FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Reduced Order State-Space Models for 2-D Systems

Diogo Pernes Cunha

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Supervisor: Paula Rocha Malonek

July 31, 2013

Resumo

Nesta dissertação, é feito um estudo sobre realização em espaço de estados de sistemas 2-D, bem como sobre a redução da ordem destes modelos. Numa primeira fase, considerando um sistema 2-D como um sistema 1-D com coeficientes definidos sobre o anel das funções racionais próprias, obtêm-se alguns resultados preliminares sobre uma via nunca antes explorada para a redução da ordem dos modelos. Numa segunda fase, apresenta-se um algoritmo da autoria de C.Beck e J.Doyle para a redução da ordem de modelos 2-D e propõe-se uma abordagem alternativa. As respetivas vantagens e desvantagens de cada uma são ilustradas através de exemplos de aplicação.

Abstract

In this thesis, a study is made on the state-space realization and model order reduction of 2-D systems. In a first stage, considering a 2-D system as a 1-D system with coefficients over the ring of proper rational functions, some preliminary results on model reduction are obtained through a novel approach. In a second stage, an algorithm by C.Beck and J.Doyle for model order reduction of 2-D systems is presented and an alternative approach is proposed. The advantages and drawbacks of each algorithm are highlighted through application examples.

Agradecimentos

À minha orientadora, Professora Paula Rocha Malonek, que foi absolutamente incansável durante todo o meu trabalho. Acredito firmemente que sem a sua orientação, dedicação e inesgotável paciência não teria sido possivel concluir esta tarefa da mesma forma.
Aos meus Amigos, que são um dos meus grandes suportes. Neste contexto, uma palavra especial para os meus também colegas Hugo Cruz, pelo muito que aprendi em cada trabalho que fizemos juntos, João Granja, pelo seu brilhantismo, perfecionismo e humildade contagiantes, e Luís Pires, pela amizade que demonstra ao me aturar muito e em todos os momentos.
À minha família, a quem devo grande parte do que sou.

Ao meu pai.

Diogo Pernes Cunha

"In formal logic, a contradiction is the signal of a defeat; but in the evolution of real knowledge it marks the first step in progress towards victory."

Alfred North Whitehead

Contents

List of Figures

List of Tables

Abbreviations and Notation

BIBO	Bounded-Input, Bounded-Output
FIR	Finite Impulse Response
IIR	Infinite Impulse Response
LMI	Linear Matrix Inequality
MIMO	Multiple-Input, Multiple-Output
MSE	Mean Squared Error
n-D	n-dimensional
PSNR	Peak Signal-to-Noise Ratio
PSVD	Pseudo Singular Value Decomposition
SISO	Single-Input, Single-Output
SVD	Singular Value Decomposition

$K[z]$	Set of polynomials in the variable z with coefficients in K
$K(z)$	Set of rational functions in the variable z with with coefficients in K
$K_p(z)$	Set of proper rational functions in the variable z with coefficients in K
$Q(R)$	Field of fractions of the ring R
$deg(p(z))$	Degree in z of the polynomial $p(z)$
$dim(M)$	Dimension of the square matrix M
$\rho(M)$	Spectral radius of the square matrix M
M^{adj}	Adjugate matrix of the square matrix M
$M^{\dagger}$	Moore-Penrose pseudoinverse of the matrix M
$M > 0$	If M is a hermitian matrix, this means that M is positive definite
$M \geq 0$	If M is a hermitian matrix, this means that M is positive semidefinite
$M > N$	If M and N are hermitian matrices of the same dimension, this means that $M - N > 0$
$M \geq N$	If M and N are hermitian matrices of the same dimension, this means that $M - N \geq 0$
$M(z)^*$	If $M(z)$ is a rational matrix in the variable z, this is the same as $M(z^{-1})^T$
n_1	Dimension of the vertical state vector
n_2	Dimension of the horizontal state vector
Z	$Z := \begin{bmatrix} z_2 I_{n_2} & 0 \\ 0 & z_1 I_{n_1} \end{bmatrix}$
$\bar{U}$	$\bar{U} := \{(z_1, z_2) \in \mathbb{C}^2 : \lvert z_1 \rvert \geq 1, \lvert z_2 \rvert \geq 1\}$
W	$W := \{(z_1, z_2) \in \mathbb{C}^2 : \lvert z_1 \rvert = 1, \lvert z_2 \rvert = 1\}$

Chapter 1

Introduction

Linear systems are rather familiar to any electrical engineer. They usually represent the dynamics of a process whose inputs and outputs evolve along time, so a model with one single independent variable (1-D model) is enough. However, in some situations, it may be necessary to consider two independent variables instead of one. Potential areas of application of this kind of systems are digital image processing, seismic analysis, 2-D encoders and even some sophisticated control systems. In digital image filtering, for instance, two spacial directions (horizontal and vertical) have to be taken into account, thus a model with only one independent variable would clearly not be adequate in this context.

There are two main approaches when dealing with digital filters, namely using the convolution with the corresponding impulse response, or then using state-space representations. The former is actually the most common choice for finite impulse response filters (FIR filters), but it becomes very inconvenient for infinite impulse response filters (IIR filters). It is also important to point out that the computation of the output of the filter becomes much simplified if a state-space representation is used. In fact, the state-space is a recursive model which allows the calculation of the output without requiring the computation of any convolution. The amount of information that has to be kept in memory is also reduced: knowledge of all previous input samples is not needed to calculate the next output sample, since it is completely determined by the current state of the system and the current input sample. These two characteristics of state-space models may be specially relevant in real-time systems, where computational time and memory available are key factors for good performance.

The practical applications and advantages just referred motivate the great amount of research that has been done in past years about 2-D state-space models (see, for instance, [1], [2], [3] and [4]). As might be expected, it turns out that the widely studied 1-D systems theory does not apply directly to the two-dimensional case. In the following, a brief presentation of 2-D systems will be made and the main differences will be highlighted.

One of the crucial issues in the context of state-space modelling is the number of state variables

which are necessary to represent a given input-output relation. Models with a large number of state variables occur in many practical applications. If the model is obtained experimentally, via system identification algorithms, this may happen due to noisy measures. If the model is rigorous, this happens simply because the world we live in is far from being simple. In any case, an approximate representation of the system with a reduced number of state variables (reduced order model) would often be more desirable. With small order models, faster simulations are performed, less memory is required, implementation on a chip is easier, among other advantages. For the one-dimensional case there are rather efficient and relatively simple methods for model order reduction. However, for 2-D systems the problem becomes more complicated, although some algorithms have been proposed over the past few years (see [5], [6], [7] and [8], among others).

