_V^* = \beta_U \beta_U^*$. Choosing $G = I_p - 2\gamma\gamma^*$, where $\gamma \in \mathbb{C}^p$ is a vector of length one which is independent of β_V and not orthogonal to β_V, it follows that $\beta_V \beta_V^* \gamma \gamma^* = \gamma\gamma^* \beta_U \beta_U^*$, which yields a contradiction. Therefore $\mathcal{F}_{U,V}$ is not injective.

Next, assume that $(k_U, k_V) \neq (0, 0)$. Let

$$J_1 = \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad J_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and consider the linear fractional transformation $\mathcal{L}$ defined by

$$\mathcal{L}(H(z)) = (J_1 H(z) + J_2)(J_2 H(z) + J_1)^{-1}.$$

This linear fractional transformation is well-defined for every $H(z)$ such that, in the block decomposition of $H(z)$

$$H(z) = \begin{bmatrix} H_1(z) & H_2(z) \\ H_3(z) & H_4(z) \end{bmatrix},$$

with $H_4(z)$ scalar, $H_4(z)$ is not identically zero. It can be computed as

$$\mathcal{L}(H) = \begin{bmatrix} H_1 - H_2 H_4^{-1} H_3 & H_2 H_4^{-1} \\ -H_4^{-1} H_3 & H_4^{-1} \end{bmatrix}.$$

The mapping $\mathcal{F}_{U,V}$ can now be written using this transformation as:

$$\mathcal{F}_{U,V}(G(z)) = \begin{bmatrix} I_p & 0 \end{bmatrix} \mathcal{L} \left(V \begin{bmatrix} 1 & 0 \\ 0 & G(z) \end{bmatrix} U^* - z \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} I_p \\ 0 \end{bmatrix}.$$

Now,

$$\mathcal{L} \left(V \begin{bmatrix} 1 & 0 \\ 0 & G(z) \end{bmatrix} U^* - z \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) =$$

$$= \left[J_1 \left(V \begin{bmatrix} 1 & 0 \\ 0 & G(z) \end{bmatrix} - z \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} U \right) + J_2 U \right] \left[J_2 \left(V \begin{bmatrix} 1 & 0 \\ 0 & G(z) \end{bmatrix} - z \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} U \right) + J_1 U \right]^{-1} =$$

$$= \begin{bmatrix} \alpha_V & M_V G \\ k_U & \beta_U^* \end{bmatrix} \begin{bmatrix} \alpha_U & M_U \\ k_V - z\,k_U & \beta_V^* G - z\,\beta_U^* \end{bmatrix}^{-1}.$$

Using the following expression for the inverse of a partitioned matrix

$$
\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} = \begin{bmatrix} 0 & M_3^{-1} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -M_3^{-1} M_4 \\ I_p \end{bmatrix} (M_2 - M_1 M_3^{-1} M_4)^{-1} \begin{bmatrix} I_p & -M_1 M_3^{-1} \end{bmatrix},
$$

in which M_2 and M_3 are square submatrices, and M_3^{-1} is assumed to exist (which holds by assumption in our case), we obtain

$$
\mathcal{F}_{U,V}(G) = \begin{bmatrix} \alpha_V & M_V\, G \end{bmatrix} \begin{bmatrix} -\frac{\beta_V^* \, G - z\, \beta_U^*}{k_V - z\, k_U} \\ I_p \end{bmatrix} \begin{bmatrix} M_U - \frac{\alpha_U \, (\beta_V^* \, G - z\, \beta_U^*)}{k_V - z\, k_U} \end{bmatrix}^{-1} =
$$

$$
= \begin{bmatrix} M_V\, G - \frac{\alpha_V \,(\beta_V^* \, G - z\, \beta_U^*)}{k_V - z\, k_U} \end{bmatrix} \begin{bmatrix} M_U - \frac{\alpha_U \,(\beta_V^* \, G - z\, \beta_U^*)}{k_V - z\, k_U} \end{bmatrix}^{-1} =
$$

$$
= \begin{bmatrix} \left(M_V - \frac{\alpha_V \, \beta_V^*}{k_V - z\, k_U} \right) G + \frac{z\, \alpha_V \, \beta_U^*}{k_V - z\, k_U} \end{bmatrix} \begin{bmatrix} -\frac{\alpha_U \, \beta_V^*}{k_V - z\, k_U}\, G + \left(M_U + \frac{z\, \alpha_U \, \beta_U^*}{k_V - z\, k_U} \right) \end{bmatrix}^{-1}.
$$

And finally, $\mathcal{F}_{U,V}(G) = \mathcal{T}_\Phi(G)$ for Φ given by

$$
\Phi(z) = \begin{bmatrix} M_U + \dfrac{z\alpha_U \beta_U^*}{k_V - zk_U} & -\dfrac{\alpha_U \beta_V^*}{k_V - zk_U} \\ \dfrac{z\alpha_V \beta_U^*}{k_V - zk_U} & M_V - \dfrac{\alpha_V \beta_V^*}{k_V - zk_U} \end{bmatrix}. \tag{66}
$$

