

where r_1 and r_2 are the input and output loads and the matrix M is the *coupling matrix*. The structure of M (non-zeros entries) specifies the way resonators are coupled to one another. The McMillan degree of R corresponds to the number of circuits, that is the number of resonant modes or else two times the number of cavities.

The problem of extracting coupling parameters from frequency scattering measurements is essential with a view to reducing the cost of hardware and CAD tuning. The direct approach consists in feeding to a generic optimizer the function evaluating the scattering matrix from the coupling parameters, in order to fit the data. However, it often depends on a favorable initial guess and substantial efforts are currently being spent to design more robust methods. Another approach consists in first identifying a rational (linear) model from the data. Then, the coupling parameters are extracted from this rational model using classical design methods. In the filter community, the so-called Cauchy method is widely used to compute the rational model [24], [1]. Let us point out three major problems encountered in this direction, and in many other methods proposed in the literature:

- there is no guarantee on the stability of the rational model, i.e. the derived model can have unstable poles;
- there is no control on the McMillan degree of the model;
- no constraint is imposed to the model outside the frequency band of measurement (broadband), which may result in unrealistic behavior there.

Many toolboxes propose input output identification, while a few deal with frequency data. The software Vector Fitting and its Matrix Fitting extension (see [17], [18] and the bibliography therein) has become popular in the electromagnetic simulation community. However, the convergence towards a stable rational approximant, optimal in some least-square sense, is not guaranteed by this algorithm. Moreover no control is given in the MIMO case on the overall McMillan degree of the result, but only on its number of distinct poles. The same problem arises with the Frequency Domain Identification Toolbox [22] which only deals with SISO systems. This is unacceptable for the application we have in mind, in which the target McMillan degree is prescribed in advance and given by the number of coupled resonators present in the equivalent circuit of the filter.

To overcome these difficulties, we have developed a two stage approach to identify a rational model from the scattering data. A stable and causal model of high degree is first computed from the data (completion stage); then, model reduction is performed to get a model of the prescribed order. The first stage will be addressed in Section 3. Then we will consider the model reduction step. We tackle this problem using rational H^2 approximation and the original approach developed over the years in [6], [14], [28]. We present here the state of the art of this

approach which includes an efficient parametrization of balanced output pairs. The exposition is definitely application oriented, so that the emphasis will be put on the effective implementation of the method.

2 The Hilbert space framework

To deal with these completion and model reduction problems, we thus favor an approach based on approximation. A relevant context to deal with approximation is that of a Hilbert space. On the other hand, stability and causality of a rational model are equivalent to the analyticity of the transfer function in the closed right half-plane (poles at finite distance in the open left half-plane). We denote by \mathbb{C}^+ and \mathbb{C}^- the open right and left half-planes. To properly handle stability and causality, we embed rational functions in a larger space of analytic functions in \mathbb{C}^+ , namely a Hardy space naturally endowed with an L^2 norm. Note that, due to the low-pass transformation, the frequency data and the model that we consider do not satisfy the conjugacy requirement. This is why we consider Hardy spaces of *complex* functions.

The usual Hardy space of the half-plane, $H^2(\mathbb{C}^+)$, consists of functions f analytic in \mathbb{C}^+ , whose L^2 -norm remains uniformly bounded on vertical lines,

$$\sup_{x>0} \int_{-\infty}^{\infty} |f(x+i\omega)|^2 d\omega < \infty.$$

The Hardy space of the left half-plane, $H^2(\mathbb{C}^-)$, is defined in a similar way. An important fact is that the Laplace transform gives an isometry from $L^2(\mathbb{R}^\pm)$ onto $H^2(\mathbb{C}^\pm)$. It allows one to consider these Hardy spaces as subspaces of $L^2(i\mathbb{R})$, the image of $L^2(\mathbb{R})$ by the Laplace transform [32]. Moreover,

$$L^2(i\mathbb{R}) = H^2(\mathbb{C}^+) \oplus H^2(\mathbb{C}^-).$$

Each function in $L^2(i\mathbb{R})$ can thus be decomposed as the sum of a function in $H^2(\mathbb{C}^+)$ (stable part) and a function in $H^2(\mathbb{C}^-)$ (anti-stable part).

However, a stable causal function which fails to be strictly proper (to be 0 at ∞) does not belong to $L^2(i\mathbb{R})$. In order to include these functions in our setting, we shall replace the usual Lebesgue measure by the weighted measure $d\mu(w) = \frac{dw}{1+w^2}$, which also has the advantage to penalize high frequencies. The associated Hardy spaces $H_\mu^2(\mathbb{C}^+)$ and $H_\mu^2(\mathbb{C}^-)$ are defined in a similar way and can be viewed as subspaces of the space $L^2(d\mu)$ of functions defined on the imaginary axis and such that

$$\|f\|_\mu^2 = \int_{-\infty}^{\infty} |f(i\omega)|^2 \frac{d\omega}{1+\omega^2} < \infty.$$

However, $H_\mu^2(\mathbb{C}^+)$ and $H_\mu^2(\mathbb{C}^-)$ fail to be orthogonal complements, since their intersection is not empty (it contains for example constant functions). For $f \in L^2(d\mu)$, we denote by $P_+(f)$ its orthogonal projection onto $H_\mu^2(\mathbb{C}^+)$ (stable part) and by $P_-(f)$ its orthogonal projection on the orthogonal complement of $H_\mu^2(\mathbb{C}^+)$ (unstable part). Hardy spaces thus provide an interesting tool to estimate causality and stability of a given transfer function.

3 Compensation of delay components and completion of the data

After the low-pass frequency transformation, we suppose that the harmonic scattering measurements of the filter yield the knowledge of a 2×2 matrix function $\tilde{S}(iw)$ defined on a strict sub-interval J of the imaginary axis. In practice this function is obtained thanks to the interpolation (splines) of a discrete set of measurement points. The mathematical model we want to identify from these measurements is given by

$$\begin{bmatrix} e^{i\frac{\alpha}{2}h(w)} & 0 \\ 0 & e^{i\frac{\beta}{2}h(w)} \end{bmatrix} R(iw) \begin{bmatrix} e^{i\frac{\alpha}{2}h(w)} & 0 \\ 0 & e^{i\frac{\beta}{2}h(w)} \end{bmatrix},$$

where R is the 2×2 rational scattering matrix of the low-pass model of the filter and the exponential terms are due to the access lines used to perform the measurements. The transformation $h(w)$ maps normalized frequencies (low-pass model) to high frequencies (original system).