The purpose of this thesis is investigating the existence and details of a possible new method to solve this problem. In order to state this purpose in a more concrete way, some preliminary notions and results from 1-D systems theory need to be reviewed. This is done in Chapter 2. Chapter 3 contains an introduction to 2-D systems, which includes a presentation of a particular two-dimensional state-space model, some notions of realization theory and on the stability of this kind of systems. Chapter 4 includes the main results of this thesis. After a brief overview of existing approaches to 2-D model order reduction, some preliminary results are derived using a completely non-explored via, by regarding 2-D systems as 1-D systems with coefficients over the ring of 1-D proper rational functions. Then, an existing 2-D model order reduction algorithm is presented in detail and a new alternative approach is developed. The chapter is concluded with the application of both algorithms to some toy examples. Finally, in Chapter 5, the order reduction approaches discussed before are applied in two case studies with practical relevance.

Chapter 2

Model order reduction for 1-D systems: a brief overview

2.1 Preliminaries

A discrete-time 1-D state-space model is given by the equations:

$$x(t+1) = Ax(t) + Bu(t) \tag{2.1}$$

$$y(t) = Cx(t) + Du(t), \tag{2.2}$$

where $t \in \mathbb{Z}_0^+$, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^p$ is the input vector, $y(t) \in \mathbb{R}^m$ is the output vector, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{m\times n}$ and $D \in \mathbb{R}^{m\times p}$. The *dimension* or *order* of the model is defined by the number of state components, i.e., it corresponds to the value of n.
The output of the model, starting from the initial state $x(0) = 0_n$, when the input is

$$u_l(t) = \delta(t) := \begin{cases} 1, & t = 0 \\ 0, & t > 0 \end{cases}, \; l = 1, ..., p, \tag{2.3}$$

is called the *impulse response* of the state-space model. It is straightforward to check that the impulse response, $h_\Sigma(t)$, of the state-space model $\Sigma = (A, B, C, D)$ is given by:

$$h_\Sigma(t) = \begin{cases} D, & t = 0 \\ CA^{t-1}B, & t > 0 \end{cases}. \tag{2.4}$$

The formal power series

$$G_\Sigma(z) := \sum_{t=0}^{\infty} h_\Sigma(t) z^{-t} \tag{2.5}$$

is the $m \times p$ model *transfer matrix*. By using (2.4) and (2.5), one can show that:

$$G_\Sigma(z) = C(zI - A)^{-1}B + D. \tag{2.6}$$

A relevant property of any model transfer matrix is the fact that all of its entries are *proper* rational functions, i.e., rational functions in which the degree of the numerator polynomial is not larger than the degree of the denominator polynomial (for a proof, see [9]). If the equality holds, the rational function is said to be *biproper*; if it doesn't, the rational function is called *strictly proper*. By abuse of terminology, if all entries of a transfer matrix are proper rational functions, that transfer matrix is itself said to be proper.
Defining $u(z)$ and $y(z)$ as

$$u(z) = \sum_{t=0}^{\infty} u(t)z^{-t}, \tag{2.7}$$

$$y(z) = \sum_{t=0}^{\infty} u(t)z^{-t}, \tag{2.8}$$

it can be shown that, assuming $x(0) = 0_n$,

$$y(z) = G_\Sigma(z)u(z). \tag{2.9}$$

A state-space model $\Sigma = (A,B,C,D)$ is said to be a *realization* of a transfer matrix $G(z)$ if the model transfer matrix, $G_\Sigma(z)$, coincides with $G(z)$, i.e., $G_\Sigma(z) = G(z)$. In a time domain formulation, $\Sigma = (A,B,C,D)$ is said to be a *realization* of an impulse response, $h(t)$, if the model impulse response, $h_\Sigma(t)$, coincides with $h(t)$, i.e., $h_\Sigma(t) = h(t)$. In either case, this means that $\Sigma = (A,B,C,D)$ is a realization of an input-output system if its input-output dynamics with zero initial state coincides with the one of the given system.
As it is well known, the same transfer matrix (or impulse response) may be realized by state-space models of different orders. However, as pointed out in the introduction, realizations with the smallest possible number of state components, i.e., *minimal realizations*, are always preferable. So, the question of how to obtain such realizations arises.
Before going into that discussion, it is useful to recall the notions of controllability, observability and stability.

Definition The system (2.1)-(2.2) or the pair (A,B) is said to be *controllable* if, for any initial state x_0 and any final state x^*, there exists a finite sequence of input samples, $u(0), u(1), \cdots, u(t^*-1)$, that transfers the system from x_0 to x^*, i.e., such that $x(0) = x_0$ and $x(t^*) = x^*$. Otherwise, the system is said to be *uncontrollable*.

Definition The system (2.1)-(2.2) or the pair (A,C) is said to be *observable* if, for zero inputs, any unknown initial state x_0 can be determined from a finite sequence of observations of the output, $y(0), y(1), \cdots, y(t^*-1)$. Otherwise, the system is said to be *unobservable*.

Definition The system (2.1)-(2.2) or the matrix A is said to be *stable* if, for zero input and for any initial condition x_0, the infinite sequence $x(0), x(1), x(2), \cdots$ converges to the zero state.

Controllability and observability of a system can be checked through the following theorems.

Theorem 2.1.1. *The following conditions are necessary and sufficient for a model $\Sigma = (A,B,C,D)$ to be controllable:*

1. *rank$(\mathscr{C}) = n$, where $\mathscr{C} = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$;*
2. *rank $\begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for every eigenvalue λ of A.*

Moreover, if all eigenvalues of A have modulus less than one, then Σ is controllable if and only the unique solution P of the Lyapunov equation

$$P = APA^T + BB^T \tag{2.10}$$

is positive definite. Such solution is called the controllability gramian *of the model.*

Theorem 2.1.2. *The following conditions are necessary and sufficient for a model $\Sigma = (A,B,C,D)$ to be observable:*

1. *rank$(\mathscr{O}) = n$, where $\mathscr{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$;*
2. *rank $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for every eigenvalue λ of A.*

Moreover, if all eigenvalues of A have modulus less than one, then Σ is observable if and only if the unique solution Q of the Lyapunov equation

$$Q = A^T QA + C^T C \tag{2.11}$$

is positive definite. Such solution is called the observability gramian *of the model.*

Theorem 2.1.3. *A necessary and sufficient condition for a model $\Sigma = (A,B,C,D)$ to be stable is that all eigenvalues of A have modulus less than one.*

Proof. A proof for these theorems can be found in [9]. □

2.1.1 Minimality

In this subsection, the problem of minimality of the number of state components will be treated. The following definition plays a major role in this discussion.