Exploiting unitarity of U and V, it then is straightforward to establish that

$$
J - \Phi(z) J \Phi(\lambda)^* = \frac{(1 - \overline{\lambda}\, z)}{(k_V - z\, k_U)(\overline{k}_V - \overline{\lambda}\,\overline{k}_U)} \begin{bmatrix} \alpha_U \\ \alpha_V \end{bmatrix} \begin{bmatrix} \alpha_U \\ \alpha_V \end{bmatrix}^*,
$$

from which it follows that $\Phi(z)$ is J-inner. The expression (66) for $\Phi(z)$ is of McMillan degree 1 having a pole at k_V / k_U (located at infinity if $k_U = 0$), except if $|k_U| = |k_V| = 1$, in which case $\Phi(z)$ is of McMillan degree 0 since $\alpha_U = \alpha_V = 0$ and $\Phi(z) = \begin{bmatrix} M_U & 0 \\ 0 & M_V \end{bmatrix}$ becomes constant.
□

The previous theorem makes clear under which conditions a mapping $\mathcal{F}_{U,V}$ with unitary matrices U and V admits a representation as an LFT associated with an elementary J-inner factor Φ. Depending on the location of the pole k_V / k_U, this factor takes the different forms described in Theorem 2.3. In the cases where such a representation exists the proof is constructive, yielding the possible expression (66) for Φ. Using Lemma 4.3 *all* the possibilities for Φ can be computed.

Lemma 6.2 *Let $\Phi(z)$ and $\Psi(z)$ be elementary J-inner factors for which $\mathcal{T}_\Phi$ and $\mathcal{T}_\Psi$ agree on $\mathcal{L}^p$. If $p > 1$ then there exists a unimodular constant ρ such that*

$$
\Phi(z) = \rho \Psi(z).
$$

If $p = 1$, then either there exists a unimodular constant ρ such that $\Phi(z) = \rho \Psi(z)$, or $\Phi(z)$ has a pole at $z = w$ with $|w| \neq 1$, $\Psi(z)$ has a pole at $z = 1/\overline{w}$ and there exists a unimodular constant ρ such that

$$
\Phi(z) = \rho b_w^\sharp(z) \Psi(z).
$$

Proof. By Lemma 4.3,

$$\Phi(z) = b(z)\Psi(z), \tag{67}$$

for some scalar function $b(z)$, and since Φ and Ψ are J-inner, $b(z)$ must be a Blaschke factor. Computing the determinants gives

$$\det \Phi(z) = b(z)^{2p} \det \Psi(z).$$

But the determinant of an elementary J-inner function is either equal to a unimodular number (if the pole is located on the unit circle) or to a Blaschke factor, having McMillan degree 1. If $p > 1$ it therefore follows that that $b(z)$ must be a unimodular constant. If $p = 1$ an alternative possibility occurs if Φ has a pole at $z = w$ with $|w| \neq 1$. In that case Ψ may have a pole at $z = 1/\overline{w}$ and $b(z)$ may also be chosen equal to $\rho b_w^\sharp(z)$ for some complex number ρ of modulus one. If Ψ is of the form

$$\Psi(z) = \left(I_2 - (1 - b_w(z))x(x^*Jx)^{-1}x^*J\right)H,$$

then

$$\Phi(z) = \rho \left(I_2 - (1 - b_w{}^\sharp(z))y(y^*Jy)^{-1}y^*J\right)H,$$

where y is a solution to $y^*Jx = 0$. $\qquad\square$

As a consequence of Theorem 6.1 and Lemma 6.2 we have the following result.

Corollary 6.3 *Let U and V be unitary matrices of size $(p+1) \times (p+1)$, partitioned as in (65). If $p > 1$, then the mapping $\mathcal{F}_{U,V}$ can be represented in the form of a mapping $\mathcal{T}_{\Theta(u,v,w,\xi,H)}$ if and only if $|k_U| < |k_V|$. Likewise it can be represented in the form of a mapping $\mathcal{T}_{\Theta(u,v,w,\xi,H)^\circ}$ if and only if $|k_U| > |k_V|$.*
If $p = 1$, a representation of $\mathcal{F}_{U,V}$ in the form of a mapping $\mathcal{T}_{\Theta(u,v,w,\xi,H)}$ exists if and only if a representation in the form of a mapping $\mathcal{T}_{\Theta(u,v,w,\xi,H)^\circ}$ exists, which holds if and only if $|k_U| \neq |k_V|$.