In order to cast the identification problem to a rational approximation problem we first need to identify the non-rational delay components, that is to evaluate α and β . We base our delay compensation procedure on analytic completion techniques [8], [7], [9]. The latter consist in extending partial frequency measurements performed on the broadband J to the whole imaginary axis under causality constraints. In the special case of our filter with rational response we will make the strong assumption that, if the delay components are properly compensated, the measurements can be extended on J_c (the complementary of J on the imaginary axis) by a polynomial of low order in the variable $1/iw$ such as to form a causal transfer function. In other words, if $S_{ij}(iw)$ denotes the measurements for which the delays have been compensated, we should be able to find a polynomial $p_{ij}(1/iw)$ of low degree such that the complemented elements $S_{ij}(iw) \vee p_{ij}(1/iw)$ have a "small" anti-causal component and have a smooth behavior at the boundaries of J . The simplicity of the extensions p_{ij} accounts for the absence of delay and the fact that the measurements on the broadband J already capture most of the complexity of the rational responses, that can therefore be represented on J_c by a short Taylor expansion p_{ij} at infinity. We expose in what follows the convex optimization problems that are considered to extract the delay

components and extend the data on the whole imaginary axis.

To a given value τ of a delay compensation, we associate the polynomial which gives the "most causal" completion

$$p_\tau = \arg \min_{p \in \mathcal{P}} \|P_-(\tilde{S}_{11}(iw)e^{-i\tau h(w)} \vee p(1/iw))\|_\mu^2, \quad (2)$$

where $\mathcal{P} = \{p; \deg p \leq n_c, \sup_{w \in J_c} |p(1/iw)| \leq 1\}$. This modular bound on p is meaningful as our filter is passive. We then choose the delay α to be the value of the compensation τ that gives the smallest discontinuities at the concatenation points between the data and p_τ . To determine β we proceed in the same manner using the measurements of \tilde{S}_{22} instead of those of \tilde{S}_{11} .

Now the delays are known, we improve the completion by relaxing (2) and imposing a better behavior near the concatenation points. We select a sub-collection of measurement indices $I = \{k, |w_k| > w_c\}$ where w_c is chosen sufficiently large (tail of the broadband J). We thus consider the optimization problem :

$$\min_{p \in \mathcal{P}} \sum_{k \in I} |p(1/iw_k) - S_{ij}(iw_k)|^2$$

under the additional constraint

$$\|P_-(S_{ij}(iw) \vee p(1/iw))\|_\mu^2 \leq E.$$

This problem has a unique optimal solution unless its admissible set is empty. This will be the case provided that $E \geq E_{min}$ where E_{min} is the optimal criterion obtained from the preceding problem. In practice the values $w_c = 2.5$ (normalized frequencies) and $n_c = 4$ seem to give very good results when the broadband is three time bigger than the passband.

If p_{ij} are the polynomial completions computed by the later method we define

$$\tilde{F}_{ij} = P_+(S_{ij}(iw) \vee p_{ij}(1/iw)).$$

Those functions can be seen as the compensated, causal, stable projections of our initial data; note that, by construction, their $L^2(d\mu)$ distance to compensated data S on J is less than \sqrt{E} .

4 From continuous-time to discrete-time

To deal with rational approximation, we shift to the disk or discrete-time framework. A good reason to do this is that the Hardy spaces of the disk are simpler in some sense than that of the half-plane. The fact that the unit disk has a finite Lebesgue measure has some nice implications, as the inclusions $L^\infty(\mathbb{T}) \subset L^2(\mathbb{T}) \subset L^1(\mathbb{T})$.

Moreover, functions in $L^2(\mathbb{T})$ can be represented by their Fourier series and we don't have to cope with sampling.

The space $L^2(\mathbb{T})$ splits into two orthogonal subspaces

$$L^2(\mathbb{T}) = H^2 \oplus H^2_{\perp},$$

where H^2 consists of functions whose Fourier coefficients of negative index are zero, while H^2_{\perp} consists of functions whose Fourier coefficients of non-negative index are zero. Hardy spaces $H^2(\mathbb{D})$ and $H^2(\mathbb{E})$ of the disk and its exterior $\mathbb{E} = \mathbb{C} \setminus \mathbb{D}$ may be defined as those of the half-planes (integrals on vertical lines are then replaced by integrals over circles). Then, by analytic continuation, H^2 can be identified with $H^2(\mathbb{D})$ and H^2_{\perp} with a strict subspace of $H^2(\mathbb{E})$ (functions vanishing at infinity). We denote by P_{H^2} and $P_{H^2_{\perp}}$ the orthogonal projections onto H^2 and H^2_{\perp} respectively. Discrete-time stable and causal rational transfer functions naturally belong to $H^2(\mathbb{E})$.

There are many ways to transform a continuous-time function into a discrete-time one. We shall use either the usual bilinear transform or a variant of it, with the Möbius transformation from the z -plane to the s -plane

$$z \mapsto s = \frac{z+1}{z-1},$$

which sends \mathbb{T} onto the imaginary axis and \mathbb{C}^+ onto \mathbb{E} .

4.1 Bilinear transform

The bilinear transform is the map

$$\tilde{F}(s) \mapsto F(z) = \tilde{F}\left(\frac{z+1}{z-1}\right).$$

This map is an isometry from $L^2(d\mu)$ onto $L^2(\mathbb{T})$ which preserves the McMillan degree. It sends the space $H^2_{\mu}(\mathbb{C}^+)$ onto $H^2(\mathbb{E})$.