Definition $\Sigma_1 = (A_1, B_1, C_1, D_1)$ and $\Sigma_2 = (A_2, B_2, C_2, D_2)$ are *equivalent state-space models* if there exists a non-singular matrix $T \in \mathbb{R}^{n \times n}$ such that:

$$A_2 = T^{-1} A_1 T, \;\; B_2 = T^{-1} B_1, \;\; C_2 = C_1 T, \;\; D_2 = D_1. \tag{2.12}$$

Theorem 2.1.4. *Any two equivalent state-space models define the same input-output dynamics for zero initial state.*

Proof. Use (2.6) to check that Σ_1 and Σ_2 yield the same transfer matrix. □

Remark Note that the converse of this theorem does not hold even for models of the same dimension, i.e., there may exist non-equivalent models (in the sense defined above) that define the same input-output dynamics for zero initial state.

Given a system transfer matrix $G(z)$, the problem of minimal realization consists, then, in finding a state-space model $\Sigma_m = (A_m, B_m, C_m, D_m)$ with minimum dimension, n_{min}, such that $G(z) = C_m(zI - A_m)^{-1} B_m + D_m$. That n_{min} is known as the *McMillan degree* of the system.

Theorem 2.1.5. *If a system has a single input and a single output (SISO system) and its transfer function $g(z)$ is written as $g(z) = n(z)/d(z)$, where $n(z)$ and $d(z)$ are coprime polynomials, then $n_{min} = deg(d(z))$.*

Proof. It is firstly proved that a realization of $g(z)$ cannot have a dimension less than $deg(d(z))$ (assuming $d(z)$ and $n(z)$ as coprime). Note that $g(z)$ can be written as

$$g(z) = \frac{1}{det(zI - A)} C(zI - A)^{adj} B + D, \tag{2.13}$$

where ^{adj} stands for the adjugate matrix. Since $det(zI - A)$ is nothing but the characteristic polynomial of A, one immediately concludes that its degree equals the size of A, $dim(A)$. Observing that the entries of $(zI - A)^{adj}$ are all polynomials in z (i.e., there are no rational functions), one concludes that:

$$det(zI - A) = d(z)p(z), \tag{2.14}$$

where $p(z)$ is a nonzero polynomial. Thus,

$$dim(A) = deg(d(z)) + deg(p(z)), \tag{2.15}$$

so $dim(A) \geq deg(d(z))$.
The rest of the theorem is proved constructively. Decompose $g(z)$ as $g(z) = \tilde{n}(z)/d(z) + k$, where $\tilde{n}(z)$ and $d(z)$ are still coprime, k is a constant and $deg(\tilde{n}(z)) < deg(d(z))$. Clearly, $\tilde{n}(z)$ and $d(z)$ can be written as follows:

$$d(z) = z^m + \alpha_{m-1} z^{m-1} + \cdots + \alpha_0, \tag{2.16}$$

$$\tilde{n}(z) = \beta_{m-1} z^{m-1} + \beta_{m-2} z^{m-2} + \cdots + \beta_0. \tag{2.17}$$

It can easily be checked that the following matrices A, B, C and D yield a realization for $g(z) = n(z)/d(z)$.

$$A = \begin{bmatrix} -\alpha_{m-1} & -\alpha_{m-2} & \cdots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.18}$$

$$C = \begin{bmatrix} \beta_{m-1} & \beta_{m-2} & \cdots & \beta_1 & \beta_0 \end{bmatrix}, \quad D = k.$$

Hence, a realization where $dim(A) = deg(d(z))$ was found, so $n_{min} = deg(d(z))$. □

For systems with multiple inputs and multiple outputs (MIMO systems), the following theorem holds.

Theorem 2.1.6. *The McMillan degree, n_{min}, of a transfer matrix $G(z)$ equals the degree of the least common denominator of all minors of $G(z)$.*

Proof. See [9]. □

By using Theorem 2.1.5 and the construction given in its proof, one can obtain a minimal realization for a transfer function of a SISO system. However, for a given realization, the question of how to check its minimality without computing its transfer matrix remains. Theorem 2.1.7 answers this question.

Theorem 2.1.7. *A realization $\Sigma = (A,B,C,D)$ is minimal if and only if (A,B) is controllable and (A,C) is observable. Furthermore, all minimal realizations of a system are equivalent.*

Proof. See [9]. □

Remark Besides giving a rather easy way of checking minimality of a realization, Theorem 2.1.7 implicitly tells that it is possible to obtain a minimal realization, from a non-minimal one, just by restricting to controllable and observable states. This can be done applying Kalman decompositions. For a detailed explanation of this procedure see [10].

2.1.2 Balanced realizations

The controllability and observability gramians, introduced in Theorems 2.1.1 and 2.1.2, are going to be very useful for the definition of balanced realizations. The following results provide explicit forms for the gramians.

Theorem 2.1.8. *The controllability gramian of a stable realization $\Sigma = (A,B,C,D)$ is given by*

$$P = \mathscr{C}_\infty \mathscr{C}_\infty^T, \tag{2.19}$$

where $\mathscr{C}_\infty = \begin{bmatrix} B & AB & \cdots & A^iB & \cdots \end{bmatrix}$.

Proof. By definition, P is the unique matrix that satisfies:

$$P = APA^T + BB^T. \tag{2.20}$$

If A is stable, then the series $\sum_{i=0}^{\infty} A^i BB^T (A^T)^i = \mathscr{C}_\infty \mathscr{C}_\infty^T$ converges. Moreover, it is a solution of (2.10), as

$$A\sum_{i=0}^{\infty} A^i BB^T (A^T)^i A^T + BB^T = A^0 BB^T (A^T)^0 + \sum_{i=1}^{\infty} A^i BB^T (A^T)^i = \sum_{i=0}^{\infty} A^i BB^T (A^T)^i. \tag{2.21}$$

This proves that $P = \mathscr{C}_\infty \mathscr{C}_\infty^T$. □

Theorem 2.1.9. *The observability gramian of a stable realization $\Sigma = (A,B,C,D)$ is given by*

$$Q = \mathscr{O}_\infty^T \mathscr{O}_\infty, \tag{2.22}$$

where $\mathscr{O}_\infty = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^i \\ \vdots \end{bmatrix}$.