Proof. The statements for the case $p > 1$ follow from the earlier results as indicated. For $p = 1$, note that $\mathcal{F}_{U,V} = \mathcal{F}_{\overline{V},\overline{U}}$ which follows from Lemma 3.3 (ii) upon complex conjugation and noting that transposition has no effect on a matrix of size 1×1. $\qquad\square$

To complete the picture it remains to characterize the elementary J-inner factor Φ analytic inside the disk in case $|k_U| < |k_V|$. As we have seen, if a mapping $\mathcal{F}_{U,V}$ is given, the expression (66) together with Lemma 6.2 provide all the possible alternatives for Φ. Conversely, if an elementary J-inner factor Φ analytic inside the disk is given, one may wonder whether $\mathcal{T}_\Phi$ admits a representation as a mapping of the form $\mathcal{F}_{U,V}$. Here it will prove to be essential to introduce the J-inner matrix function $\widehat{\Theta}(u,v,w)(z)$ defined by

$$\widehat{\Theta}(u,v,w)(z) = H(uv^*)S_{u,w}(z)H(\overline{w}uv^*). \tag{68}$$

The main result of this section can now be stated as follows.

Theorem 6.4 *Let $\Theta(z)$ be an elementary J-inner factor analytic inside the unit disk. If $\mathcal{T}_{\Theta(z)}$ coincides with a mapping $\mathcal{F}_{U,V}$ with U and V unitary, then $\Theta(z)$ can be written as*

$$\Theta(z) = \widehat{\Theta}(u,v,w)(z) \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}, \tag{69}$$

for some $u,v \in \mathbb{C}^p$ with $\|u\| = 1$, $\|v\| < 1$, some $w \in \mathbb{C}$ with $|w| < 1$, and some $p \times p$ unitary matrices P and Q.

In that case one can take $U = \widehat{U} \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix}$ and $V = \widehat{V} \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$, where

$$\widehat{U} = \begin{bmatrix} \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}u & I_p - (1 + \frac{w\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}})uu^* \\ \frac{\overline{w}\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}} & \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}u^* \end{bmatrix}, \tag{70}$$

$$\widehat{V} = \begin{bmatrix} \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}v & I_p - (1 - \frac{\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}})\frac{vv^*}{\|v\|^2} \\ \frac{\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}} & -\frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}v^* \end{bmatrix}. \tag{71}$$

Proof. Theorem 6.1 ensures that a mapping $\mathcal{F}_{U,V}$ can be represented by a linear fractional transformation for the J-inner elementary factor $\Phi(z)$ given by (66). If $p > 1$, it follows from Lemma 6.2 that $|k_U| < |k_V|$. If $p = 1$, since the mappings $\mathcal{F}_{U,V}$ and $\mathcal{F}_{\overline{V},\overline{U}}$ coincide, we still may assume $|k_U| < |k_V|$, and in both cases we have that

$$\Phi(z) = \rho\, \Theta(z),$$

for some unimodular number ρ.

Since $|k_U| < |k_V|$, $k_V \neq 0$ and $k_U \neq 1$, so that $\|\alpha_U\| \neq 0$. Let

$$w = \frac{\overline{k_U}}{\overline{k_V}}, \qquad u = \frac{\alpha_U}{\|\alpha_U\|}, \qquad v = \frac{\alpha_V}{\|\alpha_U\|}, \tag{72}$$

which satisfy $|w| < 1$, $\|u\| = 1$ and $\|v\| < 1$. The function $\Theta(z)$ is, by Proposition 5.2, of the form

$$\Theta(z) = H(uv^*)S_{u,w}(z)H(E) \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix},$$

for some strictly contractive matrix E and some unitary matrices P and Q, depending on the matrices U and V. To see this, note that if $\Theta(z)$ is an elementary J-inner factor written in the form $\Theta(u,v,w,\xi,H)(z)$ having its only zero at $z = w$, then the matrix $\Theta(w)$ has a nontrivial left kernel of dimension one, which is characterized by Eqn. (52). For the matrix function $\Phi(z)$ given by (66) it is easily shown that $\begin{bmatrix} u^* & -v^* \end{bmatrix}\Phi(w) = 0$ for w, u and v defined by Eqn. (72). Note that, the matrices U and V being unitary,

$$\|\alpha_U\|^2 + |w|^2|k_V|^2 = 1$$
$$\|\alpha_U\|^2\|v\|^2 + |k_V|^2 = 1,$$

so that

$$\|\alpha_U\| = \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}, \qquad |k_V| = \frac{\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}}. \tag{73}$$

Now, the quickest way to proceed is to compare the realizations of $\Phi^\sharp$ and $\Theta^\sharp$ (note that Θ will fail to have a realization if $w=0$). Using (66) we get

$$\Phi^\sharp(z) = \begin{bmatrix} M_U^* & 0 \\ -\frac{\beta_V \alpha_U^*}{k_V} & M_V^* - \frac{\beta_V \alpha_V^*}{k_V} \end{bmatrix} + \begin{bmatrix} \beta_U \\ -\frac{k_U}{k_V}\beta_V \end{bmatrix} (z - \frac{\overline{k_U}}{\overline{k_V}})^{-1} \begin{bmatrix} \frac{\alpha_U^*}{k_V} & \frac{\alpha_V^*}{k_V} \end{bmatrix},$$

while a state-space realization of $S_{u,w}^\sharp(z)$ is easily computed to yield that

$$\Theta^\sharp(z) = \mathcal{D} + \mathcal{C}(z-w)^{-1}\mathcal{B},$$

where

$$\mathcal{D} = \begin{bmatrix} P^* & 0 \\ 0 & Q^* \end{bmatrix} H(E) \begin{bmatrix} I_p - (1+\overline{w})uu^* & 0 \\ 0 & I_p \end{bmatrix} H(uv^*),$$

$$\mathcal{C} = \begin{bmatrix} P^* & 0 \\ 0 & Q^* \end{bmatrix} H(E) \begin{bmatrix} \sqrt{1-|w|^2}u \\ 0 \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} \sqrt{1-|w|^2}u^* & 0 \end{bmatrix} H(uv^*).$$