4.2 $H^2(\mathbb{C}^+) \rightarrow H^2_{\perp}$ isometry

The map

$$\tilde{F}(s) \mapsto F(z) = \frac{\sqrt{2}}{z-1} \tilde{F}\left(\frac{z+1}{z-1}\right). \quad (3)$$

is an isometry from $L^2(i\mathbb{R})$ onto $L^2(\mathbb{T})$. It also preserves the McMillan degree. It sends the space $H^2(\mathbb{C}^+)$ onto H^2_{\perp} . With this transformation, the formulas which allow us to derive a realization $(A, B, C, 0)$ of $F(z)$ from a realization $(\tilde{A}, \tilde{B}, \tilde{C}, 0)$ of $\tilde{F}(s)$ and reciprocally, are

completely symmetric:

$$\begin{aligned} C &= \tilde{C} & \tilde{C} &= C \\ A &= -(I - \tilde{A})^{-1}(I + \tilde{A}) & \tilde{A} &= -(I - A)^{-1}(I + A) \\ B &= \sqrt{2}(I - \tilde{A})^{-1}\tilde{B} & \tilde{B} &= \sqrt{2}(I - A)^{-1}B \end{aligned}$$

Both methods are used to transport the function \tilde{F} obtained in Section 3 from continuous-time to discrete-time. If the $H^2(\mathbb{C}^+) \rightarrow H^2_{\perp}$ isometry is used, then the value at infinity must be taken off. It will be reset after the approximation step. This method preserves the value at infinity obtained in the completion stage. If the bilinear transform is used, the value at infinity can thus be improved by the rational approximation step. This method usually gives better results.

5 Stable rational approximation of given McMillan degree

In the completion stage, we dealt with each entry separately. In rational approximation, the constraint on the McMillan degree involves the whole scattering matrix, and it is not possible to handle each entry independently. A specific approach has been developed at INRIA to cope with this problem, which is based on the following points

- the optimization range is reduced to a compact set,
- an atlas of charts is used to parametrize the optimization domain.

To simplify our writing, we keep denoting by $L^2(\mathbb{T})$, H^2 and H^2_{\perp} the spaces of matrix-valued functions with entries in $L^2(\mathbb{T})$, H^2 and H^2_{\perp} respectively. The proper dimension of the matrix will be understood from the context. The L^2 -norm of a matrix-valued function derives from the scalar product

$$\langle F, G \rangle = \frac{1}{2\pi} \text{Tr} \int_0^{2\pi} F(e^{it})G(e^{it})^* dt.$$

The rational approximation problem we consider is, given a $p \times m$ matrix-valued function $F \in H^2(\mathbb{E})$, to minimize the L^2 distance to the set of rational stable functions of McMillan degree less than or equal to n . In our application $m = p = 2$.

Using the orthogonal decomposition

$$H^2(\mathbb{E}) = \mathbb{C} \oplus H^2_{\perp},$$

we can see that any solution H must satisfy $H(\infty) = F(\infty)$. Thus, we may restrict our study to the case of

strictly proper transfer functions, that is to the space H_{\perp}^2 .

A number of qualitative results are available in the literature which assert that the problem is well-posed and which pave the way to convergent algorithms. It was proved in [5] that *the global minimum of the L^2 criterion does exist*, as well as the *normality property*: if F is not of McMillan degree strictly less than n , then any local minimum of the criterion over the set of systems of order less than or equal to n has effective order n . The problem can thus be stated as:

Rational approximation problem. Given $F \in H_{\perp}^2$ of order $\geq n$, find \hat{H} such that

$$\hat{H} = \operatorname{argmin}_{H \in \mathcal{S}_n} \|F - H\|_2^2 \quad (4)$$

where

$$\mathcal{S}_n = \{H \in H_{\perp}^2, \deg H = n\}$$

is the set of rational strictly proper stable transfer functions of exact degree n .

The following *consistency* result must also be mentioned: if F has McMillan degree n , then the only critical point of the L^2 criterion is F itself [10].

The present approach was first proposed in the SISO case [6] and then in the MIMO case [14]. The first step is the reduction of the parameters space (see Section 6). The second step is to find a clever parametrization of the reduced optimization space namely the space of stable all-pass systems of fixed order. This parametrization, an atlas of charts, takes into account the precise structure of the space, namely a *non-trivial* differentiable manifold ([2], [13]). Several atlases have been proposed in the literature (see [2], [20]). The new atlas which is used in this paper particularly suits to state-space representations and has been preferred for computation facilities.

The question of the parametrization of LTI systems has been widely studied. The non-zero entries in classical canonical forms (companion, Hessenberg, tridiagonal forms) are often used as parameters because of their simplicity. See [37, chap.7] and [31] for an overview of these parametrization issues and their properties. Compared to these representations, our parametrization guarantees

- the non-redundancy in the parameters (injective mapping)
- restrictions on the system (stability, fixed order) automatically taken into account
- numerical robustness of the algorithm due to the use of unitary matrices

6 Reduction of the optimization set

In the SISO case, it is known that if $h = \frac{p}{q}$ is a best approximant of f then the numerator p can be easily computed from the denominator q by solving linear equations. The projection theorem in a Hilbert space asserts that h must be the projection of f onto the vector space $V_q = \{\frac{p}{q}; \deg p < n\}$. This space is the orthogonal complement of BH_{\perp}^2 in H_{\perp}^2 , where $B = \prod_{i=1}^n \frac{1-\bar{a}_i z}{z-a_i}$ is the Blaschke product whose denominator is q ,

$$H_{\perp}^2 = BH_{\perp}^2 \oplus V_q.$$

This approach was already developed in [33] and can be generalized to the MIMO case. The right generalization of the fraction description to the MIMO case is the Douglas-Shapiro-Shields factorization: a $p \times m$ rational matrix function $H \in H_{\perp}^2$ can be written in the form

$$H = GP, \quad (5)$$

where G is $p \times p$ lossless of McMillan degree n and $P \in H^2$. Recall that a rational lossless matrix is a matrix Blaschke product: $G(z)$ is contractive for $z \in \mathbb{D}$ and unitary for $z \in \mathbb{T}$. Multiplication by G lossless in H_{\perp}^2 is an isometry. In (5), G is called the lossless factor and brings the pole structure of H and thus its dynamics. It is unique up to a right unitary constant matrix $U \in \mathbb{U}_p$.