Proof. The proof for this theorem is analogous to the proof of Theorem 2.1.8. □

As shown next, these gramians give a measure of "how controllable" and "how observable" a given state is. To simplify the discussion, assume that the system is a SISO one.
Consider the singular value decomposition (SVD) of P, the controllability gramian of a given controllable realization:

$$P = V_p S_p V_p^T, \tag{2.23}$$

where $S_p = diag(\sigma_{p_1}, \sigma_{p_2}, \cdots, \sigma_{p_n})$, $\sigma_{p_1} \geq \cdots \geq \sigma_{p_n} > 0$, and $V_p = \begin{bmatrix} v_{p_1} & v_{p_2} & \cdots & v_{p_n} \end{bmatrix}$ is an orthogonal matrix whose columns are the v_{p_i}, $i = 1,2,\cdots,n$. Note that the right and left singular vectors are equal because P is a symmetric matrix. Recall also that P is nonsingular, since it is positive definite.
Now suppose that the system is transferred, in finite time t^*, from the initial zero state to a final state x^*. This state can be written as a linear combination of the columns of V_p (which form an orthonormal basis of $\mathbb{R}^n$):

$$x^* = \alpha_1 v_{p_1} + \alpha_2 v_{p_2} + ... + \alpha_n v_{p_n} = V_p \alpha, \tag{2.24}$$

where $\alpha := \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix}^T$. It can easily be shown that, starting from the origin, the state of the system at time t will be

$$x(t) = \mathscr{C}_t \mathscr{U}(t), \tag{2.25}$$

where $\mathscr{C}_t := \begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix}$ and $\mathscr{U}(t) := \begin{bmatrix} u(t-1) & u(t-2) & \cdots & u(0) \end{bmatrix}^T$. Note that, if $t > n$, equation (2.25) corresponds to an underdetermined system of equations where the unknown is the input sequence $\mathscr{U}(t)$.

If one accepts to reach the state x^* in an arbitrarily large amount of time, then $x^* = x(\infty) = \mathscr{C}_\infty \mathscr{U}(\infty)$, where $\mathscr{C}_\infty$ and $\mathscr{U}(\infty)$ are, respectively, matrices $\mathscr{C}_t$ and $\mathscr{U}(t)$ when t is extended to infinity.

A way of measuring the "effort" needed to transfer the system from the origin to x^* is computing the sum of the squares of the corresponding input signal, $\sum_{t=0}^{\infty} u^2(t)$. This sum is commonly designated as the energy of the signal $u(t)$ and is here denoted by E_u. For this matter, the most interesting solution of (2.25) is the one that corresponds to a minimum input energy. It can be shown (see [11]) that the minimum 2-norm solution of (2.25) is

$$\mathscr{U}(t) = \mathscr{C}_t^T (\mathscr{C}_t \mathscr{C}_t^T)^{-1} x(t). \tag{2.26}$$

Thus, the minimum energy required to transfer the system from the initial zero state to a final state x^* is

$$\begin{aligned} E_u|_{min} &= \sum_{t=0}^{\infty} u^2(t) = \mathscr{U}(\infty)^T \mathscr{U}(\infty) = (\mathscr{C}_\infty^T (\mathscr{C}_\infty \mathscr{C}_\infty^T)^{-1} x^*)^T \mathscr{C}_\infty^T (\mathscr{C}_\infty \mathscr{C}_\infty^T)^{-1} x^* = \\ &= x^{*T} (\mathscr{C}_\infty \mathscr{C}_\infty^T)^{-1} x^* = x^{*T} P^{-1} x^*. \end{aligned} \tag{2.27}$$

Replacing x^* by the linear combination defined above and P by its SVD, gives

$$E_u|_{min} = (V_p \alpha)^T (V_p S_p V_p^T)^{-1} V_p \alpha = \alpha^T S_p^{-1} \alpha = \frac{\alpha_1^2}{\sigma_{p_1}} + \frac{\alpha_2^2}{\sigma_{p_2}} + \cdots + \frac{\alpha_n^2}{\sigma_{p_n}}. \tag{2.28}$$

Thus, states aligned with the first singular vectors of P will require a smaller energy to be reached than those which are aligned with the last singular vectors. For instance, a final state $x_1^* = u_{p_1}$ will require an energy of $1/\sigma_{p_1}$ to be reached, which is obviously smaller than $1/\sigma_{p_n}$, the energy needed to reach the state $x_2^* = u_{p_n}$.

Through an analogous procedure, one can show that the energy of the output signal $y(t)$, when the system is released from an initial state $x_0 = \beta_1 v_{q_1} + \beta_2 v_{q_2} + \cdots + \beta_n v_{q_n}$, is given by

$$E_y = \beta_1^2 \sigma_{q_1} + \beta_2^2 \sigma_{q_2} + \ldots + \beta_n^2 \sigma_{q_n}, \tag{2.29}$$

where the σ_{q_i} and the v_{q_i}, $i = 1, 2, \cdots, n$, are, respectively, the singular values and the singular vectors of Q, the observability gramian of the system, which is assumed to be observable. Now, states aligned with the first singular vectors of Q will provide more output energy than those aligned with the last singular vectors.

One of the key ideas followed by model reduction, treated in the next subsection, is that states that are difficult to achieve (i.e., states that need a big control energy to be reached) and that provide a small output energy can be discarded. Thus, it would be useful to have a realization in which

hardly reachable states are also hardly observable states and vice-versa. This corresponds exactly to a *balanced realization*, defined below.

Definition A stable and minimal realization $\Sigma = (A,B,C,D)$ is *balanced* if the corresponding gramians P and Q are such that:

$$P = Q = S = diag(\sigma_1, \sigma_2, \cdots, \sigma_n), \;\; \sigma_1 \geq \cdots \geq \sigma_n > 0 \tag{2.30}$$

The numbers σ_i, $i = 1, 2, \cdots, n$, are the *singular values of the system*.

The following algorithm to obtain a balanced realization is proposed and proved in [12].

Algorithm 1. *Balanced realization*
Input data: $\Sigma = (A,B,C,D)$ *minimal and stable*

1. *Compute the gramians P and Q by solving the Lyapunov equations (2.10) and (2.11)*
2. *Compute the SVD of Q,*
$$Q = US_qU^T. \tag{2.31}$$
3. *Compute the SVD of* $US_q^{1/2}PS_q^{1/2}U$,
$$US_q^{1/2}PS_q^{1/2}U = VS_pV^T. \tag{2.32}$$
4. *Apply the similarity transformation T to* $\Sigma = (A,B,C,D)$, *as defined in (2.12), where*
$$T = US_q^{-1/2}VS_p^{1/4}. \tag{2.33}$$

The obtained realization $\Sigma_b = (A_b, B_b, C_b, D)$ *is a balanced one.*

2.2 Common approaches for model order reduction

In the following, two methods for obtaining an approximate and reduced order system will be discussed. The first one starts from a balanced realization and reduces the order of the model by discarding the least controllable and the least observable modes (in the sense discussed in the previous section). The second one follows a completely different approach, based on a low-rank approximation of a Hankel matrix.