Comparing the (1,2)-block-entries of the expressions for the direct feedthrough yields

$$(I_p - EE^*)^{-1/2}(E - \overline{w}uv^*)(I_p - vv^*)^{-1/2} = 0$$

and thus $E = \overline{w}uv^*$ or equivalently

$$\Theta(z) = \widehat{\Theta}(u,v,w)(z) \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix},$$

as claimed. Replacing E by its value gives

$$\mathcal{D} = \begin{bmatrix} P^* & 0 \\ 0 & Q^* \end{bmatrix} \begin{bmatrix} I_p - \left(1 + \frac{\overline{w}\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}}\right)uu^* & 0 \\ \frac{1-|w|^2}{\sqrt{1-\|v\|^2}\sqrt{1-|w|^2\|v\|^2}}vu^* & I_p - \left(1 - \frac{\sqrt{1-|w|^2\|v\|^2}}{\sqrt{1-\|v\|^2}}\right)\frac{vv^*}{\|v\|^2} \end{bmatrix},$$

$$\mathcal{C} = \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}} \begin{bmatrix} P^* & 0 \\ 0 & Q^* \end{bmatrix} \begin{bmatrix} u \\ wv \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \frac{\sqrt{1-|w|^2}}{\sqrt{1-\|v\|^2}} \begin{bmatrix} u^* & v^* \end{bmatrix}.$$

The matrix M_U is therefore given by

$$M_U = \rho \left(I_p - \left(1 + \frac{w\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}}\right)uu^*\right) P.$$

From (72) and (73) we have that

$$\begin{bmatrix} \frac{\alpha_U^*}{k_V} & \frac{\alpha_V^*}{k_V} \end{bmatrix} = \frac{|k_V|}{\overline{k_V}}\mathcal{B},$$

which determines the transformation between the two realizations, so that

$$
\begin{bmatrix} \beta_U^* & -\frac{k_U}{k_V}\beta_V^* \end{bmatrix} = \rho\, \frac{\overline{k_V}}{|k_V|}\, \mathcal{C}^* = \rho\, \frac{\overline{k_V}}{|k_V|}\, \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}\, \begin{bmatrix} u^* & \overline{w}\, v^* \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}.
$$

Finally it is obtained that the matrices U and V are of the form

$$
U = \begin{bmatrix} I_p & 0 \\ 0 & \frac{\overline{k_V}}{|k_V|} \end{bmatrix} \widehat{U} \begin{bmatrix} 1 & 0 \\ 0 & \rho P \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} I_p & 0 \\ 0 & \frac{\overline{k_V}}{|k_V|} \end{bmatrix} \widehat{V} \begin{bmatrix} 1 & 0 \\ 0 & \rho Q \end{bmatrix},
$$

with $\hat{U}$ and $\hat{V}$ specified by (70) and (71). Since a pre-multiplication by $\begin{bmatrix} I_p & 0 \\ 0 & \frac{\overline{k_V}}{|k_V|} \end{bmatrix}$ and a post-multiplication by $\begin{bmatrix} 1 & 0 \\ 0 & \rho I_p \end{bmatrix}$ do not change the mapping $\mathcal{F}_{U,V}$ according to Lemma 3.3, the values of U and V given in the theorem are obtained. $\square$

7   Overlapping canonical forms for lossless and stable systems

In the tangential Schur algorithm as presented in Theorem 5.5, a lossless function $G = G^{(n)}$ of McMillan degree n is broken down to a constant unitary matrix $G^{(0)}$ using a sequence of inter-polation points w_k, direction vectors u_k, Schur vectors v_k, unimodular constants ξ_k and constant $J-$unitary matrices $H_k(u_k, v_k, w_k, \xi_k)$, for $k = 1, \ldots, n$. The set of values for w_k, u_k, ξ_k and the set of mappings H_k at the n recursion steps can serve to index a generic chart for the differen-tiable manifold of lossless functions of degree n, if the mappings H_k are sufficiently smooth. The Schur parameter vectors v_k together with the unitary matrix $G^{(0)}$ that we finally reach, then provide the local coordinates for this chart. An infinite atlas of overlapping generic charts is obtained by varying the choices for the u_k, w_k, ξ_k and H_k. Sub-atlases can be extracted from it, as the one built in [1] which employs elementary J-inner matrices of the form $\Theta(u_k, v_k, w_k, 1, I_p)$. Such a choice implies that $G^{(0)}$ is the value of the lossless function G at 1. We may even take a finite atlas by fixing all the interpolation points at zero and letting the values of the u_k vary among a canonical basis of $\mathbb{C}^p$ (cf. the proof of Theorem 5.5). However it is interesting to keep in mind the richness of the possibilities, including the one to mix interpolation conditions on the columns (51) with interpolation conditions on the rows (64). This could be of interest in order to describe lossless transfer functions with a particular structure.

