A New Parametrization of Rational Inner Functions of Fixed Degree: Schur Parameters and Realizations.

Andrea Gombani* and Martine Olivi†

Abstract. We present a new parametrization of inner functions based on the Schur algorithm. We make use of state space formulas (in practice we obtain a new parametrization of observable pairs). The main advantage of our parametrization is that for each chart the observability gramian is constant: this leads to a very good behavior in some approximation problems.

Key words. Rational inner functions, Parametrizations, Schur algorithm, Realizations, Continuous time systems, Smooth manifolds.

1 Introduction

In recent years quite a lot of attention has been devoted to the problem of parametrizing efficiently $p \times p$ matrix valued inner functions of a given degree n. The problem has a relatively long history: in fact it is intrinsically related to the problem of parametrizing linear systems of degree n whose first solution was provided by Hazewinkel and Kalman [HK]. Nevertheless, the atlas presented in this work (nice selection) is not very practical from a computational point of view, since the derivatives, the domain of the charts and the change of charts are quite difficult to compute and not very well behaved numerically. More recently, the research of better representations has been fostered by the use of matrix inner functions in approximation problems in L^2 norm (see [BO]) and in some new orthonormal basis used in signal processing (see, e.g., [HBV]). Among the papers on the subject we would like to recall the work of Ober (see [O]), who in a seminal paper on balanced realizations derives a canonical form for inner functions: this form, although it has a nice behavior in many respects (in particular it has good numerical properties), is not an atlas. More recently Hanzon and Ober ([HO1] and [HO2]) have obtained a canonical form based on balanced realizations which is an atlas. Still, Kronecker indexes are needed, and the realization has a rather complicated form. Fuhrmann and Helmke (see [FH]) have provided another parametrization by means of geometric control tools. Also in [ABG] an atlas is provided and it has been successfully used for actual computations in some model reduction problems (see [FO]). Nevertheless, also here, although the domain of the chart is described very simply, the actual computations of derivatives are quite involved.

*LADSEB-CNR, Corso Stati Uniti 4, 35127 Padova, ITALY, e-mail:gombani@ladseb.pd.cnr.it

†INRIA, BP 93, 06902 Sophia–Antipolis Cedex, FRANCE, e-mail:olivi@sophia.inria.fr

The atlas we present here has a very nice behavior with respect to differentiation and change of charts. Also the problem of the domain of the charts, although not as simple as in [ABG], can be handled relatively easily.

The main idea is to use an implicit version of the Schur algorithm to parametrize our functions. It will be seen that the Schur algorithm yields in a rather canonical way a realization which has a diagonal observability gramian. The Schur algorithm is generally used to solve a recursive interpolation problem: in particular, n interpolation points $\omega_1, \ldots, \omega_n$ together with some interpolating conditions are given; the algorithm consists in constructing recursively a sequence of functions $Q^{(i)}$ for $i = 1, \ldots, n$ which satisfy (among other conditions):

$$u_i^* Q^{(i)}(\omega_i) = v_i^* \tag{1}$$

where (u_i, v_i) for $i = 1, \ldots, n$ are suitably given (this process will be described in detail in section 4). The coordinates in the chart are then the vectors v_i, $i = 1, \ldots, n$. In practice, with this choice of charts the observability gramian will change together with the v_i. This leads to an unsatisfactory behavior when we differentiate.

To avoid this problem, we do assume that for each chart the observability gramian is constant and then show that there exist, under the proper assumptions, unique points $\{\omega_i; i = 1, ..., n\}$ such that (1) is satisfied for $i = 1, \ldots, n$. It should be pointed out that we do not actually need to compute the value of ω_i to construct our inner functions, so that all computations become quite straightforward. The price to pay for this nice behavior is that we need, in theory, an infinite number of charts as we approach the boundary of the manifold (see the example in section 6.2). Nevertheless, this problem does not arise when we use this atlas for model reduction in L^2 norm: under some extra assumptions that hold generically, the critical points of the L^2 criterion are always in the interior of the manifold and finite in number (see [FO] for more details in the set-up of the disk). They can therefore be reached with a finite number of changes of chart.

We would like to remark that we are actually parametrizing the inverses of inner functions (so the Schur algorithm is turned around). This is because by using the standard Schur algorithm in our context (i.e. for the parametrization of inner functions) it becomes natural, for the construction of the charts, to consider also interpolation points in a vertical strip of the left-half plane; this procedure clearly introduces a discontinuity along the imaginary axis. Although this approach is feasible, it leads to unnecessary complications which can be avoided by the present choice of parameters.

The paper is structured as follows: in section 2 we give some preliminaries and in section 3 we present some background material about inner and $J-$inner functions. In section 4 a review of the Schur algorithm is presented and realization formulas in terms of Schur parameters are provided. In section 5 we give our new atlas. Section 6 is devoted to examples and section 7 to an application to model reduction.

2 Preliminaries and notations

We shall denote by Π^+ the open right half-plane and by $\mathcal{H}_2$ the corresponding Hardy space of vector or matrix valued functions (the proper dimension will be understood from the context). The space $\mathcal{H}_2$ is naturally endowed with the scalar product,

$$< F, G >= \frac{1}{2\pi} \mathrm{Tr} \int_{-\infty}^{\infty} F(iy)G(iy)^* \, dy, \tag{2}$$

and we shall denote by $\| \, \|_2$ the associated norm. Note that if M is a complex matrix, Tr stands for its trace, M^T for its transposed and M^* for its conjugate transposed.

The prefix $\mathcal{R}$ in front of the name of some set of vector or matrix functions will indicate that we consider the *real* subspace of functions F whose Fourier coefficients are real, or equivalently which satisfy the relation $F(\bar{s}) = \overline{F(s)}$. Such functions are relevant in most applications. However, the natural framework for our study is the complex case which plainly includes the real case by restriction. When necessary, the results will be stated for real transfer functions.

We say that a $p \times p$ rational matrix function Q analytic in Π^+ is inner if

$$Q(s)Q(s)^* = I_p, \qquad s \in i\mathbb{R},$$

where $i\mathbb{R}$ denotes the imaginary axis. As usual, the space of $\mathbb{C}^p$-valued functions, $H(Q)$, is defined by

$$H(Q) := \mathcal{H}_2 \ominus Q\mathcal{H}_2. \tag{3}$$

We say (following Dym [D1]) that a Hilbert space H of $\mathbb{C}^p$-valued functions analytic in an open domain Ω is a Reproducing Kernel Hilbert Space (RKHS) if there exists a $\mathbb{C}^{p\times p}$-valued function $K(s,\omega)$ such that $K(\cdot,\omega)\xi \in H$ for $\xi \in \mathbb{C}^p$ and $\omega \in \Omega$ and, for any f in H, we have that $\langle f, K(\cdot,\omega)\xi \rangle_H = \xi^* f(\omega)$. It is easily checked that K satisfies

$$\sum_{i,j=1}^{r} \xi_i^* K(s_i, s_j)\xi_j \geq 0, \tag{4}$$

for every choice of points $s_1, \dots, s_r \in \Omega$ and vectors $\xi_1, \dots, \xi_r \in \mathbb{C}^p$, and that for a given space the reproducing kernel is unique. An important result of Aronszajn (see e.g. [D1, th.2.1]) ensures that, if $K(s,\omega)$ is a function satisfying condition (4), then there exists a Hilbert space having K as its reproducing kernel. This space is defined as

$$H = \overline{span}\{K(s,\omega)\xi; \omega \in \Omega, \xi \in \mathbb{C}^p\}$$

where the symbol $\overline{span}$ is the closure, of the span, in the induced inner product.

Let $J_{p,q} := \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$ for p, q nonnegative integers. We say that a $(p+q) \times (p+q)$ rational matrix Θ is $J_{p,q}$-unitary if

$$\Theta(s) J_{p,q} \Theta(s)^* = J_{p,q}, \qquad s \in i\mathbb{R}. \tag{5}$$

A $J_{p,q}-$unitary function is $J_{p,q}-$inner if

$$\Theta(s)J_{p,q}\Theta(s)^* \leq J_{p,q}, \qquad \operatorname{Re} s \geq 0 \;\; a.e. \tag{6}$$

For a $J_{p,q}-$inner function Θ we set $H(\Theta)$ to be the RKHS with kernel

$$K(s,\omega) := \frac{J_{p,q} - \Theta(s)J_{p,q}\Theta(\omega)^*}{s+\overline{\omega}} \tag{7}$$

(it is easily verified that (4) holds [D1, ch.2]). Then, $H(\Theta)$ is a subspace of $\mathcal{H}_2$ (of $\mathbb{C}^p$-valued functions) endowed with the J inner product [D1, th.2.8.]

$$\langle f, g\rangle_{H(\Theta)} = \langle f, J_{p,q}\, g\rangle. \tag{8}$$

Observe that in the case $J_{p,q} = I_p$, Θ is actually inner (see [D1, ch.1]) and the present definition of the Hilbert space $H(\Theta)$ is consistent with the one given by (3).