2.2.1 Model order reduction via balanced truncation

Consider a balanced realization $\Sigma = (A,B,C,D)$ and assume that the controllability and observability gramians are partitioned as:

$$P = Q = S = \begin{bmatrix} S_r & 0 \\ 0 & S_{n-r} \end{bmatrix}, \tag{2.34}$$

$$S_r = diag(\sigma_1, \cdots, \sigma_r), \tag{2.35}$$

$$S_{n-r} = diag(\sigma_{r+1}, \cdots, \sigma_n), \tag{2.36}$$

with $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} \geq \cdots \geq \sigma_n > 0$. The corresponding (A,B,C) can be partitioned accordingly as:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \; B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \; C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \tag{2.37}$$

where $A_{11} \in \mathbb{R}^{r\times r}$, $A_{12} \in \mathbb{R}^{r\times(n-r)}$, $A_{21} \in \mathbb{R}^{(n-r)\times r}$, $A_{22} \in \mathbb{R}^{(n-r)\times(n-r)}$, $B_1 \in \mathbb{R}^{r\times p}$, $B_2 \in \mathbb{R}^{(n-r)\times p}$, $C_1 \in \mathbb{R}^{m\times r}$ and $C_2 \in \mathbb{R}^{m\times(n-r)}$. Roughly speaking, if σ_{r+1} is sufficiently small, then the input-output system realized by $\Sigma_r = (A_{11}, B_1, C_1, D)$ will be a good approximation of the original one, in the sense that the corresponding transfer matrices are close. The following theorem states this affirmation rigorously.

Theorem 2.2.1. *If $\Sigma = (A,B,C,D)$ is a balanced realization with transfer matrix $G(z)$, then the transfer matrix $G_r(z) = C_1(zI - A_{11})^{-1}B_1 + D$, where A_{11}, B_1 and C_1 are as defined in (2.37), has the following properties:*

1. *It is stable;*
2. $\|G(z) - G_r(z)\|_{\mathscr{H}_\infty} := \max_{\forall\omega\in\mathbb{R}} \|G(e^{j\omega}) - G_r(e^{j\omega})\|_2 \leq 2(\sigma_{r+1} + \cdots + \sigma_n)$;
3. *$\Sigma_r = (A_{11}, B_1, C_1, D)$ is a minimal realization of $G_r(z)$.*

Proof. See [10] □

2.2.2 Model order reduction via Hankel matrix truncation

Consider the impulse response, $h(t)$, of a system realized by $\Sigma = (A,B,C,D)$. Using $h(t)$, the following matrix can be built:

$$H_{\alpha,\beta} = \begin{bmatrix} h(1) & h(2) & \cdots & h(\beta) \\ h(2) & h(3) & \cdots & h(\beta+1) \\ \vdots & \vdots & \ddots & \vdots \\ h(\alpha) & h(\alpha+1) & \cdots & h(\alpha+\beta-1) \end{bmatrix} = \begin{bmatrix} CB & CAB & \cdots & CA^{\beta-1}B \\ CAB & CA^2B & \cdots & CA^{\beta}B \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-1}B & CA^{\alpha}B & \cdots & CA^{\alpha+\beta-2}B \end{bmatrix}. \tag{2.38}$$

It has the structure of a *Hankel matrix*. If $\alpha = \beta = k$, $H_{\alpha,\beta}$ is simply denoted by H_k. The notation H_∞ will be used to denote that matrix when k is extended to infinity and $rank(H_\infty) := \sup_k rank(H_k)$.

Theorem 2.2.2. *The McMillan degree, n_{min}, of a system with impulse response $h(t)$ equals $rank(H_\infty)$.*

Proof. The proof will be only given for SISO systems.
Firstly, it is proved that $rank(H_\infty) \leq n_{min}$. Suppose that $\Sigma = (A,B,C,D)$ is a minimal realization. Then, by the Cayley-Hamilton theorem,

$$A^{n_{min}} = \alpha_0 I + \alpha_1 A + ... + \alpha_{n_{min}-1} A^{n_{min}-1}. \tag{2.39}$$

Thus,

$$CA^i B = \alpha_0 CA^{i-n_{min}} B + \alpha_1 CA^{i+1-n_{min}} B + ... + \alpha_{n_{min}-1} CA^{i-1} B, \;\; i = n_{min}+1, n_{min}+2, \cdots \tag{2.40}$$

Consider the matrix H_k, where $k > n_{min}$ is arbitrarily large. Equation (2.40) shows that the $(n_{min}+1)$-th column of H_k can be written as a linear combination of the first n_{min} columns. The $(n_{min}+2)$-th column of H_k can be written as a linear combination of the previous n_{min} columns and so on. Hence, one concludes that, from the $(n_{min}+1)$-th on, all columns can be written as a linear combination of the first n_{min} columns, therefore $rank(H_\infty) \leq n_{min}$.
Now suppose that $rank(H_\infty) < n_{min}$. Then, there exists $n^* \leq n_{min}$ such that the n^*-th column of H_k could be written as a linear combination of the first $n^* - 1$ columns, that is:

$$h(t) = \beta_1 h(t-1) + \beta_2 h(t-2) + \cdots + \beta_{n_{min}-1} h(t - n_{min} + 1), \;\; t = n_{min}, n_{min}+1, \cdots \tag{2.41}$$

This recursion implies that the corresponding system is described by a difference equation like the following:

$$y(t) = \beta_1 y(t-1) + \beta_2 y(t-2) + \cdots + \beta_{n_{min}-1} y(t - n^* + 1) + f(u(t), u(t-1), \cdots, u(t-n^*+1)), \tag{2.42}$$

where f is a linear function. As can be readily verified, the McMillan degree of a system described by (2.42) would be, at most, $n^* - 1 \leq n_{min} - 1$. Since the McMillan degree is n_{min}, it is impossible that $rank(H_\infty) < n_{min}$, thus $rank(H_\infty) = n_{min}$. □

As it should be clear now, $H_{n_{min}+1}$ and D completely define the system. Hence, it must be possible to use them in order to find a minimal realization. Start by noting that:

$$H_{\alpha,\beta} = \mathscr{O}_\alpha \mathscr{C}_\beta, \tag{2.43}$$

where, if $\alpha, \beta > n_{min}$, $\mathscr{O}_\alpha$ and $\mathscr{C}_\beta$ are, respectively, the extended observability and controllability matrices of some realization of the system. In particular, if $H_{n_{min}+1} = \mathscr{O}_{n_{min}+1} \mathscr{C}_{n_{min}+1}$. with $\mathscr{O}_{n_{min}+1} \in \mathbb{R}^{m(n_{min}+1) \times n_{min}}$ and $\mathscr{C}_{n_{min}+1} \in \mathbb{R}^{n_{min} \times p(n_{min}+1)}$, these matrices correspond to a minimal realization. Thus, matrices B and C of that minimal realization can be immediately obtained by extracting, respectively, the first p columns of $\mathscr{C}_{n_{min}+1}$ and the first m rows of $\mathscr{O}_{n_{min}+1}$.