Now let constants ξ_k be chosen with $|\xi_k| = 1$ for all $k = 1, \ldots, n$, and let the mappings

$$
H_k(u, v, w, \xi) = H(uv^*)S_{u,w}(\xi)H(\bar{w}uv^*)
$$

be chosen so that $\Theta(u, v, w, \xi_k, H_k(u, v, w, \xi_k)) = \widehat{\Theta}(u, v, w)$ for all $u \in \mathbb{C}^p$, $\|u\| = 1$, $v \in \mathbb{C}^p$, $\|v\| < 1$, $w \in \mathbb{C}$, $|w| < 1$. Then one obtains an atlas for the manifold of lossless systems that can be described directly in terms of state-space realizations using the $\mathcal{F}_{U,V}$ mappings.

Let us now describe such an atlas for the manifold $\mathcal{L}_n^p$ of $p \times p$ lossless functions of McMillan degree n in more detail. Choose $\xi_k = 1$ for all $k = 1, \ldots, n$. For a set of interpolation points

$w_1, w_2, \ldots, w_n$ in the open unit disk, a set of direction vectors $u_1, u_2, \ldots, u_n \in \mathbb{C}^p$ of length one, and a coordinate chart $(\mathcal{W}, \vartheta)$ for the set $\mathcal{U}_p$ of unitary $p \times p$ matrices, we define a chart $(\mathcal{V}, \varphi)$ by its domain

$$\mathcal{V}(w_1, w_2, \ldots, w_n, u_1, u_2, \ldots, u_n, \mathcal{W}) = \{G \in \mathcal{L}_n^p \,|\, \|G^{(k)}(1/\overline{w}_k)u_k\| < 1, G^{(0)} \in \mathcal{W}\},$$

and its coordinate map

$$\varphi: \quad G \quad \rightarrow \quad (v_1, v_2, \ldots, v_n, \vartheta(G^{(0)}))$$

where, for $k = 1, \ldots, n$, the lossless functions $G^{(k)}$ and the Schur parameter vectors v_k are recursively defined in Theorem 5.5, in which we choose

$$\Theta_k = \widehat{\Theta}(u_k, v_k, w_k).$$

Theorem 7.1 *The family $(\mathcal{V}, \varphi)$ defines a C^∞-atlas on $\mathcal{L}_n^p$.*

The proof is left to the reader; see also [1] and [10].

In approximation problems, the Douglas-Shapiro-Shields factorization is often used to represent a transfer function. Since in such problems the lossless factor is only determined up to a unitary constant matrix, we are interested in a parametrization of the quotient space of lossless functions by unitary matrices. To be precise, two lossless functions $G_1(z)$ and $G_2(z)$ in this quotient are equivalent if there exists a unitary matrix X such that $G_1(z) = X G_2(z)$. To deal with such an equivalence relation, we have the identity

$$\mathcal{T}_{\widehat{\Theta}(w, u, X\,v)}(X\,G) = X\,\mathcal{T}_{\widehat{\Theta}(w, u, v)}(G),$$

which is satisfied for every lossless G and every unitary X. Therefore, if $G(z)$ has a set of Schur parameters $(v_1, \ldots, v_n)$ and a unitary matrix $G^{(0)} = D_0$ in a chart defined by given sequences $(w_1, \ldots, w_n)$ of interpolation points and $(u_1, \ldots, u_n)$ of direction vectors, then $X\,G(z)$ has the Schur parameters $(X\,v_1, \ldots, X\,v_n)$ and the unitary matrix $G^{(0)} = X\,D_0$ in the same chart. An atlas of the quotient space can therefore be obtained by fixing the unitary matrix $G^{(0)}$ in each chart instead of letting it vary. For $w_1, w_2, \ldots, w_n$ in the open unit disk, $u_1, u_2, \ldots, u_n \in \mathbb{C}^p$ of unit length *and* D_0, a fixed $p \times p$ unitary matrix, we define the chart $(\widetilde{\mathcal{V}}, \widetilde{\varphi})$ by its domain

$$\widetilde{\mathcal{V}}(w_1, w_2, \ldots, w_n, u_1, u_2, \ldots, u_n, D_0) = \{G \in \mathcal{L}_n^p \,|\, \|G^{(k)}(1/\overline{w}_k)u_k\| < 1, G^{(0)} = D_0\},$$

and its coordinate map

$$\widetilde{\varphi}: \quad G \quad \rightarrow \quad (v_1, v_2, \ldots, v_n)$$

where the lossless functions $G^{(k)}$ and the Schur parameter vectors v_k are again those of Theorem 5.5, in which $\Theta_k = \widehat{\Theta}(u_k, v_k, w_k)$.