Now if $H = GP$ is a local approximant of F , thus H is completely determined from G as the orthogonal projection of F onto V_G

$$V_G = \{H \in H_{\perp}^2; H = GP, P \in H^2\}.$$

Equivalently, using multiplication by $G^{\sharp}(z) = G(z)^{-1}$ which is an isometry in $L^2(\mathbb{T})$, we get

$$\langle F - H, GH^2 \rangle = \langle G^{\sharp}F - C, H^2 \rangle = 0,$$

so that P is the orthogonal projection of $G^{\sharp}F$ onto H^2 ,

$$P = P_{H^2}(G^{\sharp}F).$$

The rational approximation problem is then to minimize the criterion

$$\psi_n : G \mapsto \|F - GP_{H^2}(G^{\sharp}F)\|^2 \quad (6)$$

over $\mathbb{L}_n^p/\mathbb{U}_p$ the right quotient of the set \mathbb{L}_n^p of $p \times p$ lossless functions of McMillan degree n by unitary constant matrices. In least-square optimization and using a state-space formulation, this elimination step is classical and known under the name of Separable Least Square. It presents some important advantages: the dimension of the parameter space is reduced and mostly, lossless functions enter the picture.

Let us derive the state-space formulation which is implemented in our software RARL2 [26]. Let

$$F(z) = \mathcal{C}(zI_N - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}. \quad (7)$$

be a realization of F and let $\hat{H}(z) = \mathcal{D} + \hat{C}(zI - \hat{A})^{-1}\hat{B}$ be a local approximant of F . The error $F - \hat{H}$ has realization

$$\tilde{A} = \begin{bmatrix} \mathcal{A} & 0 \\ 0 & \hat{A} \end{bmatrix}, \tilde{B} = \begin{bmatrix} \mathcal{B} \\ \hat{B} \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} \mathcal{C} & -\hat{C} \end{bmatrix}, \tilde{D} = 0,$$

and the square of the L^2 error can be computed as

$$\|F - H\|_2^2 = \text{Tr}(\tilde{B}^* \tilde{Q} \tilde{B}) = \text{Tr}(\tilde{C} \tilde{P} \tilde{C}^*),$$

where \tilde{P} and \tilde{Q} are the reachability and observability gramians of the error. Necessary conditions for optimality may easily be found by computing the gradient of the square of the L^2 -norm with respect to the state-space parameters. Partitioning the gramians in the same way as \tilde{A} ,

$$\tilde{P} = \begin{bmatrix} \mathcal{P} & P_{12} \\ P_{12}^* & \hat{P} \end{bmatrix}; \quad \tilde{Q} = \begin{bmatrix} \mathcal{Q} & Q_{12} \\ Q_{12}^* & \hat{Q} \end{bmatrix},$$

we get

$$\begin{aligned} Q_{12}^* \mathcal{B} &= -\hat{Q} \hat{B} \\ \mathcal{C} P_{12} &= \hat{C} \hat{P} \\ Q_{12}^* \mathcal{A} P_{12} &= -\hat{Q} \hat{A} \hat{P}. \end{aligned} \quad (8)$$

These necessary conditions were first obtained by Wilson [38] and many model reduction algorithms were proposed in the literature based on these conditions (see [29], [3], [21], [36], [19], [16] and [11]). However, to our knowledge no other algorithm is available which guarantees the stability of the approximant and works in the MIMO case.

The reduction of the optimization set translates in this state-space setting as follows. First choose a realization of \hat{H} such that the observable pair (\hat{C}, \hat{A}) is output normal

$$\hat{A}^* \hat{A} + \hat{C}^* \hat{C} = I,$$

which means that \hat{Q} is the identity. Then, the first necessary condition yields $\hat{B} = -Q_{12}^* \mathcal{B}$. The rational approximation problem is thus to minimize the criterion

$$J_n(C, A) = \|F\|_2^2 - \text{Tr}(\mathcal{B}^* Q_{12} Q_{12}^* \mathcal{B}) \quad (9)$$

over the set of output normal observable pairs (C, A) . Note that Q_{12} is determined from A and C as the solution of the Stein equation

$$\mathcal{A}^* Q_{12} \mathcal{A} + \mathcal{C}^* \mathcal{C} = Q_{12}. \quad (10)$$

The derivative of the criterion with respect to some parameter λ can be computed as

$$\frac{dJ_n}{d\lambda} = 2 \text{Re} \text{Tr} \left(\frac{dQ_{12}}{d\lambda} \mathcal{B} \mathcal{B}^* \right).$$

Using

$$\mathcal{A} P_{12}^* \mathcal{A}^* + \mathcal{B} \mathcal{B}^* = P_{12}^*,$$

we get

$$\frac{dJ_n}{d\lambda} = 2 \text{Re} \text{Tr} \left(P_{12}^* \left[\frac{dQ_{12}}{d\lambda} - \mathcal{A}^* \frac{dQ_{12}}{d\lambda} \mathcal{A} \right] \right).$$

Now differentiating (10) with respect to A and C , we get the relations

$$\begin{aligned} \frac{\partial Q_{12}}{\partial A} - \mathcal{A}^* \frac{\partial Q_{12}}{\partial A} \mathcal{A} &= \mathcal{A}^* Q_{12}, \\ \frac{\partial Q_{12}}{\partial C} - \mathcal{A}^* \frac{\partial Q_{12}}{\partial C} \mathcal{A} &= \mathcal{C}^* \end{aligned}$$

so that we finally have

$$\frac{dJ_n}{d\lambda} = 2 \text{Re} \text{Tr} \left(P_{12}^* \left[\mathcal{A}^* Q_{12} \frac{\partial A}{\partial \lambda} + \mathcal{C}^* \frac{\partial C}{\partial \lambda} \right] \right). \quad (11)$$

The connection between observable pairs and lossless functions is stressed by the following result (see [25]).

Proposition 1 (Lossless embedding) *Given an observable pair (C, A) with A asymptotically stable, let Q be its observability gramian. Then, the rational matrix $G(z) = D + C(zI - A)^{-1}B$, with*

$$\begin{aligned} B &= -(A - \nu I)Q^{-1}(I - \nu A^*)^{-1}C^* \\ D &= I - CQ^{-1}(I - \nu A^*)^{-1}C^* \end{aligned}$$

is lossless for every ν such that $|\nu| = 1$.

The lossless function G satisfies $G(\nu) = I$. The map $(C, A) \mapsto G$ is a one-to-one correspondence between the set of observable pairs (C, A) , A asymptotically stable, up to similarity, and the set of lossless functions up to a right unitary matrix.