To find A, note that:

$$\begin{bmatrix} AB & A^2B & \cdots & A^{n_{min}}B \end{bmatrix} = A\mathscr{C}_{n_{min}}, \tag{2.44}$$

$$\begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^{n_{min}} \end{bmatrix} = \mathscr{O}_{n_{min}}A. \tag{2.45}$$

Thus, A can be found by any of the following formulas:

$$A = |\mathscr{C}_{n_{min}+1}\,\mathscr{C}^{\dagger}_{n_{min}}, \tag{2.46}$$

$$A = \mathscr{O}^{\dagger}_{n_{min}}\bar{\mathscr{O}}_{n_{min}+1}, \tag{2.47}$$

where $|\mathscr{C}_{n_{min}+1}$ is formed by the last pn_{min} columns of $\mathscr{C}_{n_{min}+1}$ and $\bar{\mathscr{O}}_{n_{min}+1}$ is formed by the last mn_{min} rows of $\mathscr{O}_{n_{min}+1}$. The symbol † stands for the Moore-Penrose pseudo-inverse. The only question that remains unanswered is how to obtain matrices $\mathscr{O}_{n_{min}+1}$ and $\mathscr{C}_{n_{min}+1}$. Indeed, any factorization of $H_{n_{min}+1}$ such that both matrices have full column and row rank, respectively, is valid. This can be obtained, for example, through the SVD of $H_{n_{min}+1}$.
The following algorithm, known as the *Ho-Kalman algorithm*, describes a way of obtaining a minimal realization given the impulse response of the system. The proof of this fact will not be explicitly given here, because it follows exactly the same reasoning used above.

Algorithm 2. *Ho-Kalman algorithm*
Input data: Impulse response of the system, $h(t)$

1. *Build the matrix $H_{\alpha,\beta}$, as defined in (2.38), for some arbitrary α and β*

2. *Compute the SVD of $H_{\alpha,\beta}$, which should look like the following:*

$$H_{\alpha,\beta} = \begin{bmatrix} U_{n_{min}} & \bar{U}_{n_{min}} \end{bmatrix} \begin{bmatrix} S_{n_{min}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V^T_{n_{min}} \\ \bar{V}^T_{n_{min}} \end{bmatrix} = U_{n_{min}} S_{n_{min}} V^T_{n_{min}}, \tag{2.48}$$

   where $S_{n_{min}} = diag(\sigma_1, \cdots, \sigma_{n_{min}})$, $\sigma_i \neq 0$, $i = 1, \cdots, n_{min}$. If the zero matrices above have null dimension, go back to 1. and consider larger values for α and β.

3. *Calculate the extended observability and controllability matrices as:*

$$\mathscr{O}_{\alpha} = U_{n_{min}} S^{1/2}_{n_{min}}, \quad \mathscr{C}_{\beta} = S^{1/2}_{n_{min}} V^T_{n_{min}}. \tag{2.49}$$

4. *Obtain A, B, C, and D using:*

$$A = |\mathscr{C}_\beta \, \mathscr{C}^\dagger_{\beta-1} \tag{2.50}$$

$$B = \textit{first } p \textit{ columns of } \mathscr{C}_\beta \tag{2.51}$$

$$C = \textit{first } m \textit{ rows of } \mathscr{O}_\alpha \tag{2.52}$$

$$D = h(0), \tag{2.53}$$

where $|\mathscr{C}_\beta$ is formed by the last $p(\beta-1)$ lines of $\mathscr{O}_\alpha$. $\Sigma = (A,B,C,D)$ is a minimal realization for the system whose impulse response is $h(t)$.

The realization given by this algorithm is obviously exact. However, with a slight modification one can obtain a reduced order realization. In order to do that, the following theorem will be useful.

Theorem 2.2.3. *Let $H \in \mathbb{R}^{k\times q}$ be a full rank matrix and $H = USV^T$ its SVD, where*

$$S = \begin{cases} diag(\sigma_1, \cdots, \sigma_r, \sigma_{r+1}, \cdots, \sigma_k), \text{ if } k \leq q \\ diag(\sigma_1, \cdots, \sigma_r, \sigma_{r+1}, \cdots, \sigma_q), \text{ if } k \geq q \end{cases} \quad \textit{(note that S is not necessarily square)}$$

The $k \times q$ matrix $\tilde{H} = US_rV^T$, where $S_r = diag(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0)$, has the following properties:

1. *It has rank r;*
2. *It minimizes $\|H - X\|_2$ for all possible matrices X whose rank is not larger than r;*
3. *$\|H - \tilde{H}\|_2 = \sigma_{r+1}$.*

$\tilde{H}$ is said to be the low-rank approximation *of H of order r.*

Proof. See [12] □

Thus, to obtain a reduced order system, one can simply compute a low-rank approximation of $H_{\alpha,\beta}$ of the desired order and then find a realization for the new matrix $\tilde{H}_{\alpha,\beta}$, by using Ho-Kalman or any other algorithm. It is important to point out that this procedure is possible because the low-rank approximation preserves the Hankel matrix structure, i.e., $\tilde{H}_{\alpha,\beta}$ is also a Hankel matrix.

Chapter 3

2-D Systems: Roesser's state-space model

In the two-dimensional case, a state-space model may be defined by the equations:

$$\begin{bmatrix} x_h(i+1,j) \\ x_v(i,j+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ x_v(i,j) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(i,j) \tag{3.1}$$

$$y(i,j) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ x_v(i,j) \end{bmatrix} + Du(i,j), \tag{3.2}$$

where $(i,j) \in \mathbb{Z}_0^{+^2}$, $x(i,j) = \begin{bmatrix} x_h(i,j)^T & x_v(i,j)^T \end{bmatrix}^T \in \mathbb{R}^n$ is the state vector, $u(i,j) \in \mathbb{R}^p$ is the input vector, $y(i,j) \in \mathbb{R}^m$ is the output vector and all matrices are real. Once again, the dimension or order of the system is defined as the dimension of the state vector, n. This model was presented by Robert P. Roesser in [2], so it is usually designated as the Roesser model. Although this is apparently the most used when dealing with 2-D systems, other state-space models exist. For completeness, another widely used model, proposed by Fornasini and Marchesini in [3], is given here:

$$x(i+1,j+1) = A_1 x(i+1,j) + A_2 x(i,j+1) + B_1 u(i+1,j) + B_2 u(i,j+1) \tag{3.3}$$

$$y(i,j) = Cx(i,j) + Du(i,j). \tag{3.4}$$

Models (3.1)-(3.2) and (3.3)-(3.4) have a common updating structure in the sense that the state at each grid point $(i+1,j+1)$ is computed from the states and inputs at the previous nearest neighbours $(i+1,j)$, $(i,j+1)$. In fact, these models are equivalent, since one can be rewritten in the form of the other by possibly redefining the state variables (see [3]). As stated before, Roesser's is the most popular model, so, from now on, only this will be considered.