Theorem 7.2 *The family $(\widetilde{\mathcal{V}}, \widetilde{\varphi})$ defines a C^∞-atlas on the quotient space $\mathcal{L}_n^p/\mathcal{U}_p$.*

Moreover, a unique balanced realization computed from the Schur parameters corresponds to every lossless function in the domain $\widetilde{\mathcal{V}}$ of a chart, by iterating formula (34) in which U and V are the unitary matrices given in Theorem 6.4. These 'Schur balanced realizations' are really in a (local) canonical form. With each (A, B, C, D) this method associates (for a given choice of interpolation points and direction vectors) a unique realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$. Attaching a canonical form with each chart, we obtain a set of overlapping canonical forms. This generalizes the results of [14] to the multivariable case and opens up possibilities for multivariable stable all-pass model reduction and approximation methods.

Note that, except for particular choices of the interpolation points and direction vectors, these overlapping canonical forms in general cannot be characterized in a simple way, and cannot be computed directly from a given realization. However, if the interpolation points w_k are the poles of $G(z)$ with the direction vectors u_k spanning the associated kernels of the singular matrices $G(1/\overline{w}_k)$, then the Schur vectors v_k are all zero and the Schur balanced realization is in Schur form (i.e., the matrix A is triangular). If the interpolation points w_k are all equal to zero and the vectors u_k are all equal to $e_1 = [1, 0, \ldots, 0]^*$, then the Schur balanced realization is positive upper Hessenberg.

The overlapping balanced canonical forms for stable all-pass systems obtained in the way described above, also give rise to output-normal canonical forms, resp. input-normal canonical forms, as follows. Consider an arbitrary (asymptotically) stable system in minimal state-space form, given by (A, B, C, D). It is not hard to show that one can find a pair $(\widetilde{B}, \widetilde{D})$ such that $(A, \widetilde{B}, C, \widetilde{D})$ is stable all-pass. One can then choose a state-space transformation that brings this system into one of the balanced canonical forms presented in this paper, which we denote by (A_b, B_b, C_b, D_b). The resulting state-space transformation T and the resulting pair $(A_b, C_b) = (TAT^{-1}, CT^{-1})$ are independent of the choice of $\widetilde{B}, \widetilde{D}$.

Therefore applying the same state-space transformation to (A, B, C, D) one obtains a canonical form (A_n, B_n, C_n, D_n), where $A_n = A_b$, $C_n = C_b$, $B_n = TB$, $D_n = D$. Obviously the observability Gramian of this canonical form is the identity matrix, hence it is an output-normal canonical form. By duality one can also construct input-normal canonical forms in the same fashion. These output-normal and input-normal canonical forms are of importance in model order reduction and system identification problems. An analogous canonical form in continuous time was used in [13] to construct an algorithm to find the stable system of order k, $k < n$, which is closest in the H_2-norm to a given stable system of order n. In system identification output-normal canonical forms are used in the context of the so-called separable least squares approach (cf., e.g., [6]).

The parametrizations described in this paper have been implemented in a software program named RARL2, dedicated to the rational approximation of *multivariable* discrete-time transfer functions in L^2 norm. Just as the software program Hyperion, it is based on the algorithm described in [10]. However, in contrast to Hyperion it has the particularity to deal with state-space formulations. This state-space approach is of great interest due to the good numerical behavior of the recursive construction of balanced realizations presented in this paper. Presently the main domain of application of RARL2 is the synthesis and identification of filters from partial frequency data.

A   Realizations and linear fractional transformations

In this appendix, we give the realization formulas associated with the linear fractional transformations studied in Section 4.

Proposition A.1 *Let $S(z)$ be a $p \times p$ proper rational function with state space realization (A, B, C, D) of degree n and let $\Phi(z)$ be a $2p \times 2p$ proper rational function with m-dimensional state space realization*

$$\Phi(z) = \left[\begin{array}{cc} \mathcal{D}_1 & \mathcal{D}_2 \\ \mathcal{D}_3 & \mathcal{D}_4 \end{array} \right] + \left[\begin{array}{c} \mathcal{C}_1 \\ \mathcal{C}_2 \end{array} \right] (zI_m - \mathcal{A})^{-1} \left[\begin{array}{cc} \mathcal{B}_1 & \mathcal{B}_2 \end{array} \right].$$