This correspondence is in fact a diffeomorphism (see Cor.2.1 in [2]). If in addition the pair (C, A) is output normal, the matrix $[A \ C]^T$ has orthonormal columns and the lossless embedding consists in completing it into a unitary matrix

$$\begin{bmatrix} A \\ C \end{bmatrix} \mapsto \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

We shall call *unitary realization*, a realization (A, B, C, D) such that A is asymptotically stable and the corresponding realization matrix unitary. Unitary realizations are precisely the balanced realizations of lossless functions

(see [20, prop.3.2]). In what follows, we shall use balanced realizations and the associated unitary matrix to represent lossless matrices. This is of course very advantageous from a numerical viewpoint.

7 Minimization over a manifold.

We said that an atlas of charts happens to be the right representation in order to use differential tools (as a gradient algorithm). A differential manifold is precisely a set endowed with a differential structure by means of an atlas of charts. An atlas is a collection of charts or (coordinate) maps $\phi_i : \mathcal{D}_i \rightarrow \mathbb{R}^d$, where \mathcal{D}_i is an open subset of the manifold, which satisfy some compatibility conditions: the union of the \mathcal{D}_i covers the manifold and the transition maps or changes of coordinates $\phi_i \circ \phi_j^{-1}$ are smooth. By means of the coordinate maps, differential calculus on \mathbb{R}^d can be carried over to the manifold. The coordinate maps then become diffeomorphisms. The dimension of the manifold is d . An atlas of charts is thus the way to parametrize a non-trivial manifold in a local smooth manner. In the next section, we describe such an atlas of charts for the manifold \mathbb{L}_n^p of dimension $2np$.

7.1 Lossless mutual encoding

Let $\Omega = (W, X, Y, Z)$ be a unitary realization. The idea is to attach to Ω a chart or coordinate map.

Let $G(z) = D + C(zI - A)^{-1}B$ be a balanced realization of G lossless. Note that such a realization is unique up to a state isometry. Let Λ be the unique solution to the Stein equation

$$\Lambda - A^*\Lambda W = C^*Y \quad (12)$$

and

$$V = D^*Y + B^*\Lambda W \quad (13)$$

Remark. In fact, the matrix V satisfies

$$\frac{1}{2i\pi} \int_{\mathbb{T}} G^\sharp(z) Y (zI - W)^{-1} dz = V, \quad (14)$$

which is known as a Nudelman interpolation condition for G^\sharp (see [27]).

Formulas (12) and (13) can be rewritten in a matrix form

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \Lambda \\ V \end{bmatrix} = \begin{bmatrix} SW \\ Y \end{bmatrix}. \quad (15)$$

Note that the realization of G being balanced by assumption, $P = \Lambda^*\Lambda$ satisfies

$$Y^*Y + W^*PW = V^*V + P. \quad (16)$$

We now assume that P is positive definite, a condition which allows for a parametrization of the set of solutions of the Nudelman problem (14) [4]. The matrix Λ is thus invertible and we may normalize the triple (Y, W, V) as

$$(\tilde{Y}, \tilde{W}, \tilde{V}) = (Y\Lambda^{-1}, \Lambda W\Lambda^{-1}, V\Lambda^{-1}). \quad (17)$$

We get from (15)

$$\begin{bmatrix} \tilde{W} \\ \tilde{Y} \end{bmatrix}^* \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I \\ \tilde{V} \end{bmatrix} = \tilde{W}^*\tilde{W} + \tilde{Y}^*\tilde{Y}.$$

Under the assumption $P > 0$,

$$K = \tilde{W}^*\tilde{W} + \tilde{Y}^*\tilde{Y} = \tilde{V}^*\tilde{V} + I \quad (18)$$

is positive definite and if $K^{1/2}$ denotes its Hermitian square root, then

$$\begin{bmatrix} \tilde{W}K^{-1/2} \\ \tilde{Y}K^{-1/2} \end{bmatrix}^* \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} K^{-1/2} \\ \tilde{V}K^{-1/2} \end{bmatrix} = I.$$

We will now specify from V and Ω two unitary completions \mathcal{U}, \mathcal{V} of these orthonormal columns,

$$\mathcal{U} = \begin{bmatrix} \tilde{W}K^{-1/2} & * \\ \tilde{Y}K^{-1/2} & * \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} K^{-1/2} & * \\ \tilde{V}K^{-1/2} & * \end{bmatrix}$$

and define a map

$$\phi_\Omega : (A, B, C, D) \mapsto (V, D_0),$$

where D_0 is the unitary matrix such that

$$\mathcal{U}^* \begin{bmatrix} A & B \\ C & D \end{bmatrix} \mathcal{V} = \begin{bmatrix} I & 0 \\ 0 & D_0 \end{bmatrix}. \quad (19)$$

The map ϕ_Ω will be then invertible.

We must first fix the balanced realization (A, B, C, D) of G we start with. We will say that a realization of G is in canonical form with respect to Ω iff Λ given by (12) is positive definite and Hermitian. Then $\Lambda = P^{1/2}$ where P is the solution of (16). We denote by \mathcal{D}_Ω the set of unitary realizations in canonical form with respect to Ω .

The matrix \mathcal{V} is chosen according to Proposition 1 (with $\nu = -1$)

$$\mathcal{V} = \begin{bmatrix} (I + \tilde{V}^*\tilde{V})^{-1/2} & -\tilde{V}^*(I + \tilde{V}\tilde{V}^*)^{-1/2} \\ \tilde{V}(I + \tilde{V}^*\tilde{V})^{-1/2} & (I + \tilde{V}\tilde{V}^*)^{-1/2} \end{bmatrix}. \quad (20)$$

The matrix \mathcal{U} is computed from Ω as follows:

- perform the state isomorphism with matrix $\Lambda = P^{1/2}$,
 $(W, X, Y, Z) \longrightarrow (\tilde{W}, \tilde{X}, \tilde{Y}, Z)$:

$$(\tilde{W}, \tilde{X}, \tilde{Y}, Z) = (\Lambda W \Lambda^{-1}, \Lambda X, Y \Lambda^{-1}, Z).$$

- compute a Cholesky factorization of the matrix

$$\begin{bmatrix} K & L \\ L^* & N \end{bmatrix} = \begin{bmatrix} \tilde{W} & \tilde{X} \\ \tilde{Y} & Z \end{bmatrix}^* \begin{bmatrix} \tilde{W} & \tilde{X} \\ \tilde{Y} & Z \end{bmatrix}. \quad (21)$$

using the well-known formula [12, Sec. 0.2],

$$\begin{aligned} \begin{bmatrix} K & L \\ L^* & N \end{bmatrix} &= \begin{bmatrix} I & 0 \\ L^* K^{-1} & I \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} I & K^{-1} L \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} K^{1/2} & K^{-1/2} L \\ 0 & M^{-1/2} \end{bmatrix}^* \begin{bmatrix} K^{1/2} & K^{-1/2} L \\ 0 & M^{-1/2} \end{bmatrix} \end{aligned}$$

where $M^{-1} = N - L^* K^{-1} L$ can be computed by inverting the matrix (21) [12, Formula (0.8)].