3 State space formulas for the Schur linear fractional transformation.

We introduce here some material about $J_{p,q}$-inner functions. The proofs can be found in [BGR] and [D1]. If $F(s) = C(sI_n - A)^{-1}B + D$ is a realization of some proper rational function F, we shall write

$$F = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right).$$

Lemma 1 *Let Θ be a $(p+q)\times(p+q)$ rational function analytic at ∞, of McMillan degree n.*

1. *Θ is a $J_{p,q}-$unitary function if and only if for any minimal realization*

$$\Theta = \left(\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array}\right), \tag{9}$$

we have $\mathcal{D}J_{p,q}\mathcal{D}^* = J_{p,q}$ and there exists a unique invertible solution $\mathcal{P}$ to the Lyapunov equation

$$\mathcal{A}^*\mathcal{P} + \mathcal{P}\mathcal{A} + \mathcal{C}^*J_{p,q}\mathcal{C} = 0 \tag{10}$$

for which

$$\mathcal{B} = -\mathcal{P}^{-1}\mathcal{C}^*J_{p,q}\mathcal{D}. \tag{11}$$

2. *The function Θ is $J_{p,q}$–inner if and only if $\mathcal{P}$ in* (10) *is positive definite. In this case, the columns of*

$$\mathcal{C}(sI_n - \mathcal{A})^{-1} \tag{12}$$

form a basis for the space $H(\Theta)$ and its reproducing kernel can be written as

$$K(s,\omega) = \mathcal{C}(sI_n - \mathcal{A})^{-1}\mathcal{P}^{-1}(\bar{\omega}I_n - \mathcal{A}^*)^{-1}\mathcal{C}^*. \tag{13}$$

Moreover, $\mathcal{P}$ is the Gram matrix associated to this basis with respect to the $J_{p,q}$-inner product (8).

These facts are well known and can be found in e.g. [BGR, th.6.1.1, th.6.2.2.] and [GVKDM].

Corollary 1 *The $\mathbb{C}^p$-valued function Q is inner if and only if for any minimal realization*

$$Q = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right),$$

we have $DD^* = D^*D = I_p$ and there exists a unique positive definite solution P to

$$A^*P + PA + C^*C = 0,$$

for which $B = -P^{-1}C^*D$.

Proof: it follows from Lemma 1 if we take $J_{p,q} = I_p$. □

In this case, P is the observability Gramian associated to the realization.

In the sequel, we will assume that $J = \begin{pmatrix} I_p & 0 \\ 0 & -I_p \end{pmatrix}$ and $\Theta(\infty) = I_{2p}$, so that any J-inner matrix is uniquely determined by the pair $\mathcal{A}, \mathcal{C}$. Now, let us write $\mathcal{C} = \begin{pmatrix} U \\ V \end{pmatrix}$, where U and V have dimension $p \times n$ and let

$$\Theta \;=\; \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix} \tag{14}$$

be the block partition of Θ with each block of size $p \times p$; let S be a $(p \times p)$-rational inner function. We define the linear fractional transformation $T_\Theta(S)$ as

$$T_\Theta(S) \;=\; (\Theta_{11}S + \Theta_{12})\;(\Theta_{21}S + \Theta_{22})^{-1}. \tag{15}$$

Then $T_\Theta(S)$ is still inner and we have

$$\delta\,(T_\Theta(S)) \;=\; \delta\,(\Theta) \;+\; \delta\,(S),$$

where δ denotes the McMillan degree. Also note that if S and Θ have real coefficients, so does $T_\Theta(S)$. These results can be found in [ABG] in the set-up of the disk.

From a realization point of view, the linear fractional transformation T_Θ, for a J-inner function Θ, works as follows

Lemma 2 *Let Θ be a $J-$inner function such that $\Theta(\infty) = I_{2p}$:*

$$\Theta = \left(\begin{array}{c|cc} \mathcal{A} & -\mathcal{P}^{-1}U^* & \mathcal{P}^{-1}V^* \\ \hline U & I_p & 0 \\ V & 0 & I_p \end{array} \right),$$

where $\mathcal{P}$ is the positive definite solution to the Lyapunov equation

$$\mathcal{A}^*\mathcal{P} + \mathcal{P}\mathcal{A} + U^*U - V^*V = 0.$$

Let $S = \left(\begin{array}{c|c} A & -P^{-1}C^* \\ \hline C & I_p \end{array} \right)$ be an inner function, where P is the solution to the Lyapunov equation $A^*P + PA + C^*C = 0$. Then the inner function $T_\Theta(S)$ has minimal realization:

$$T_\Theta(S) = \left(\begin{array}{cc|c} A & P^{-1}C^*V & -P^{-1}C^* \\ -\mathcal{P}^{-1}U^*C & \mathcal{A} + \mathcal{P}^{-1}(U^* - V^*)V & -\mathcal{P}^{-1}(U^* - V^*) \\ \hline C & U - V & I_p \end{array} \right) \tag{16}$$

and $[T_\Theta(S)](\infty) = I_p$.

Again these are well known facts [K]. Taking $S = I_p$, we immediately get the following special but important case:

Corollary 2 *The function $Q = T_\Theta(I_p)$ is inner and an n-dimensional (minimal) realization is given by*

$$Q = \left(\begin{array}{c|c} \mathcal{A} + \mathcal{P}^{-1}[U^* - V^*]V & -\mathcal{P}^{-1}[U^* - V^*] \\ \hline U - V & I_p \end{array} \right). \tag{17}$$

The observability gramian associated with this realization is $\mathcal{P}$.

4 The Schur algorithm

In the sequel, we shall denote by $\mathcal{I}_n^p(\infty)$ the set of $(p \times p)$-rational inner functions of McMillan degree n which satisfy the condition $Q(\infty) = I_p$. This normalization is quite standard in the literature: in fact, in most applications (inner-outer factorization, Douglas-Shapiro-Shields factorization, etc.) the inner function is only determined up to a constant unitary factor.

We consider now a particular case of the *tangential Schur algorithm* [D1] in which degree one $J-$inner factors analytic in Π^+ are involved. In the set-up of the disk, such an algorithm is used in [ABG] to construct an atlas of charts for the manifold of inner functions of fixed McMillan degree. We shall follow this approach, but now in the set-up of the left half plane. Let $\mathcal{A} = -\bar{\omega}$,

$\omega \in \Pi^+$ and $\mathcal{C} = \begin{bmatrix} u \\ v \end{bmatrix}$, with $u, v \in \mathbb{C}^p$ such that $\|v\| < \|u\|$ (where $\|\ \|$ denotes the usual Euclidean norm), and define $\Theta(\omega, u, v)$ as

$$\Theta(\omega, u, v) := \left(\begin{array}{c|cc} -\bar{\omega} & -\sigma^{-1}u^* & \sigma^{-1}v^* \\ \hline u & I_p & 0 \\ v & 0 & I_p \end{array} \right), \quad \sigma = \frac{\|u\|^2 - \|v\|^2}{\omega + \bar{\omega}}. \tag{18}$$

The tangential Schur algorithm consists in constructing from an inner function $Q \in \mathcal{I}_n^p(\infty)$, a sequence of inner functions

$$Q^{(n)} = Q, Q^{(n-1)}, \ldots, Q^{(1)}, Q^{(0)} = I_p,$$

of decreasing McMillan degree: assume that $Q^{(i)}$ of McMillan degree i has been constructed; let $\omega_i \in \Pi^+$ and find $u_i \in \mathbb{C}^p$ such that the vector $v_i \in \mathbb{C}^p$ defined by the interpolation condition

$$v_i := Q^{(i)}(\omega_i)^* u_i, \quad i = 1, \ldots n, \tag{19}$$

satisfies $\|v_i\| < \|u_i\|$. The term *tangential* comes from the fact that the interpolation is taken in some direction u_i. Such a vector u_i can always be found since otherwise the matrix $Q^{(i)}$ would be constant. Then, it can be proved that

$$Q^{(i)} = T_{\Theta_i}(Q^{(i-1)}), \quad i = 2, \ldots n, \tag{20}$$

for some inner matrix $Q^{(i-1)}$ of McMillan degree $i - 1$, where $\Theta_i = \Theta(\omega_i, u_i, v_i)$ is the J-inner matrix given by (18). Finally,

$$Q = T_{\Theta_n} \circ T_{\Theta_{n-1}} \cdots T_{\Theta_1}(I_p) = T_{\Theta_n \cdots \Theta_1}(I_p).$$