Note that, contrarily to what happens in the one-dimensional model, the initial state $x(0,0) = [x_h(0,0)^T, x_v(0,0)^T]^T$ is not enough to define the evolution of the system for a given input. Knowledge of $x_h(0,j)$ and $x_v(i,0)$, for all $i,j \geq 0$, is required to compute the evolution of the state vector

and, consequently, the evolution of the output.
The impulse response, $h_\Sigma(i,j)$, of a 2-D state-space model $\Sigma = (A_{11},A_{12},A_{21},A_{22},B_1,B_2,C_1,C_2,D)$ is defined as the output of the system when the input

$$u_l(i,j) = \delta(i,j) := \begin{cases} 1, & i=j=0 \\ 0, & \text{otherwise} \end{cases} \quad , l = 1,...,p, \tag{3.5}$$

is applied and zero initial conditions are considered. The 2-D state-space model transfer matrix is the 2-D formal power series

$$G_\Sigma(z_1,z_2) := \sum_{i=0,j=0}^{\infty,\infty} h_\Sigma(i,j) z_1^{-j} z_2^{-i}. \tag{3.6}$$

It can also be expressed as:

$$G_\Sigma(z_1,z_2) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} z_2 I - A_{11} & -A_{12} \\ -A_{21} & z_1 I - A_{22} \end{bmatrix}^{-1} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + D. \tag{3.7}$$

Doing some algebraic manipulation, it is possible to show that:

$$\begin{aligned} G_\Sigma(z_1,z_2) = (C_1 + C_2(z_1 I - A_{22})^{-1} A_{21})(z_2 I - A_{11} - A_{12}(z_1 I - A_{22})^{-1} A_{21})^{-1}(B_1 + A_{12}(z_1 I - A_{22})^{-1} B_2) + \\ + C_2(z_1 I - A_{22})^{-1} B_2 + D. \end{aligned} \tag{3.8}$$

Like in the 1-D case, all entries of a 2-D state-space model transfer matrix are proper rational functions. However, since here those functions depend on two variables, a new definition of *properness* has to be considered.

Definition A rational function $g(z_1,z_2) = n(z_1,z_2)/d(z_1,z_2)$, where $n(z_1,z_2)$ and $d(z_1,z_2)$ are polynomials, is said to be *proper* if:

1. The degree in z_1 of $d(z_1,z_2)$ is not less than the degree in z_1 of $n(z_1,z_2)$.
2. The degree in z_2 of $d(z_1,z_2)$ is not less than the degree in z_2 of $n(z_1,z_2)$.
3. Denoting the degrees of $d(z_1,z_2)$ in z_1 and in z_2 by k and q, respectively, the coefficient of the monomial $z_1^k z_2^q$ of $d(z_1,z_2)$ is nonzero.

If, in 1. and 2., the expression "not less" can be replaced by "larger", $g(z_1,z_2)$ is called *strictly proper*.

Once again, if all entries of a transfer matrix are proper rational functions, that transfer matrix is also said to be proper.
Analogously to what was defined for 1-D systems, $\Sigma = (A_{11},A_{12},A_{21},A_{22},B_1,B_2,C_1,C_2,D)$ is said to be a *realization* of a transfer matrix $G(z_1,z_2)$ or impulse response $h(i,j)$ if, respectively,

$G_\Sigma(z_1,z_2) = G(z_1,z_2)$ or $h_\Sigma(i,j) = h(i,j)$.
Assuming that a 2-D proper transfer matrix, $G(z_1,z_2)$, is given, one can always see it as a 1-D transfer matrix, $F(z_2) = (G(z_1))(z_2)$, where the coefficients are in the ring[1] of proper rational functions in the variable z_1 with coefficients in $\mathbb{R}$, here denoted by $\mathbb{R}_p(z_1)$. This statement, which is the key idea of R. Eising's algebraic approach for 2-D systems (see [1]), motivates the discussion of the following section.

3.1 First level and second level realizations

In this section, it is shown in detail how to obtain a Roesser's state-space realization of a given 2-D transfer matrix. The approach to 2-D systems as 1-D systems over a ring will be explored.

3.1.1 First level realizations as 1-D systems over a ring

The set of polynomials in z_1 and z_2 with real coefficients is denoted by $\mathbb{R}[z_1,z_2]$, the set of rational functions in z_1 and z_2 with real coefficients is denoted by $\mathbb{R}(z_1,z_2)$ and $\mathbb{R}_p^{m\times p}(z_1,z_2)$ stands for the set of 2-D proper rational $m\times p$ matrices.
If $G(z_1,z_2)\in\mathbb{R}_p^{m\times p}(z_1,z_2)$, then it can be written as $G(z_1,z_2) = N(z_1,z_2)/D(z_1,z_2)$, where $N(z_1,z_2)\in \mathbb{R}^{m\times p}[z_1,z_2]$ and $D(z_1,z_2) = d_0(z_1) + d_1(z_1)z_2 + \cdots + d_q(z_1)z_2^q \in \mathbb{R}[z_1,z_2]$, with $d_q(z_1)$ nonzero. Dividing $N(z_1,z_2)$ and $D(z_1,z_2)$ by $d_q(z_1)$ makes all coefficients of $F(z_2) = (G(z_1))(z_2)$ become proper rational functions in z_1 (recall that $G(z_1,z_2)$ was assumed to be proper). Furthermore, the denominator polynomial becomes *monic*, i.e., its leading coefficient equals one.
Now, $G(z_1,z_2)$ can be seen as an element of $(\mathbb{R}_p^{m\times p}(z_1))_p(z_2)$, i.e., as a 1-D transfer matrix whose coefficients are in the ring of 1-D proper transfer functions, $\mathbb{R}_p(z_1)$. One can also see the corresponding impulse response of the system, $h(i,j)$, as a 1-D impulse response T with $T(i)\in\mathbb{R}_p(z_1)$. Thus, if one has an algorithm to obtain a (preferably) minimal realization of $F(z_2)$ or T over $\mathbb{R}_p(z_1)$, then it will be possible to find a realization $\Sigma = (A(z_1),B(z_1),C(z_1),D(z_1))$ such that:

$$G(z_1,z_2) = C(z_1)(z_2 I - A(z_1))^{-1}B(z_1) + D(z_1). \tag{3.9}$$

Such Σ is called a *first level realization* of $G(z_1,z_2)$. Clearly, the state-space model for a first level realization follows a structure that is completely identical to the classic 1-D state-space model:

$$x_{i+1}(z_1) = A(z_1)x_i(z_1) + B(z_1)u_i(z_1) \tag{3.10}$$

$$y_i(z_1) = C(z_1)x_i(z_1) + D(z_1)u_i(z_1), \tag{3.11}$$

[1] See Appendix A for the definition of ring

where $u_i(z_1)$ and $y_i(z_1)$ are related with the input, $u(i,j)$, and the output, $y(i,j)$, of the 2-D system by:

$$u_i(z_1) = \sum_{j=0}^{\infty} u(i,j) z_1^{-j} \tag{3.12}$$

$$y_i(z_1) = \sum_{j=0}^{\infty} y(i,j) z_1^{-j}. \tag{3.13}$$

Remark Since the variables z_1 and z_2 play an identical role in the transfer matrix description, $G(z_1,z_2)$ may also be seen as an element of $(\mathbb{R}_p^{m\times p}(z_2))_p(z_1)$. Thus, the discussion above remains valid if the roles of these variables are interchanged.