(i) *Then $\widetilde{S}(z) = \mathcal{T}_{\Phi(z)}(S(z))$ is well-defined if $\mathcal{D}_2 D + \mathcal{D}_1$ is invertible and a state-space realization $(\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D})$ of $\widetilde{S}(z)$ with $(n + m)$-dimensional state-space is given by:*

$$\widetilde{A} = \left[\begin{array}{cc} \mathcal{A} - (\mathcal{B}_2 D + \mathcal{B}_1)(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \mathcal{C}_1 & \left[\mathcal{B}_2 - (\mathcal{B}_2 D + \mathcal{B}_1)(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \mathcal{D}_2 \right] C \\ -B(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \mathcal{C}_1 & A - B(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \mathcal{D}_2 C \end{array} \right],$$

$$\widetilde{B} = \left[\begin{array}{c} (\mathcal{B}_2 D + \mathcal{B}_1)(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \\ B(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \end{array} \right],$$

$$\widetilde{C} = \left[\begin{array}{cc} \mathcal{C}_2 - (\mathcal{D}_4 D + \mathcal{D}_3)(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \mathcal{C}_1 & (\mathcal{D}_4 - (\mathcal{D}_4 D + \mathcal{D}_3)(\mathcal{D}_2 D + \mathcal{D}_1)^{-1} \mathcal{D}_2) C \end{array} \right],$$

$$\widetilde{D} = (\mathcal{D}_4 D + \mathcal{D}_3)(\mathcal{D}_2 D + \mathcal{D}_1)^{-1}.$$

(74)

(ii) *Then $\widehat{S}(z) = \widehat{\mathcal{T}}_{\Phi(z)}(S(z))$ is well-defined if $D \mathcal{D}_2 + \mathcal{D}_4$ is invertible and a state-space realization $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ of $\widehat{S}(z)$ with $(n + m)$-dimensional state-space is given by:*

$$\widehat{A} = \left[\begin{array}{cc} \mathcal{A} - \mathcal{B}_2 (D \mathcal{D}_2 + \mathcal{D}_4)^{-1} (D \mathcal{C}_1 + \mathcal{C}_2) & -\mathcal{B}_2 (D \mathcal{D}_2 + \mathcal{D}_4)^{-1} C \\ B \mathcal{C}_1 - B \mathcal{D}_2 (D \mathcal{D}_2 + \mathcal{D}_4)^{-1} (D \mathcal{C}_1 + \mathcal{C}_2) & A - B \mathcal{D}_2 (D \mathcal{D}_2 + \mathcal{D}_4)^{-1} C \end{array} \right],$$

$$\widehat{B} = \left[\begin{array}{c} \mathcal{B}_1 - \mathcal{B}_2 (D \mathcal{D}_2 + \mathcal{D}_4)^{-1}(D \mathcal{D}_1 + \mathcal{D}_3) \\ B \left(\mathcal{D}_1 - \mathcal{D}_2 (D \mathcal{D}_2 + \mathcal{D}_4)^{-1}(D \mathcal{D}_1 + \mathcal{D}_3) \right) \end{array} \right],$$

(75)

$$\widehat{C} = \left[\begin{array}{cc} (D \mathcal{D}_2 + \mathcal{D}_4)^{-1} (D \mathcal{C}_1 + \mathcal{C}_2) & (D \mathcal{D}_2 + \mathcal{D}_4)^{-1} C \end{array} \right],$$

$$\widehat{D} = (D \mathcal{D}_2 + \mathcal{D}_4)^{-1}(D \mathcal{D}_1 + \mathcal{D}_3).$$

Proof. We prove (ii), the proof of (i) is similar. Let Φ be block-partitioned analogously to the block-partitioning of Φ in Section 4. We have (cf. Eqn. 40):

$$\widehat{S}(z) = (S(z)\Phi_2(z) + \Phi_4(z))^{-1}(S(z)\Phi_1(z) + \Phi_3(z)).$$

Then, $(S\Phi_1 + \Phi_3)(z)$ and $(S\Phi_2 + \Phi_4)(z)$ have realizations

$$\begin{array}{rcl} (S\Phi_1 + \Phi_3)(z) & = & H(zI_{n+m} - F)^{-1}G_1 + K_1, \\ (S\Phi_2 + \Phi_4)(z) & = & H(zI_{n+m} - F)^{-1}G_2 + K_2, \end{array}$$

where

$$F = \begin{bmatrix} \mathcal{A} & 0 \\ B\,\mathcal{C}_1 & A \end{bmatrix}, \qquad G_1 = \begin{bmatrix} \mathcal{B}_1 \\ B\,\mathcal{D}_1 \end{bmatrix}, \quad K_1 = D\,\mathcal{D}_1 + \mathcal{D}_3,$$

$$H = \begin{bmatrix} D\,\mathcal{C}_1 + \mathcal{C}_2 & C \end{bmatrix}, \quad G_2 = \begin{bmatrix} \mathcal{B}_2 \\ B\,\mathcal{D}_2 \end{bmatrix}, \quad K_2 = D\,\mathcal{D}_2 + \mathcal{D}_4.$$

The inverse $(G\Phi_2 + \Phi_4)(z)^{-1}$ then has realization

$$(G\Phi_2 + \Phi_4)^{-1}(z) = K_2^{-1} - K_2^{-1}H\,(zI_{n+m} - (F - G_2K_2^{-1}H))^{-1}G_2K_2^{-1}.$$