The tangential Schur algorithm enables us to construct explicit charts for $\mathcal{I}_n^p(\infty)$ as follows: given $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$, $\omega_i \in \Pi^+$, $i = 1, \ldots, n$, and $U = (u_1, u_2, \ldots, u_n)$, $\|u_i\| = 1$, $i = 1, \ldots, n$, define

$$\mathcal{V}_{(\omega,U)} = \{Q \in \mathcal{I}_n^p(\infty) \ / \ \|Q^{(i)}(\omega_i)^* u_i\| < 1, i = 1, \ldots, n\}.$$

This set is open in $\mathcal{I}_n^p(\infty)$ for the topology induced by the norm $\|\ \|_\infty$, and each function $Q \in \mathcal{I}_n^p(\infty)$ belongs to one of these sets. We thus have a covering of $\mathcal{I}_n^p(\infty)$ by open neighborhoods of the form $\mathcal{V}_{(\omega,U)}$ and these "coordinate neighborhoods" correspond to open subsets of $\mathbb{R}^{2np}$ by the local homeomorphism

$$\varphi_{(\omega,U)}: \begin{array}{ccc} \mathcal{V}_{(\omega,U)} & \rightarrow & \mathcal{B}_n^p \\ Q & \rightarrow & (v_1, v_2, \ldots, v_n) \end{array},$$

where the matrices $Q^{(i)}$ and the vectors v_i are computed recursively by (19) and (20), and $\mathcal{B}_n^p$ denotes the product of n copies of the open unit ball of $\mathbb{C}^p$. The details can be adapted from [ABG]. Note that an atlas for $\mathcal{RI}_n^p(\infty)$ can be obtained in a similar way: we simply have to impose the constraint that the points $\{\omega_i; i = 1, ..., n\}$ belong to the positive real axis $\mathbb{R}^+$ and the vectors u_i and v_i have real components. The range of the charts is thus the product of n copies of the open unit ball of $\mathbb{R}^p$.

The following Lemma links the Schur algorithm to the general transformation T_Θ discussed in the previous section:

Lemma 3 *Let $\omega_i \in \Pi_+$ and $u_i, v_i \in \mathbb{C}^p$ with $\|v_i\| < \|u_i\|$ for $i = 1, ..., n$. The J-inner matrix of degree n*

$$\Theta = \Theta_n \Theta_{n-1} \dots \Theta_i \dots \Theta_1,$$

where $\Theta_i = \Theta(\omega_i, u_i, v_i)$, $i = 1, \dots, n$, is the J-inner matrix given by (18)*, has minimal realization*

$$\Theta = \left(\begin{array}{cccc|c} -\bar{\omega}_1 & & & & \mathcal{B}_1 \\ \mathcal{B}_2\mathcal{C}_1 & \ddots & & & \mathcal{B}_2 \\ \vdots & & & & \vdots \\ \mathcal{B}_n\mathcal{C}_1 & \dots & \mathcal{B}_n\mathcal{C}_{n-1} & -\bar{\omega}_n & \mathcal{B}_n \\ \hline \mathcal{C}_1 & \dots & \dots & \mathcal{C}_n & I_{2p} \end{array} \right) \tag{21}$$

where for $i = 1, \dots, n$,

$$\mathcal{C}_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix}, \qquad \mathcal{B}_i = -\sigma_i^{-1}\mathcal{C}_i^* J, \qquad \sigma_i = (\|u_i\|^2 - \|v_i\|^2)/(\omega_i + \bar{\omega}_i).$$

The Gram matrix associated with this realization is diagonal and equal to:

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{pmatrix} \tag{22}$$

Proof: By induction over i. Clearly the statement is true for $i = 1$ (by definition). Let now

$$\Theta^{(i)} := \Theta_i \ \Theta_{i-1} \cdots \Theta_1$$

and denote a realization of $\Theta^{(i)}$ as

$$\Theta^{(i)} = \left(\begin{array}{c|c} \mathcal{A}^{(i)} & -[\Sigma^{(i)}]^{-1}[\mathcal{C}^{(i)}]^* J \\ \hline \mathcal{C}^{(i)} & I_{2p} \end{array} \right)$$

Then the cascade of two systems yields:

$$\begin{aligned} \Theta^{(i)} = \Theta_i \Theta^{(i-1)} &= \left(\begin{array}{c|c} -\bar{\omega}_i & -\sigma_i^{-1}\mathcal{C}_i^* J \\ \hline \mathcal{C}_i & I_{2p} \end{array} \right) \left(\begin{array}{c|c} \mathcal{A}^{(i-1)} & -[\Sigma^{(i-1)}]^{-1}[\mathcal{C}^{(i-1)}]^* J \\ \hline \mathcal{C}^{(i-1)} & I_{2p} \end{array} \right) \\ &= \left(\begin{array}{cc|c} \mathcal{A}^{(i-1)} & 0 & -[\Sigma^{(i-1)}]^{-1}[\mathcal{C}^{(i-1)}]^* J \\ -\sigma_i^{-1}\mathcal{C}_i^* J \mathcal{C}^{(i-1)} & -\bar{\omega}_i & -\sigma_i^{-1}\mathcal{C}_i^* J \\ \hline \mathcal{C}^{(i-1)} & \mathcal{C}_i & I_{2p} \end{array} \right) \end{aligned}$$

It is then easily proved that $\Sigma = \mathrm{diag}\{\sigma_1, \dots, \sigma_n\}$ is the unique solution to the Lyapunov equation (10). □

It turns out that every $J-$inner function Θ can be represented in the above form: that means that the Schur algorithm can be used to represent any rational transformation T_Θ. This could be seen by using the Potapov factorization (see [P]) of a $J-$inner function and then applying the above Lemma. Nevertheless we prefer to exhibit a direct proof, since it gives a constructive method which will be needed in the sequel (a discrete time version of this proof can be found in [LK]).

Lemma 4 *(Potapov factorization) Let $\Theta(s)$ be a rational $J-$inner function of McMillan degree n. Then there exists a realization $\Theta = \left(\begin{array}{c|c} \mathcal{A} & -\Sigma^{-1}\mathcal{C}^* J \\ \hline \mathcal{C} & I_{2p} \end{array}\right)$ of the form (21), where $\Sigma =$* $\mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$ *is a prescribed diagonal positive definite matrix.*

Proof: First observe that by Lemma 1 and equation (10) a minimal realization

$$\Theta = \left(\begin{array}{c|c} \mathcal{A} & -\mathcal{P}^{-1}\mathcal{C}^* J \\ \hline \mathcal{C} & I_{2p} \end{array}\right)$$

of a J-inner function in which $\mathcal{A}$ is lower triangular and $\mathcal{P}$ diagonal is necessarily of the form described in Lemma 3.

Let $\Theta = \left(\begin{array}{c|c} \mathcal{A}_0 & -\mathcal{P}^{-1}\mathcal{C}_0^* J \\ \hline \mathcal{C}_0 & I_{2p} \end{array}\right)$ be a minimal realization. Then we can diagonalize $\mathcal{P}$ by means of a unitary congruence transformation and, since $\mathcal{P} > 0$, we can make it equal to the identity by a diagonal congruence transformation. The induced state space transformation will yield $\Theta = \left(\begin{array}{c|c} \mathcal{A}' & -(\mathcal{C}')^* J \\ \hline \mathcal{C}' & I_{2p} \end{array}\right)$. Now we can find a unitary matrix M which makes $\mathcal{A}'' = M^*\mathcal{A}'M$ lower triangular. Since the identity is invariant by unitary congruence transformation, we obtain a realization $\Theta = \left(\begin{array}{c|c} \mathcal{A}'' & -(\mathcal{C}'')^* J \\ \hline \mathcal{C}'' & I_{2p} \end{array}\right)$ of the form (21) and with a Gram matrix equal to the identity. Then, with the change of basis $T = \Sigma^{-1/2}$, we obtain the desired realization with $\mathcal{A} = \Sigma^{-1/2}\mathcal{A}''\Sigma^{1/2}$, and $\mathcal{C} = \mathcal{C}''\Sigma^{1/2}$. □

In view of what we saw above, the Schur algorithm can be expressed in a very compact and simple form, as the following Proposition shows.