We thus define

$$\mathcal{U} = \begin{bmatrix} \tilde{W} & \tilde{X} \\ \tilde{Y} & Z \end{bmatrix} \begin{bmatrix} K^{-1/2} & -K^{-1} L M^{1/2} \\ 0 & M^{1/2} \end{bmatrix}. \quad (22)$$

The matrices L and Z are given by

$$L = \tilde{Y}^* Z + \tilde{W}^* \tilde{X} \quad (23)$$

$$M = Z^* Z + X^* P^{-1} X. \quad (24)$$

Theorem 2 Let $\Omega = (W, X, Y, Z)$ be a unitary realization and \mathcal{D}_Ω the set of unitary realizations in canonical form with respect to Ω (Λ in (12) is positive definite and Hermitian). The map

$$\begin{aligned} \phi_\Omega : \mathcal{D}_\Omega &\longrightarrow \mathbb{R}^{2np} \times \mathbb{U}_p, \\ (A, B, C, D) &\longmapsto (V, D_0) \end{aligned}$$

where $V, \mathcal{V}, \mathcal{U}$ and D_0 are successively computed by (13), (20), (22) and (19), is a chart of \mathbb{L}_n^p .

The map $\phi_\Omega^{-1} : (V, D_0) \longmapsto \mathcal{U} \text{diag}(I, D_0) \mathcal{V}^*$ is a local canonical form.

The family $(\mathcal{D}_\Omega, \phi_\Omega)$, Ω unitary realization forms an atlas of \mathbb{L}_n^p .

Note that $\Omega \in \mathcal{D}_\Omega$ and has parameters $V = 0$ and $D_0 = I$. The chart is centered on G_Ω and is called an *adapted chart* for G_Ω . A parametrization of the quotient space $\mathbb{L}_n^p / \mathbb{U}_p$ is obtained by fixing D_0 within the chart and letting only V vary.

7.2 Illustration

Up to a right constant unitary matrix, any 2×2 lossless matrices of McMillan degree 1 with real coefficients can be written in the form

$$B_{a,\phi} = I + (\zeta_a - 1) \begin{bmatrix} \cos \frac{\phi}{2} \\ \sin \frac{\phi}{2} \end{bmatrix} \begin{bmatrix} \cos \frac{\phi}{2} \\ \sin \frac{\phi}{2} \end{bmatrix}^T, \quad \phi \in [0, 2\pi[,$$

where ζ_a is a normalized Blaschke factor

$$\zeta_a(z) = \nu \frac{1 - \bar{a}z}{z - a}, \quad \begin{cases} \nu = -\frac{a}{|a|}, & 0 < |a| < 1, \\ \nu = 1, & a = 0. \end{cases} \quad (25)$$

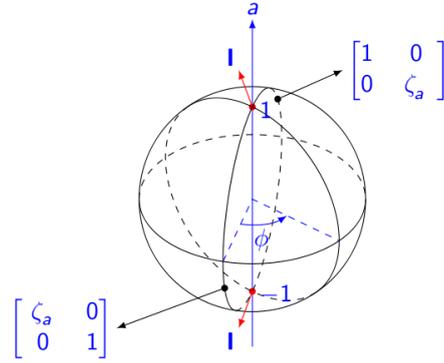


Fig. 3. The quotient space of 2×2 , degree 1, real lossless functions

The dimension of this manifold is 2 and it can be represented by a sphere of which the poles have been excluded (see Picture 3). The poles of the sphere correspond to a drop of degree.

Consider the chart centered at $B_{0,0}(z) = \text{diag}(1/z, 0)$. The parameter V is a 2-vector and $\Lambda = \sqrt{1 - \|V\|^2}$, so that the parameter domain is just the open unit disk ($\|V\| < 1$). The canonical form is

$$V = \begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} -x & 1 - \sigma x x^* & -\sigma x^* y \\ \Lambda & x^* & y^* \\ -y & -\sigma y^* x & 1 - \sigma y y^* \end{bmatrix}.$$

The corresponding lossless function is of the form $B_{a,\phi} U$ with $a = -x$, $\cos \phi/2 = \Lambda$, $\sin \phi/2 = y$ and U a unitary matrix. Since Λ cannot be zero, the matrices on the meridian $\phi = \pi$ are not represented in this chart. They show up on the boundary of the chart except for $x = \pm 1$ and $y = 0$ which correspond to the poles of the sphere. As for the sphere, two charts at least are needed to represent the whole set. Adding the chart centered at $B_{0,\pi}$ we get an atlas.

7.3 Optimization in RARL2.

The software RARL2 is a Matlab based software which performs rational approximation following the principle we just described. It divides into two libraries

- `ar12lib` contains all the computations concerning the L^2 criterion and its gradient. The function can be given by a realization or Fourier coefficients as in our application. In this case, the matrix \hat{B} (see Section 6) is computed by $\sum_{k \geq 0}^{N-1} F_{k+1} C A^k$, where the $(F_k)_{0 \leq k \leq N}$ are the matrix Fourier coefficients.
- `boplib` is concerned with the parametrization of lossless functions (balanced output pairs) by means of the lossless mutual encoding method described in section 7.1. It also provides a minimization process which could handle **any criterion** defined over the manifold $\mathbb{L}_n^p / \mathbb{U}_p$.