Finally, $\widehat{G}(z)$ has the realization $\widehat{G}(z) = C'\,(zI_{2(n+m)} - A')^{-1}B' + D'$, given by:

$$A' = \begin{bmatrix} F & 0 \\ -G_2K_2^{-1}H & F - G_2K_2^{-1}H \end{bmatrix}$$

$$B' = \begin{bmatrix} G_1 \\ -G_2K_2^{-1}K_1 \end{bmatrix}$$

$$C' = \begin{bmatrix} K_2^{-1}H & K_2^{-1}H \end{bmatrix}$$

$$D' = K_2^{-1}K_1.$$

However, this realization with a state space of dimension $2(n+m)$ is not minimal. It can be reduced to a realization with a state space of dimension $n + m$ by applying the state space basis transformation matrix T defined by $T = \begin{bmatrix} I_{n+m} & 0 \\ I_{n+m} & I_{n+m} \end{bmatrix}$ and truncating the unobservable $n+m$ state components. This yields the realization (75).   □

References

[1] D. Alpay, L. Baratchart and A. Gombani, On the differential structure of matrix-valued rational inner functions, *Operator Theory: Adv. and Appl.*, **73**, 30–66, 1994.

[2] M. Bakonyi and T. Constantinescu, *Schur's algorithm and several applications*, Longman Sci. Tech., Harlow, 1992.

[3] J.A. Ball, I. Gohberg and L. Rodman, Interpolation of rational matrix functions, *Operator Theory: Advances and Applications*, **45**, Birkhäuser, 1990.

[4] L. Baratchart, On the topological structure of inner functions and its use in identification, in: *Analysis of Controlled Dynamical Systems, Lyon, France, 1990, Progress in Systems and Control Theory*, Vol. **8**, Birkhäuser, 51–59, 1990.

[5] L. Baratchart and M. Olivi, Inner-unstable factorization of stable rational transfer functions in: *Progress in System and Control Theory: Modeling, Estimation and Control of Systems with Uncertainty*, Vol. **10**, Birkhäuser, 22–39, 1991.

[6] J. Bruls, C.T. Chou, B.R.J. Haverkamp and M. Verhaegen, *Linear and non-linear system identification using separable least-squares*, Technical University Delft, Dept. Electrical Engineering, Systems- and Control Laboratory, Report TUD/ET/SCE96.009, 1997.

[7] R. Douglas, H. Shapiro and A. Shields, Cyclic vectors and invariant subspaces for the backward shift operator, *Annales de l'Institut Fourier, (Grenoble)*, **20**, 37–76, 1970.

[8] H. Dym, *J*-contractive matrix functions, reproducing kernel spaces and interpolation, *CBMS lecture notes*, **71**, American Mathematical Society, Rhode Island, 1989.

[9] H. Dym, On reproducing kernel spaces, *J*-unitary matrix functions, interpolation and displacement rank, *Operator Theory : Advances and Applications*, Vol. **41**, 173–239, 1989.

[10] P. Fulcheri and M. Olivi, Matrix Rational H^2-Approximation: A Gradient Algorithm Based On Schur Analysis, *SIAM Journal on Control and Optimisation*, Vol. **36**, No. 6, 2103–2127, 1998.

[11] J. Garnett, *Bounded Analytic Functions*, Academic Press, 1981.

[12] Y. Genin, P. Van Dooren, T. Kailath, J.-M. Delosme and M. Morf, On Σ-lossless transfer functions and related questions, *Linear Algebra and its Applications*, **50**, 251–275, 1983.

[13] B. Hanzon and J.M. Maciejowski, Constructive algebra methods for the L_2 problem for stable linear systems, *Automatica*, Vol. **32**, No. 12, 1645–1657, 1996.

[14] B. Hanzon and R.L.M. Peeters, Balanced Parametrizations of Stable SISO All-Pass Systems in Discrete-Time, *MCSS*, **13**, 240–276, 2000.

[15] B. Hanzon and R.J. Ober, Overlapping block-balanced canonical forms for various classes of linear systems, *Linear Algebra and its Applications*, **281**, 171–225, 1998.

[16] T. Kailath, *Linear Systems*, Prentice-Hall, 1980.

[17] J. Leblond and M. Olivi, Weighted H^2 Approximation of Transfer Functions, *MCSS*, **11**, 28–39, 1998.

[18] B.C. Moore, Principal component analysis in linear systems: Controllability, observability, and model reduction, *IEEE Trans. Autom. Control* **AC-26**, 17–32, 1981.

[19] R.J. Ober, Balanced realizations: canonical form, parametrization, model reduction, *Int. J. Control*, **46**, 643–670, 1987.

[20] V.P. Potapov, The multiplicative structure of *J*-contractive matrix functions, *Amer. Math. Soc. Transl. (2)*, **15**, 131–243, 1960.

[21] V.P. Potapov, Linear Fractional Transformations of Matrices, *Amer. Math. Soc. Transl. (2)*, **138**, 21–35, 1988.

[22] L. Pernebo, L.M. Silverman, Model reduction via balanced state space representations. *IEEE Trans. Autom. Control*, **AC-27**, 382–387, 1982.

[23] N.J. Young, Linear fractional transformations in rings and modules, *Linear Algebra Appl.*, **56**, 251–290, 1984.