Proposition 1 *Let $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$, where $\omega_i \in \Pi^+$, and $U = (u_1, u_2, \ldots, u_n)$, where u_i is a unit vector in $\mathbb{C}^p$. The local homeomorphism $\varphi^{-1}_{(\omega,U)}$ has the explicit representation*

$$\varphi^{-1}_{(\omega,U)}(v_1, v_2, \ldots, v_n) = \left(\begin{array}{c|c} A & -\Sigma^{-1}C^* \\ \hline C & I_{2p} \end{array}\right),$$

where the observability gramian of $Q = \varphi^{-1}_{(\omega,U)}(v_1, v_2, \ldots, v_n)$ is equal the diagonal matrix $\Sigma =$ $\mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$, *with*

$$\sigma_i = \frac{u_i^* u_i - v_i^* v_i}{\omega_i + \bar{\omega}_i},$$

the matrix A has entries

$$[A]_{ij} = \begin{cases} \quad - \; (\omega_i + \bar{\omega}_i)\dfrac{u_i^*(u_j - v_j)}{\|u_i\|^2 - \|v_i\|^2} & i > j, \\ -\bar{\omega}_i \; + \; (\omega_i + \bar{\omega}_i)\dfrac{(u_i^* v_i - v_i^* v_i)}{\|u_i\|^2 - \|v_i\|^2} & i = j, \\ \quad (\omega_i + \bar{\omega}_i)\dfrac{(u_i^* - v_i^*)v_j}{\|u_i\|^2 - \|v_i\|^2} & i < j, \end{cases} \tag{23}$$

and

$$C = \begin{pmatrix} u_1 - v_1 & \dots & u_n - v_n \end{pmatrix}.$$

Proof: this is an immediate consequence of Corollary 2 and Lemma 3. □

The realization of Q given by Proposition 1 is relatively simple to use in computations by state space formulas. Nevertheless, the values σ_i can be arbitrarily small in a given chart, and this leads to a bad behavior in numerical computations of the inverse of Σ (which is well-known to be the controllability gramian). Such computations are useful in many situations such as for example the rational approximation problem addressed in section 7. To avoid this unsatisfactory numerical behavior, a different atlas is constructed in the next section by imposing that Σ is constant.

5 A new atlas.

In the atlas described in section 4 we take the vector ω and the matrix U to parametrize the charts, and the matrix V to construct the coordinates in each chart. This choice is made to build that particular atlas; however it is not the only possible one. In fact, as long as we have a set of points $\omega_i \in \Pi^+$ for $i = 1, ..., n$ and a corresponding set of pairs $u_i, v_i \in \mathbb{C}^p$ such that $\|v_i\| < \|u_i\|$, we can construct an inner function Q satisfying (19). Conversely, given an inner function $Q^{(i)}$, whenever we have such a triple (u_i, v_i, ω_i) satisfying (19), we can find $Q^{(i-1)}$ satisfying (20). So, there is a lot of freedom in the way these triples (u_i, v_i, ω_i) are chosen at each iteration to build an atlas. We will make use of this freedom to define a new atlas: we will impose that the Gram matrix Σ in (22) is constant and is given by $\Sigma = \mathrm{diag}\{\sigma_1, \dots, \sigma_n\}$ where the σ_i are positive numbers; we take these values instead of the interpolation points ω_i to be the parameters of our charts. We will show the following result: if we impose that, for $i = 1, ..., n$, each ω_i is *real*, then ω_i can be determined as the unique solution to the equation

$$\frac{\|u_i\|^2 - \|Q^{(i)}(\omega_i)^* u_i\|^2}{2\omega_i} = \sigma_i.$$

In this manner, we construct a new atlas which presents a better computational behavior but in which the domain of the charts is shortened. As illustrated in section 6.1, this can be improved by taking the values v_i in (19) to be the parameters of our charts while the vectors u_i can be used as coordinates.

Note that, since $Q^{(i)}$ is inner, $\left[Q^{(i)}\right]^{-1}(s) = Q^{(i)}(-\bar{s})^*$ a.e. and, if ω_i is not a zero of $Q^{(i)}$, the interpolation condition (19) can be rewritten as:

$$u_i = Q^{(i)}(-\bar{\omega}_i)v_i. \tag{24}$$

In order to proceed we need to see under which conditions we can find such a ω_i, given v_i and σ_i. This will give us the domain of our charts. This is the purpose of the next two lemmas.

Lemma 5 *Let* $Q = \left(\begin{array}{c|c} A & -P^{-1}C^* \\ \hline C & I_p \end{array}\right)$ *be a minimal realization of the inner function* Q*, the observability gramian* P *being positive definite, and let* s_Q *be the smallest real number such that* $2s_Q P - C^*C \geq 0$*. Then, for any* $v \in \mathbb{C}^p$ *such that* $v \notin \ker C^*$*, the function*

$$g_{Q,v}(s) = v^* \frac{Q^{-1}(s)Q^{-1}(s)^* - I_p}{2s} v \tag{25}$$

is continuously differentiable and strictly decreasing on the open interval $(s_Q, +\infty)$*.*

Proof: Observe that $Q^{-1} = \left(\begin{array}{c|c} -P^{-1}A^*P & P^{-1}C^* \\ \hline C & I_p \end{array}\right)$ is a minimal realization of Q^{-1} and that $K(s,\omega) = \frac{Q^{-1}(s)Q^{-1}(w)^* - I_p}{s+\bar{\omega}}$ is the reproducing kernel of $H(Q^{-1})$ where Q^{-1} is viewed as a $(-I_p)$-inner function; thus this kernel can be represented by (13) with $\mathcal{D} = I_p$ and $J_{p,q} = -I_p$. Consider the restriction of the kernel to the set $\{(s,s) | s \in (s_Q, +\infty)\}$:

$$L(s) := K(s,s) = \frac{Q^{-1}(s)Q^{-1}(s)^* - I_p}{2s};$$

then $g_{Q,v}(s) = v^* L(s)\, v$ and we get

$$\begin{aligned}
\frac{d}{ds}L(s) &= \frac{d}{ds}\left[C(sI_n + P^{-1}A^*P)^{-1}P^{-1}(sI_n + PAP^{-1})^{-1}C^*\right] \\
&= \frac{d}{ds}\left[CP^{-1}(sI_n + A^*)^{-1}P(sI_n + A)^{-1}P^{-1}C^*\right] \\
&= -CP^{-1}(sI_n + A^*)^{-2}P(sI_n + A)^{-1}P^{-1}C^* \\
&\quad -CP^{-1}(sI_n + A^*)^{-1}P(sI_n + A)^{-2}P^{-1}C^* \\
&= -CP^{-1}(sI_n + A^*)^{-2}[(sI_n + A^*)P + P(sI_n + A)](sI_n + A)^{-2}P^{-1}C^* \\
&= -CP^{-1}(sI_n + A^*)^{-2}[2sP - C^*C](sI_n + A)^{-2}P^{-1}C^*.
\end{aligned}$$

Now, since we assume $s > s_Q$, it follows that $2sP - C^*C > 0$. Moreover, $2sP - C^*C = (sI_n + A^*)P + P(sI_n + A)$, so that, if η is an eigenvector of A associated to the eigenvalue λ, then $2(s + Re(\lambda))\eta^* P\eta > 0$ and thus $s > -Re(\lambda)$. This implies that for $s > s_Q$, $(sI_n + A)^{-1}$ is well-defined. Finally, $v^* \frac{d}{ds}L(s)v = 0$ is equivalent to $v \in \ker C^*$. If this is not the case, the function $v^* L(s) v$ is strictly decreasing on (s_Q, ∞); hence the conclusion. □

Remark that s_Q only depends on Q and not on the realization. If $v \in \ker C^*$, then the function $g_{Q,v}$ vanishes identically; this condition is still independent of the realization and we have: $v \in \ker C^*$ if and only if $Q(s)v = v$ for all s.