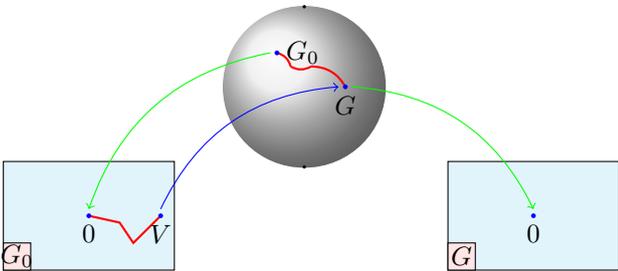


Fig. 4. Optimization over a manifold

The minimization process makes use of the Matlab solver `fmincon`. It starts at some initial point G_0 which is encoded in its adapted chart $\Omega = (A, B, C, D)$ (see section 7.1). Then `fmincon` performs the optimization of the criterion submitted to the nonlinear constraint $P > 0$, where P is the solution to

$$P - A^* P A = C^* C - V^* V$$

and V the parameter of the current point in the chart. This constraint ensures we remain within the domain of the chart. When a constraint violation occurs, a new adapted chart is computed and the optimization pursues within this new chart (see Figure 4) until a minimum is reached.

The convergence of the algorithm has been proved under mild assumptions in the SISO case [6] but never in the MIMO case. The main obstruction to the convergence is if the boundary of the manifold is reached, that is to say if the constraint violation (P singular) corresponds to the non-minimality of the canonical realization, that is to a drop of degree for the lossless functions. This would result in changing chart indefinitely.

7.4 Initialization

Since the criterion may possess many local minima, the choice of an initial point in the optimization process is essential. Projection based model reduction proposes a panel of low-cost computational methods to get a reasonable starting point. In particular, optimal Hankel norm model reduction [15] or balanced truncation [30] presents guarantees of quality and error bound on the result. In RARL2, the balanced truncation method of [23] has been implemented and is used as a starting point for the identification of microwave filters.

8 Results and conclusion.

A long-standing cooperation with the space agency CNES resulted in an original method to extract coupling parameters from frequency scattering measurements, and in two dedicated software programs which are now fully integrated in the design and tuning process. In this paper we have described in detail the identification step performed by the software PRESTO-HF [35] (which includes the rational approximation software RARL2). In Figures 5,6,7 the results of our procedure are shown at hand of a real-life example provided by the CNES. It consists of measurements of a microwave filter of 8th order in 800 frequency points. As shown by the error function in Figure 7, an excellent agreement is obtained between the final rational model and the measurements. The latter is obtained in less than 15 seconds on a Intel Core I7 processor, which makes our approach compatible with a real-time tuning procedure of the filter. The software PRESTO-HF [35] is currently used for this purpose by several of our industrial partners.

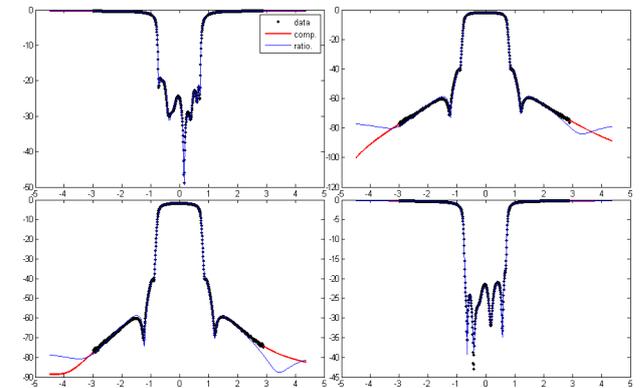


Fig. 5. CNES 2×2 hyperfrequency filter; Bode diagram of data (dots), completion (red) and approximant at order 8 (blue line).

In a second step we extract the coupling parameters from the identified model. This step is performed by the software DEDALE-HF [34]. Based on computer algebra methods, this software computes a realization of the form (1) in which the coupling geometry of M has been specified.

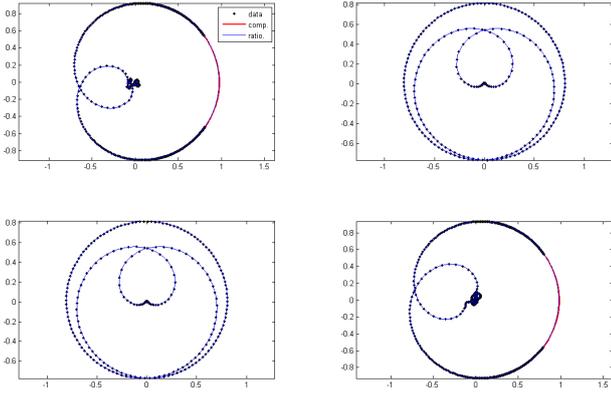


Fig. 6. CNES 2×2 hyperfrequency filter: data (dots), completion (red), and approximant (blue) at order 8 (Nyquist diagram).

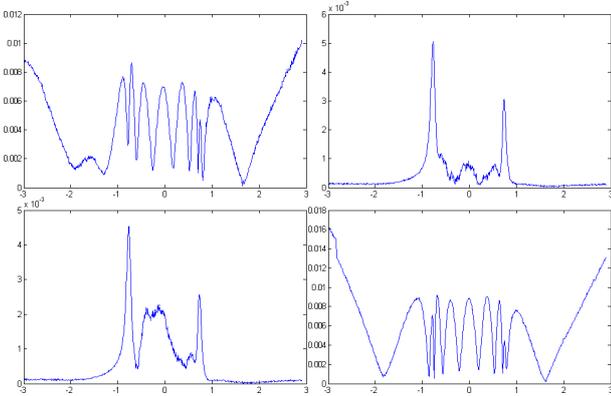


Fig. 7. CNES 2×2 hyperfrequency filter: Magnitude of the point-wise error between rational approximation (including delay components) and measurements for each entry.

The case of output multiplexers (OMUX) where several filters of the previous type are coupled on a common guide has also been considered. The model is obtained upon chaining the corresponding scattering matrices, and mixes up rational elements and complex exponentials (because of the delays). This makes the identification much more challenging. Nowadays, the new trend is to remove the waveguides that are an important part of the mass and bulk, in order to integrate more strongly these devices and to simplify their architecture. The devices obtained in this way (compact OMUX) differ from a single filter by the number of ports m greater than 2. They are described by a scattering matrix of order m . We are currently investigating the possibility to apply our methods to such devices. If the model reduction step directly applies to this case, the completion as well as the extraction of coupling parameters from the model require new developments.