Lemma 6 *Let* $\sigma > 0$*, let* $v \neq 0$ *in* $\mathbb{C}^p$*, and define*

$$\mathcal{D}_{\sigma,v} := \{Q \text{ inner}; \quad \lim_{s \downarrow s_Q} g_{Q,v}(s) > \sigma\}. \tag{26}$$

For $Q \in \mathcal{D}_{\sigma,v}$*, the function* $g_{Q,v}$ *is invertible on* $(0, \lim_{s\downarrow s_Q} g_{Q,v}(s))$ *and the inverse* $g_{Q,v}^{-1}$ *is continuously differentiable.*

Proof: In view of Lemma 5, for all $Q \in \mathcal{D}_{\sigma,v}$ the derivative of $g_{Q,v}$ is always strictly negative and thus the function is invertible on its range $g_{Q,v}(s_Q, \infty)$. But this range is easily seen to be $(0, \lim_{s \downarrow s_Q} g_{Q,v}(s))$ since

$$\lim_{s \to +\infty} v^* \frac{Q(-s)^* Q(-s) - I_p}{2s} v = 0.$$

Since the derivative never vanishes, the continuous differentiability of $g_{Q,v}^{-1}$ now follows from the differentiability and the invertibility of $g_{Q,v}$. □

Define the function

$$\begin{aligned} e_{\sigma,v} &: \mathcal{D}_{\sigma,v} \mapsto \mathbb{R}^+ \\ e_{\sigma,v}(Q) &:= g_{Q,v}^{-1}(\sigma) \end{aligned} \tag{27}$$

Let now $Q \in \mathcal{I}_n^p(\infty)$ and let $\Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_n\}$ and $V = (v_1, \ldots, v_n)$, where $\sigma_i > 0$ and $v_i \in \mathbb{C}^p$ with $\|v_i\| = 1$ for $i = 1, \ldots, n$, be given. We would like to construct a chart having as parameters Σ and V. The chart will be constructed using the Schur algorithm described below. So, under which conditions can we define a sequence of Schur functions $Q^{(i)}$ from (Σ, V)? The answer is quite simple: we set $Q^{(n)} = Q$. If $Q \notin \mathcal{D}_{\sigma_n, v_n}$, then we stop since the function Q is not in our chart and therefore the sequence cannot be constructed. Otherwise, the inductive construction goes as follows:

1. if $Q^{(i)} \notin \mathcal{D}_{\sigma_i, v_i}$, then stop. Otherwise, set $\omega_i := e_{\sigma_i, v_i}(Q^{(i)})$
2. define $u_i := Q^{(i)}(-\omega_i) v_i$, (see (24))
   then, since $\sigma_i > 0$, $\|v_i\| < \|u_i\|$, and the Schur algorithm allows to construct $Q^{(i-1)}$ from $Q^{(i)}$:
3. $Q^{(i-1)} := T_{\Theta_i}^{-1} Q^{(i)}$, where $\Theta_i = \Theta(\omega_i, u_i, v_i)$ is still given by (18).

Theorem 1 *Let $\Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_n\}$, with $\sigma_i > 0, i = 1, \ldots, n$, and $V \in \mathbb{C}^{p \times n}$, with each column v_i of norm 1 (i.e. $v_i^* v_i = 1$); define, for each pair (Σ, V), the set $\mathcal{V}_{(\Sigma,V)}$ and the function $\varphi_{(\Sigma,V)}$ (defined on $\mathcal{V}_{(\Sigma,V)}$) as follows:*

$$\begin{aligned} \mathcal{V}_{(\Sigma,V)} &:= \{Q \in \mathcal{I}_n^p(\infty); Q^{(i)} \in \mathcal{D}_{\sigma_i, v_i},\ i = 1, \ldots, n\} \\ \varphi_{(\Sigma,V)}(Q) &:= \left[Q^{(1)}(-e_{\sigma_1, v_1}(Q^{(1)}))v_1, \ldots, Q^{(n)}(-e_{\sigma_n, v_n}(Q^{(n)}))v_n\right] \end{aligned}$$

Then the family $(\mathcal{V}_{(\Sigma,V)}, \varphi_{(\Sigma,V)})$ forms an atlas for the set $\mathcal{I}_n^p(\infty)$ whose topology coincides with the one induced by the H^∞ topology.

Proof: In view of what we said above, the sets $\mathcal{V}_{(\Sigma,V)}$ are well defined and not empty. Now, we show that they are open in the H^∞-induced topology. In fact, $\mathcal{D}_{\sigma_i, v_i}$ is the inverse image of the open set $(\sigma_i, +\infty)$ by the application f

$$f(Q^{(i)}) \quad := \quad \lim_{s \downarrow s_Q} v^* \frac{Q^{(i)}(-s)^* Q^{(i)}(-s) - I_p}{2s} v$$

and this is a continuous function on $\mathcal{I}_n^p(\infty)$, since the evaluation of rational function is continuous with respect to the topology induced on $\mathcal{I}_n^p(\infty)$ by the H^∞ topology (rational inner functions do not have poles on the boundary).

The family $\mathcal{V}_{(\Sigma,V)}$ covers $\mathcal{I}_n^p(\infty)$. To prove this, let $Q \in \mathcal{I}_n^p(\infty)$ and run the Schur algorithm described in section 4. At each step ω_i can be chosen arbitrarily so that we may assume that it is real and it is not a zero of $Q^{(i)}$. Then $v_i = Q(w_i)^* u_i$ cannot be equal to zero, and dividing by the norm, we may assume that v_i is a unit vector. Set $\sigma_i = (u_i^* u_i - v_i^* v_i)/2\omega_i$, $\Sigma := \operatorname{diag}\{\sigma_1, \ldots, \sigma_n\}$ and $V := [v_1, \ldots, v_n]$. Then clearly $Q \in \mathcal{V}_{(\Sigma,V)}$.

Next we prove that $\varphi_{(\Sigma',V')} \circ \varphi_{(\Sigma,V)}^{-1}$ is a diffeomorphism. We show first that $\varphi_{(\Sigma,V)}(Q)$ is invertible and that the inverse is continuously differentiable with respect to the H^∞-induced topology. In fact, given Σ, V, Q, set $U := \varphi_{(\Sigma,V)}(Q)$. Then, from Σ, U, V we can construct a unique Θ representing the Schur interpolation conditions as in Lemma 3. Since, from the formula (17),

$$Q = \left(\begin{array}{c|c} \mathcal{A} + \Sigma^{-1}(U^* - V^*)V & -\Sigma^{-1}(U^* - V^*) \\ \hline U - V & I_p \end{array} \right),$$

where

$$[\mathcal{A}]_{ij} = \begin{cases} -\dfrac{u_i^* u_j - v_i^* v_j}{\sigma_i} & i > j, \\ -\dfrac{\|u_i\|^2 - \|v_i\|^2}{2\sigma_i} & i = j, \\ 0 & i < j, \end{cases}$$

the function $\varphi_{(\Sigma,V)}$ is clearly invertible, and the inverse is a rational function of the coefficients of U. Therefore it is continuously differentiable with respect to U (always in the H^∞ topology on $\mathcal{I}_n^p(\infty)$). To see that $\varphi_{(\Sigma',V')} \circ \varphi_{(\Sigma,V)}^{-1}$ is a diffeomorphism, observe that, in view of Lemma 6 and definition (27), $e_{\sigma_n',v_n'}$ is a smooth function, and thus also $e_{\sigma_n',v_n'}(\varphi_{(\Sigma,V)}^{-1})$ is smooth. Since

$$Q^{(i)} = T_{\Theta_{i+1}}^{-1} \circ T_{\Theta_{i+2}}^{-1} \circ \ldots \circ T_{\Theta_n}^{-1}(\varphi_{(\Sigma,V)}^{-1}(U)) =: \left[\varphi_{(\Sigma,V)}^{-1}(U)\right]^{(i)}$$

and the Θ_i are rational functions, also $Q^{(i)}$ is a smooth function of U, and thus so is $e_{\sigma_i',v_i'}$ for $i = 1, \ldots, n$ which proves that for $i = 1, \ldots, n$ the column

$$u_i' = \left[\left(\varphi_{(\Sigma',V')} \circ \varphi_{(\Sigma,V)}^{-1}\right)(U)\right]_i = \left[\left(\varphi_{(\Sigma,V)}^{-1}(U)\right)^{(i)} \left(-e_{\sigma_i',v_i'}\left(\left(\varphi_{(\Sigma,V)}^{-1}(U)\right)^{(i)}\right)\right)\right] v_i'$$

is continuously differentiable. Since we can interchange the role of the two maps, this is actually a diffeomorphism.

The sets $\varphi_{(\Sigma,V)}^{-1}(\mathcal{V}_{(\Sigma',V')} \cap \mathcal{V}_{(\Sigma,V)})$ are open in the H^∞ topology, since $\varphi_{(\Sigma,V)}^{-1}$ is continuous and $\mathcal{V}_{(\Sigma,V)}$ is open in $\mathcal{I}_n^p(\infty)$. Since also $\varphi_{(\Sigma,V)}^{-1}$ is continuous with respect to the H^∞ topology, we immediately deduce that the topology induced by our atlas actually coincides with the H^∞-induced topology. $\square$

As previously, an atlas of $\mathcal{RI}_n^p(\infty)$ is obtained if we impose the constraint that the vectors u_i and v_i have real components.

Given Σ, U, V, it's easy to construct a minimal realization of $Q = \varphi_{(\Sigma,U)}(V)$:

Proposition 2 *Let $\Sigma = \mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$, $\sigma_i > 0$, $i = 1, \ldots, n$, and $V = (v_1, v_2, \ldots, v_n)$, $\|v_i\| = 1$, $i = 1, \ldots, n$. The local homeomorphism $\varphi^{-1}_{(\Sigma,V)}$ has the explicit representation*

$$\varphi^{-1}_{(\Sigma,V)}(u_1, u_2, \ldots, u_n) = \left(\begin{array}{c|c} A & -\Sigma^{-1}C^* \\ \hline C & I \end{array}\right),$$

where the matrix A has entries

$$[A]_{ij} = \begin{cases} -\dfrac{u_i^*(u_j - v_j)}{\sigma_i} & i > j \\ \dfrac{(u_i^* - v_i^*)v_i}{2\,\sigma_i} - \dfrac{u_i^*(u_i - v_i)}{2\,\sigma_i} & i = j \\ \dfrac{(u_i^* - v_i^*)v_j}{\sigma_i} & i < j \end{cases} \tag{28}$$

and

$$C = \begin{pmatrix} u_1 - v_1 & \ldots & u_n - v_n \end{pmatrix}.$$

More difficult is to check the domain. To this end we have the following simple Lemma. Let $\Sigma^{(i)} := \mathrm{diag}\{\sigma_1, .., \sigma_i\}$ and $C^{(i)} = \begin{pmatrix} u_1 - v_1 & \ldots & u_i - v_i \end{pmatrix}$.

Lemma 7 *Let $(\mathcal{V}_{(\Sigma,V)}, \varphi_{(\Sigma,V)})$ be the atlas defined in Theorem 1. Then U belongs to the domain of the parameters $\mathcal{U}_{(\Sigma,V)} = \varphi_{(\Sigma,V)}(\mathcal{V}_{(\Sigma,V)})$ if and only if for $i = 1, \ldots, n$*

$$\frac{u_i^* u_i - v_i^* v_i}{\sigma_i} > \rho\left([\Sigma^{(i)}]^{-1/2}[C^{(i)}]^* C^{(i)}[\Sigma^{(i)}]^{-1/2}\right) = \rho\left(C^{(i)}[\Sigma^{(i)}]^{-1}[C^{(i)}]^*\right), \tag{29}$$

where ρ denotes the spectral radius.

Proof: A necessary and sufficient condition for U to belong to $\mathcal{U}_{(\Sigma,V)}$ is that $U = \varphi_{(\Sigma,V)}(Q)$ for some Q such that $Q^{(i)}$ belongs to $\mathcal{D}_{\sigma_i,v_i}$, for $i = 1, \ldots, n$. Recall that $Q^{(i)}$ has a realization of the form

$$Q^{(i)} = \left(\begin{array}{c|c} A^{(i)} & -[\Sigma^{(i)}]^{-1}[C^{(i)}]^* \\ \hline C^{(i)} & I_p \end{array}\right).$$

By Lemma 5, if $Q^{(i)}$ belongs to $\mathcal{D}_{\sigma_i,v_i}$, then $w_i > s_{Q^{(i)}}$, where $s_{Q^{(i)}}$ is the smallest positive number such that

$$2s_{Q^{(i)}}\Sigma^{(i)} - [C^{(i)}]^* C^{(i)} > 0.$$

Conversely, if $w_i > s_{Q^{(i)}}$, since $g_{Q^{(i)},v_i}$ decreases for $s > s_{Q^{(i)}}$, we then have that $Q^{(i)} \in \mathcal{D}_{\sigma_i,v_i}$. Now, let us compute $s_{Q^{(i)}}$, which is by definition the smallest positive number such that

$$2s_{Q^{(i)}} > \frac{\|C^{(i)}\xi\|^2}{\xi^*\Sigma^{(i)}\xi},$$

for any column vector ξ. Therefore, putting $\eta = [\Sigma^{(i)}]^{1/2}\xi$, we have that

$$2s_{Q^{(i)}} = \sup_{\eta} \frac{\eta^* \left([\Sigma^{(i)}]^{-1/2}[C^{(i)}]^* C^{(i)}[\Sigma^{(i)}]^{-1/2}\right)\eta}{\|\eta\|^2},$$

so that

$$s_{Q^{(i)}} = \frac{1}{2}\,\rho\left([\Sigma^{(i)}]^{-1/2}[C^{(i)}]^* C^{(i)}[\Sigma^{(i)}]^{-1/2}\right) = \frac{1}{2}\rho\left(C^{(i)}[\Sigma^{(i)}]^{-1}[C^{(i)}]^*\right). \tag{30}$$

Since $w_i = \frac{u_i^* u_i - v_i^* v_i}{2\sigma_i}$ the proof is achieved. □

The condition (29) is quite complicated to check. But it is very easy to derive slightly more restrictive conditions: for example, if $\omega_n < \omega_i$ for $i = 1, ..., n-1$, it is sufficient to verify the condition only for $i = n$.

What usually occurs in applications, is that we minimize a function over the manifold $\mathcal{I}_n^p(\infty)$ by making use of an iterative scheme which selects different points in a chart of the atlas; after a while we might get at some inner function Q close to boundary of the chart: then we have to change chart at this point Q; to do that, we have to compute the coordinates of the inner function Q in a different chart. An apparently difficult point is that, since the ω_i are defined by implicit equations, when we change chart we run into problems because we have to solve those equations. But when a new chart is needed, we do not have, in the actual implementation of the algorithm, to go through the procedure outlined in Theorem 1. In fact, we can use the following well-known result (for the proof we refer again to [D1]):

Proposition 3 *(Nevanlinna-Pick problem) Let Q be an inner function normalized at ∞, let $\omega_1, \omega_2, \ldots, \omega_n$ be n distinct points in $\mathbb{C}^+$ and let $x_1, x_2, \ldots, x_n \in \mathbb{C}^p$ be given. Define $\mathcal{A}_0 := diag\{-\bar{\omega}_1, -\bar{\omega}_2, \ldots, -\bar{\omega}_n\}$, $X := [x_1, x_2, \ldots, x_n]$ and*

$$Y = [y_1, y_2, \ldots, y_n] := [(Q(\omega_1))^* x_1, (Q(\omega_2))^* x_2, \ldots, (Q(\omega_n))^* x_n].$$

If the Pick matrix $\mathcal{P}$ given by

$$\mathcal{P}_{ij} = \frac{x_i^* x_j - y_i^* y_j}{w_i + \bar{w}_j} \tag{31}$$

is positive definite, then $Q = T_\Theta(I_p)$, where $\Theta = \left(\begin{array}{c|cc} \mathcal{A}_0 & -\mathcal{P}^{-1}X^* & \mathcal{P}^{-1}Y^* \\ \hline X & I_p & 0 \\ Y & 0 & I_p \end{array}\right).$

In other words, given any inner matrix Q, it's very easy to obtain a J-inner matrix Θ such that $Q = T_\Theta(I_p)$; indeed, the Pick matrix is the gramian of the projection $P_{H(Q)}$ on $H(Q)$ of the set

$$E = \left\{\frac{x_i}{s + \bar{\omega}_i}\right\}_{i=1,\ldots,n};$$

this is easily shown from the relations

$$P_{H(Q)}\left(\frac{x_i}{s + \bar{\omega}_i}\right) = \frac{I_p - Q(s)Q(\omega_i)^*}{s + \bar{\omega}_i} x_i,$$

and

$$< \frac{I_p - Q(s)Q(\omega_i)^*}{s+\bar{\omega}_i}x_i, \frac{I_p - Q(s)Q(\omega_j)^*}{s+\bar{\omega}_j}x_j >= x_i^* \frac{I_p - Q(\omega_i)Q(\omega_j)^*}{\omega_i+\bar{\omega}_j}x_j.$$

Thus, the Pick matrix is positive definite if and only if the projection of E on $H(Q)$ is injective. This is equivalent to saying that $E \cap Q\mathcal{H}_2 = 0$. But E is a subspace in $\mathcal{H}_2$ coinvariant for the shift (its orthogonal complement is invariant), of dimension n, while $Q\mathcal{H}_2$ is an invariant subspace of codimension n. Thus, generically, the intersection is zero: if we choose $\omega_1, \dots, \omega_n$ distinct positive numbers and vectors $x_1, \dots, x_n$, we will generically get a positive definite Pick matrix and we will have $Q = T_\Theta(I_p)$, where Θ is the J-inner matrix given in Proposition 3.