References

[1] R.S. Adve, T.K. Sarkar, S.M. Rao, E.K. Miller, and R. Pflug. Application of the Cauchy method for

extrapolating/interpolating narrow-band system responses. *IEEE Transactions on Microwave Theory and Techniques*, 45:837–845, 1997.

- [2] 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.
- [3] J. D. Aplevich. Gradient methods for optimal linear system reduction. *Int. J. Control*, 18, 1973.
- [4] J.A. Ball, I. Gohberg, and L. Rodman. *Interpolation of rational matrix functions*, volume 45 of *Operator Theory: Advances and Applications*. Birkhäuser, 1990.
- [5] L. Baratchart. Existence and generic properties for L^2 approximants of linear systems. *I.M.A. Journal of Math. Control and Identification*, 3:89–101, 1986.
- [6] L. Baratchart, M. Cardelli, and M. Olivi. Identification and rational L^2 approximation: a gradient algorithm. *Automatica*, 27(2):413–418, 1991.
- [7] L. Baratchart and J. Leblond. Hardy approximation to L^p functions on subsets of the circle with $1 \leq p < \infty$. *Constructive Approximation*, 14:41–56, 1998.
- [8] L. Baratchart, J. Leblond, J. R. Partington, and N. Torkhani. Robust identification from band-limited data. *IEEE Trans. on Autom. Control*, 42(9):1318–1325, 1997.
- [9] L. Baratchart, J. Leblond, and F. Seyfert. Extremal problems of mixed type in H^2 of the circle. Rapport de recherche RR-7087, INRIA, 2009.
- [10] L. Baratchart and M. Olivi. Critical points and error rank in best H^2 matrix rational approximation. *Constructive Approximation*, 14:273–300, 1998.
- [11] P. Van Dooren, K.A. Gallivan, and P.-A. Absil. H^2 -optimal model reduction of MIMO systems. *Applied Mathematics Letters*, 21, 2008.
- [12] H. Dym. *J-contractive matrix functions, reproducing kernel spaces and interpolation*, volume 71 of *CBMS lecture notes*. American Mathematical Society, Rhode Island, 1989.
- [13] P. Fuhrmann and U. Helmke. Homeomorphism between observable pairs and conditioned invariant subspaces. *Systems and Control Letters*, 30:217–223, 1997.
- [14] 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.
- [15] K. Glover. All optimal Hankel norm approximations of linear multivariable systems and their L^∞ -error bounds. *Int. J. Contr.*, 39:1115–1193, 1984.
- [16] S. Gugercin, A.C. Antoulas, and Beattie C. H^2 model reduction for large-scale linear dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30, 2008.
- [17] B. Gustavsen and A. Semlyen. Rational approximation of frequency domain responses by vector fitting. *IEEE Trans. Power Delivery*, 14(3):1052–1061, 1999.
- [18] B. Gustavsen and A. Semlyen. A robust approach for system identification in the frequency domain. *IEEE Trans. Power Delivery*, 19(3):1167–1173, 2004.
- [19] B. Hanzon and J.M. Maciejowski. Constructive algebra methods for the L^2 -problem for stable linear systems. *Automatica*, 32(12), 1996.
- [20] B. Hanzon, M. Olivi, and R.L.M. Peeters. Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm. *Linear Algebra and its Applications*, 2006.

- [21] D.C. Hyland and D.S. Bernstein. The optimal projection equations for model reduction and the relationships among the method of Wilson, Skelton and Moore. *IEEE Trans. Aut. Control*, 29, 1985.
- [22] I. Kollár. Frequency domain system identification toolbox for matlab. Gamax Ltd, Budapest, 2004.
- [23] S.-Y. Kung and D.W. Lin. Optimal hankel-norm model reductions: Multivariable systems. *IEEE trans. on Automatic control*, AC-26(4):832–852, 1981.
- [24] A.G. Lampérez, T.K. Sarkar, and M.S. Palma. Filter model generation from scattering parameters using the Cauchy method. In *European Microwave Conference*, 2002.
- [25] 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.
- [26] J. P. Marmorat and M. Olivi. RARL2: a Matlab based software for H^2 rational approximation. <http://www-sop.inria.fr/apics/RARL2/rarl2.html>, 2004.
- [27] J.-P. Marmorat and M. Olivi. Nudelman interpolation, parametrizations of lossless functions and balanced realizations. *Automatica*, 43:1329–1338, 2007.
- [28] J.-P. Marmorat, M. Olivi, B. Hanzon, and R.L.M. Peeters. Matrix rational H^2 approximation: a state-space approach using Schur parameters. In *Proceedings of the CDC, Las-Vegas, USA*, 2002.
- [29] L. Meier and D. G. Luenberger. Approximation of linear constant systems. *IEEE Trans. Aut. Control*, 12:585–587, 1967.
- [30] B.C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, AC-26:17–32, 1981.
- [31] P. Van Overschee, B. De Moor M. Gevers, and G. Li. *Parametrizations in control, estimation and filtering problems*. Springer-Verlag, 1993.
- [32] J. Partington. *Interpolation, Identification and Sampling*. Oxford University Press, 1997.
- [33] G. Ruckebusch. Sur l’approximation rationnelle des filtres. Technical report, CMA, Ecole polytechnique, 1978.
- [34] F. Seyfert. DEDALE-HF: a Matlab toolbox dedicated to the equivalent network synthesis for microwave filters. <http://www-sop.inria.fr/apics/Dedale/WebPages/>.
- [35] F. Seyfert. PRESTO-HF: a Matlab based toolbox dedicated to the identification problem of low pass coupling parameters of band pass microwave filters, 2004.
- [36] J.T. Spanos, M.H. Milman, and D.L. Mingori. A new algorithm for l^2 optimal model reduction. *Automatica*, 28:897–909, 1992.
- [37] M. Verhaegen and V. Verdult. *Filtering and system identification: a least-square approach*. Cambridge university press, 2007.
- [38] D. Wilson and R. Mishra. Optimal reduction of multivariable systems. *Int. J. Control*, 29(2):267–278, 1979.