Next, using the procedure described in the proof of Lemma 4, we can change basis to obtain Θ in Schur form (21) with a Gram matrix equal to the identity; The two realizations are linked by the relations

$$\mathcal{A} = T\mathcal{A}_0T^{-1}, \quad UT = X, \quad VT = Y, \quad \mathcal{P} = T^*T,$$

where T is the Cholewski factor of $\mathcal{P}$ (i.e. $\mathcal{P} = T^*T$). If none of the vectors v_i vanishes for $i = 1, ..., n$, (and this is still a generic condition), we can set $s_i := (v_i^* v_i)^{-1/2}$ and $S := \mathrm{diag}\{s_1, \dots, s_n\}$; we can eventually set $\mathcal{A}'' := S^{-1}\mathcal{A}S$, $\mathcal{C}'' := \mathcal{C}S$. Then $\Sigma = S^2$ and we obtain a chart of the new atlas which contains Q.

6 Examples.

6.1 Inner matrices of size 2×2 and McMillan degree 1.

It is easily proved that an inner matrix $Q \in \mathcal{I}_1^2(\infty)$ is necessarily of the form

$$Q(s) = \frac{\begin{pmatrix} p_1(s) & p_2(s) \\ -\widetilde{p}_2(s) & \widetilde{p}_1(s) \end{pmatrix}}{q(s)}$$

where p_1 and q are monic polynomials of degree n, q having all its roots in the left half-plane, p_2 is a polynomial of degree strictly less than n, $\widetilde{p}$ is defined for a polynomial p of formal degree n by $\widetilde{p}(s) = (-1)^n p(-s)$, and the following equation is satisfied: $p_1\widetilde{p}_1 + p_2\widetilde{p}_2 = q\widetilde{q}$. When $n = 1$ and Q is *real*,

$$Q(s) = \frac{\begin{pmatrix} s+\alpha_1 & \alpha_2 \\ \alpha_2 & s-\alpha_1 \end{pmatrix}}{s+\alpha}, \tag{32}$$

where $\alpha_1 \in \mathbb{R}$, $\alpha_2 \in \mathbb{R}$ and $\alpha = \sqrt{\alpha_1^2 + \alpha_2^2}$ must be strictly positive. In conclusion the set $\mathcal{RI}_1^2(\infty)$ is completely and uniquely described as (α_1, α_2) varies in $\mathbb{R}^2 \backslash \{(0,0)\}$.

Now, let us compute such an inner matrix using our Schur algorithm. Let

$$\sigma \in \mathbb{R}, \quad u = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, \quad v = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2, \quad \|v\| < \|u\|.$$

Using Proposition 2, we obtain

$$[\varphi^{-1}_{(\sigma,v)}(x_1,x_2)](s) = \frac{\begin{pmatrix} 2\sigma s-(x_1-y_1)^2+(x_2-y_2)^2 & -2(x_1-y_1)(x_2-y_2) \\ \\ -2(x_1-y_1)(x_2-y_2) & 2\sigma s+(x_1-y_1)^2-(x_2-y_2)^2 \end{pmatrix}}{2\sigma s+(x_1-y_1)^2+(x_2-y_2)^2}.$$

This matrix is of the form (32) with

$$\alpha = \frac{(x_1-y_1)^2+(x_2-y_2)^2}{2\sigma} = \frac{\|u-v\|^2}{2\sigma}.$$

Remark. Observe that if we fix u to have norm 1 and let v vary, the condition $\|v\| < \|u\|$ implies that α remains bounded by $1/\sigma$, while when we fix v and let u vary, then α may go to infinity, and this leads to a larger domain for the chart. This is why we choose the vectors v_i as chart parameters and the vectors u_i as Schur parameters.

Now, let for example $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Using any realization of the inner matrix (32), we can see that the smallest real s_Q of Lemma 5 is equal to α, and the function

$$g_{Q,v}(s) = \frac{\alpha-\alpha_1}{(s-\alpha)^2}$$

actually decreases on (α,∞). For Q to belong to $\mathcal{D}_{\sigma,v}$, the limit of $g_{Q,v}(s)$ as s tends to α must be greater that σ. But $\lim_{s\to\alpha^+} g_{Q,v}(s) = \infty$ unless $\alpha = \alpha_1$. Since $\alpha = \sqrt{\alpha_1^2+\alpha_2^2}$, this can only happen if $\alpha_2 = 0$ and $\alpha_1 > 0$. In this case $g_{Q,v}$ is the constant null function. Consequently, the domain of the chart is

$$\mathcal{V}_{(\sigma,v)} = \mathbb{R}^2\backslash\{(\alpha_1,\alpha_2);\ \alpha_1>0,\ \alpha_2=0\}.$$

By Lemma 7, the range of the chart is the open set

$$\mathcal{U} \ = \ \{u\in\mathbb{C};\quad u^*v>1\} = \{(x_1,x_2)\in\mathbb{R}^2;\quad x_1>1\}.$$

According to the theory, it can be verified directly that the map $\varphi^{-1}_{(\sigma,v)}$ is a diffeomorphism,

$$\varphi^{-1}_{(\sigma,v)}:\ \mathcal{U}\to\mathcal{RI}_1^2(\infty)\backslash\left\{\begin{pmatrix} \frac{s+\alpha}{s-\alpha} & 0 \\ 0 & 1 \end{pmatrix}, \alpha>0\right\}.$$

In order to describe the whole manifold $\mathcal{RI}_1^2(\infty)$, we need another chart, for instance the chart indexed by σ and $v' = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ which describes all the manifold except for the matrices of the form $\begin{pmatrix} 1 & 0 \\ 0 & \frac{s+\alpha}{s-\alpha} \end{pmatrix}$.

6.2 Scalar inner functions of McMillan degree 2.

A scalar inner *real* function can be written in the form

$$Q = \frac{\widetilde{q}}{q}, \quad \widetilde{q}(s) = (-1)^n \, q(-s),$$

where q is a monic polynomial of degree n which has all its roots in the left half-plane. The function Q is completely determined by the coefficients of q which range over the subset of $\mathbb{R}^n$:

$$\Delta_n = \{(q_{n-1}, \ldots, q_0) / \; q_k > 0, \; k = 0, \ldots, n-1\}.$$

Now let $\Sigma = \mathrm{diag}\{\sigma_1, \sigma_2\}$, $V = \begin{pmatrix} 1, 1 \end{pmatrix}$, and $U = \begin{pmatrix} u_1, u_2 \end{pmatrix}$, $\sigma_1, \; \sigma_2, \; u_1, \; u_2$ being real numbers. Using Proposition 2, we can compute the denominator $q(s) = det(sI_2 - A)$ of $\varphi_{(\Sigma,V)}^{-1}(U)$:

$$q(s) = s^2 + \left(\frac{(u_1-1)^2}{2\sigma_1} + \frac{(u_2-1)^2}{2\sigma_2}\right) s + \left(\frac{(u_1-1)^2}{2\sigma_1} \frac{(u_2+1)^2}{2\sigma_2}\right).$$

By Lemma 7, the set of parameters is given by

$$\mathcal{U}_{(\Sigma,V)} = \left\{ U = (u_1, u_2), \quad u_1 > 1, \;\; u_2 > 1 + \frac{\sigma_2(u_1-1)^2}{2\sigma_1} \right\}.$$

Now the domain of the chart, that is the set of couples (q_0, q_1) which satisfy

$$\left\{ \begin{array}{lcl} q_0 & = & \frac{(u_1-1)^2}{2\sigma_1} \frac{(u_2+1)^2}{2\sigma_2}, \\ q_1 & = & \frac{(u_1-1)^2}{2\sigma_1} + \frac{(u_2-1)^2}{2\sigma_2}, \end{array} \right. \tag{33}$$

for $(u_1, u_2) \in \mathcal{U}_{(\Sigma,V)}$, can be obtained either directly or following the line of section 5: from any realization

$$Q(s) = 1 - C(sI_2 - A)^{-1} P^{-1} C^* \quad \text{of} \quad Q(s) = \frac{s^2 - q_1 s + q_0}{s^2 + q_1 s + q_0},$$

we can see that the smallest positive number $s_Q^{(2)}$ such that $2s_Q^{(2)} P + C^* C > 0$ is equal to q_1. Since

$$g_{Q,v_2}(s) = \frac{2q_1(s^2+q_0)}{(s^2 - q_1 s + q_0)^2},$$

for Q to belong to $\mathcal{D}_{\sigma_2, v_2}$, we must have

$$\frac{2q_1(q_1^2 + q_0)}{q_0^2} > \sigma_2. \tag{34}$$

Studying the sign of the polynomial $\sigma_2 q_0^2 - 2q_1 q_0 - 2q_1^3$ in q_0, we can see that (34) is equivalent to

$$q_0 < \frac{q_1}{\sigma_2}(1 + \sqrt{1 + 2\sigma_2 q_1}).$$

Now, observe that a scalar inner function of degree one belongs to $\mathcal{D}_{\sigma,v}$, for all $\sigma > 0$ and $v \neq 0$ in $\mathbb{C}^p$. The domain of the chart is thus given by

$$\mathcal{V}_{(\Sigma,V)} = \{(q_0, q_1), \quad 0 < q_0 < \frac{q_1}{\sigma_2}(1 + \sqrt{1 + 2\sigma_2 q_1})\}.$$

In order to describe the whole manifold, an infinite number of charts associated with a sequence of positive numbers σ_2 that goes to zero, is necessary. This is the price to pay for having a constant Gramian within one chart (see the introduction).

7 Application to the L^2 rational approximation problem.

The L^2-approximation problem can be stated as follows: given $F \in \mathcal{H}_2$ of McMillan degree N, find H which minimizes

$$\|F - R\|_2^2,$$

as R ranges over the set of rational functions analytic in Π^+ of McMillan degree at most n. Let

$$H \ = \ Q\, G$$

be the Douglas–Shapiro–Shields factorization (see [F]) of such a best rational approximation H to F, Q being inner, while $G \in \mathcal{H}_2^-$, the "left half-plane analogue" of $\mathcal{H}_2$. It is then obvious that the columns of H must be the projection of the columns of F onto the space $H(Q)$. The following proposition allows to compute this projection by solving a Lyapunov equation. We will denote by $F_{j,k} \in \mathcal{H}_2$ the entry (j,k) of F, by $[M]^k$ the $k-$th column of a matrix M and by $[M]_j$ its $j-$th row.

Proposition 4 *Let H be a best rational approximation to F of McMillan degree n, and let*

$$H = \left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right)$$

be a minimal realization of H. Let Σ be the observability gramian solution to the Lyapunov equation

$$A^*\Sigma + \Sigma A + C^*C = 0.$$

Then, the columns of B are given by:

$$[B]^k = \Sigma^{-1} \sum_{j=1}^{p} F_{jk}(-A^*)[C^*]^j. \tag{35}$$

Moreover, we have

$$\|F - H\|_2^2 \ = \ \|F\|_2^2 - \sum_{j=1}^{p}\sum_{k=1}^{m}\sum_{l=1}^{p} [C]_j F_{jk}(-A^*)^* \Sigma^{-1} F_{lk}(-A^*)[C^*]^l. \tag{36}$$

Proof. Since the columns of H are the projection of that of F onto the RKHS $H(Q)$, we have

$$\begin{aligned}
\xi^*[H(\omega)]^k \ &= \ < [F(.)]^k, K_Q(.,\omega)\xi > \\
&= \ \frac{1}{2\pi}\mathrm{Tr}\int_{-\infty}^{\infty} [F(iy)]^k [K_Q(iy,\omega)\xi]^*\, dy \\
&= \ \frac{1}{2\pi}\int_{-\infty}^{\infty} \xi^* K_Q(iy,\omega)^*[F(iy)]^k\, dy \\
&= \ \frac{1}{2\pi}\int_{-\infty}^{\infty} \xi^* C(\omega I - A)^{-1}\Sigma^{-1}(-iy - A^*)^{-1}C^*[F(iy)]^k\, dy,
\end{aligned}$$

using formula (13). Changing iy into s, we get

$$B = \Sigma^{-1} \frac{1}{2i\pi} \int_{i\mathbb{R}} (sI + A^*)^{-1} C^* F(s)\, ds.$$

Finally,

$$\begin{aligned} [B]^k &= \Sigma^{-1} \sum_{j=1}^{p} \frac{1}{2i\pi} \int_{i\mathbb{R}} (sI + A^*)^{-1} [C^*]^j F_{jk}(s)\, ds \\ &= \Sigma^{-1} \sum_{j=1}^{p} F_{jk}(-A^*) [C^*]^j, \end{aligned}$$

using the Riesz-Dunford functional calculus (see [F, II.2]). Now,

$$\|F - H\|_2^2 = \|F\|_2^2 - < F, H >,$$

and it is easily proved that $< F, H >= \mathrm{Tr}(B^* \Sigma B)$, which gives (36). □

Our approximation problem can then be approached by minimizing the criterion $\Psi_n(C, A)$ given by (36) over the set of equivalence classes of observable pairs (C, A) for which the spectrum of A is in the left-half plane. This set is diffeomorphic to the set $\mathcal{I}_n^p(\infty)$ and can be described using our atlas and the formulas given in Proposition 2. Then, the optimization can be tackled by using a search algorithm (gradient, Newton, etc.) through the manifold as a whole, using the charts to describe the manifold locally and changing from one chart to another when required (see [FO]). The gradient can be computed as

$$\frac{\partial \Psi_n}{\partial u_i} = -1/2 \sum_{j=1}^{p} \sum_{k=1}^{m} \left[\begin{array}{cc} 0 & C \end{array}\right]_j (F_{jk} - H_{jk}) \left(- \left[\begin{array}{cc} A & 0 \\ -\frac{\partial A}{\partial u_i} & A \end{array}\right] \right) \left[\begin{array}{c} B \\ 0 \end{array}\right]^k,$$

where B is still given by (35) and then also depends on Σ^{-1}. It is therefore clear that the parametrization of inner functions and then observable pairs described above is a very natural one for this problem since the observability gramian Σ is constant so that inversion and differentiation become much simpler to compute. An extensive study of this approximation problem, including numerical examples, will be provided in [GO].

References

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

[BO] L.Baratchart and M.Olivi, Critical points and error rank in best H^2 matrix rational approximation of fixed McMillan degree, *Constructive Approximation* 14:273–300, 1998.

[BGR] J. Ball, I. Gohberg, and L. Rodman, *Interpolation of rational matrix functions*, Birkhäuser Verlag, Basel, 1990.

[D1] H. Dym. *J-contractive matrix functions, reproducing kernel spaces and interpolation,CBMS lecture notes*, Vol. 71, American mathematical society, Rhodes island, 1989.

[F] P. A. Fuhrmann , *Linear Systems and Operators in Hilbert Space*, McGraw-Hill, 1981.

[FH] P. A. Furhmann and U. Helmke, *Homeomorphism between observable pairs and conditioned invariant subspaces*, Systems and Control Letters, 30:217-223, 1997.

[FO] P. Fulcheri and M. Olivi, Matrix rational H^2–approximation: a gradient algorithm based on Schur analysis, *SIAM Journal on Control and Optimization*, 36(6):2103–2127, 1998.

[GVKDM] Y. Genin, P. Van Doren, T. Kaylath, J.-M. Delosme, and M. Morf, On σ lossless transfer functions and related questions,*Linear Algebra and its Applications*, 50:251–275, 1983.

[GO] A. Gombani and M. Olivi, State space L^2 approximation, in preparation.

[H] K. Hoffman, *Banach Spaces of Analytic Functions,* Prentice-Hall, Englewood Cliffs, 1962.

[HO1] B. Hanzon and R. Ober, Overlapping Block-Balanced Canonical Forms and Parametrizations: the SISO case, *SIAM Journal on Control and Optimization*, 35:228-242, 1997.

[HO2] B. Hanzon and R. Ober, Overlapping Block-Balanced Canonical Forms for Various Classes of Linear Systems, *Linear Algebra and its Applications*, 281:171-225, 1998.

[HK] M. Hazewinkel and R. E. Kalman, On invariants, canonical forms and moduli for linear constant finite dimensional dynamical systems, in: G. Marchesini and S.K. Mitter (eds.), *Proceedings of the International Symposium on Mathematical System Theory, Udine, Italy,* Lecture Notes in Economics and Mathematical Systems, Berlin: Springer–Verlag, 131:48–60, 1976.

[K] H. Kimura, Chain-scattering approach to H^∞ control, Systems and Control: Foundations and Applications. Boston: Birkhaeuser, 1997.

[LK] H. Lev-Ari and T. Kailath, State-Space Approach to factorization of Lossless Transfer Functions and Structured Matrices, *Linear Algebra and its Applications*, 162–164:273–295 , 1992.

[O] R. Ober, Balanced parametrization of classes of linear systems, *SIAM Journal of Control and Optimization*, 29:1251-1287, 1991.

[P] Yu. Potapov, The multiplicative structure of J-contractive matrix functions, *Amer. Math. Soc. Trans. Ser. 2*, 15:131–244, 1960.

[HBV] P.S.C. Heuberger, O.H. Bosgra and P.M.J. Van den Hof, A generalized orthonormal basis for linear dynamical systems, *IEEE Trans. Autom. Control,*, 40(3):451-465, 1995.