Next, an algorithm to find a minimal realization over a principal ideal domain[2] R from an impulse response over R will be given. This algorithm was presented by R. Eising in [1], who also proved that the ring $\mathbb{R}_p(z_1)$ is actually a principal ideal domain.

Algorithm 3. *Minimal realization over a principal ideal domain R*
Input data: impulse response h with $h(i) \in R$

1. *Build the Hankel matrix* $H_{\alpha,\beta}$
2. *Factorize* $H_{\alpha,\beta}$ *as a Smith form, i.e.,*

$$H_{\alpha,\beta} = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V, \tag{3.14}$$

   where U and V are unimodular matrices (i.e., invertible over R) and D is $n_{min} \times n_{min}$ *diagonal. This factorization is always possible due to the fact that* $H_{\alpha,\beta}$ *is a matrix over a principal ideal domain (see [1]).*
3. *Define matrices* $\mathscr{O}_\alpha$ *and* $\mathscr{C}_\beta$ *as*

$$\mathscr{O}_\alpha := U \begin{bmatrix} D \\ 0 \end{bmatrix}, \quad \mathscr{C}_\beta := \begin{bmatrix} I & 0 \end{bmatrix} V. \tag{3.15}$$

   Note that a right inverse of $\mathscr{C}_\beta$ *is* $\mathscr{C}_\beta^\dagger = V^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix}$ *and* $\mathscr{O}_\alpha \mathscr{C}_\beta = H_{\alpha,\beta}$.
4. *Obtain* $A \in R^{n_{min}\times n_{min}}$, $B \in R^{n_{min}\times p}$, $C \in R^{m\times n_{min}}$ *and* $D \in R^{m\times p}$ *as in Algorithm 2 (subsection 2.2.2).* $\Sigma = (A,B,C,D)$ *is a realization of h over R.*

Remark It is shown in [1] that any realization algorithm that starts from a factorization $\mathscr{O}_\alpha \mathscr{C}_\beta$, over R, such that $\mathscr{O}_\alpha$ is right regular (i.e., the zero vector is the only vector with entries in R that belongs to the kernel of $\mathscr{O}_\alpha$) and $\mathscr{C}_\beta$ is right invertible, will produce a realization over R.

[2] See Appendix A for the definition of principal ideal domain

The algorithm just presented implicitly leads to the following theorem.

Theorem 3.1.1. *Let R be a principal ideal domain and h with $h(i) \in R^{m\times p}$ an impulse response whose minimal realizations over $Q(R)$, the field[3] of fractions of R, have dimension n_{min}. Then, there exists a minimal realization of h over R with dimension n_{min}.*

Proof. See [1]. □

For the case treated in this thesis, it could be more useful to have a realization algorithm that starts from the transfer matrix instead of the impulse response. Such algorithm is also presented in [1]. However, as it is going to be shown in the following example, obtaining the impulse response from a transfer matrix is not that hard, so Algorithm 3 will be enough for this discussion.

Example Consider the following SISO proper transfer function:

$$g(z_1,z_2) := \frac{z_2}{z_2^2 z_1 + 2z_1}. \tag{3.16}$$

This transfer function can be rewritten as

$$g(z_1,z_2) = \frac{\frac{z_2}{z_1}}{z_2^2+2}. \tag{3.17}$$

Now, $g(z_1,z_2)$ can be seen as an element of $(\mathbb{R}_p(z_1))_p(z_2)$. Since the numerator and the denominator polynomials in z_2 are coprime, one concludes, from Theorems 2.1.5 and 3.1.1, that the minimal dimension of the first level realization is $n_{min} = 2$.
It is easy to expand $g(z_1,z_2)$ as a formal power series in z_2^{-k}, $k = 0,1,2,...$:

$$\begin{aligned} g(z_1,z_2) &= \frac{1}{z_1 z_2}\frac{1}{1-\left(\frac{-2}{z_2^2}\right)} = \frac{1}{z_1 z_2}\left(1 - \frac{2}{z_2^2} + \left(\frac{2}{z_2^2}\right)^2 + \cdots + \left(\frac{-2}{z_2^2}\right)^k + \cdots\right) = \\ &= \frac{1}{z_1} z_2^{-1} + \frac{-2}{z_1} z_2^{-3} + \cdots + \frac{(-2)^k}{z_1} z_2^{-(2k+1)} + \cdots \end{aligned} \tag{3.18}$$

If $g(z_1,z_2)$ is thought as a function of the complex variable z_2, this expansion corresponds to the Laurent series at $z_2 = \infty$.
Note that, by the definition of a transfer function, the value of the impulse response at "time" k, $h(k)$, corresponds to the coefficient of z_2^{-k} in the above series. Since $n_{min} = 2$, the Hankel matrix $H_{2,3}$ contains all the data needed to find a minimal realization.

$$H_{2,3} = \begin{bmatrix} 1/z_1 & 0 & -2/z_1 \\ 0 & -2/z_1 & 0 \end{bmatrix} \tag{3.19}$$

[3] See Appendix A for the definition of field

Now one could compute a Smith decomposition for $H_{2,3}$, but the following factorization is rather obvious and satisfies the conditions stated in the previous remark:

$$H_{2,3} = \begin{bmatrix} 1/z_1 & 0 \\ 0 & -2/z_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix}. \tag{3.20}$$

Thus, $\mathscr{O}_2 = \begin{bmatrix} 1/z_1 & 0 \\ 0 & -2/z_1 \end{bmatrix}$ and $\mathscr{C}_3 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix}$.
A first level realization $\Sigma = (A(z_1), B(z_1), C(z_1), D(z_1))$ for $g(z_1, z_2)$ is finally obtained using the formulas indicated in the step 4. of Algorithm 2:

$$\begin{aligned} A &= \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}, \\ B &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; C = \begin{bmatrix} 1/z_1 & 0 \end{bmatrix}, \; D = 0. \end{aligned} \tag{3.21}$$

Remark Compare the obtained realization with the structure of the one given in the proof of Theorem 2.1.5, which was stated for a system over a field. It is not by chance that the realization obtained here follows exactly that structure. In fact, for a SISO system, one can always find a minimal realization over $\mathbb{R}(z_1)$ by simply extracting the coefficients of the 2-D transfer function, when it is written like in (3.17). Then, one just has to build matrices A, B, C and D as done in that proof. This gives a very fast way of finding a minimal first level realization of a 2-D SISO transfer function.

3.1.2 Second level realizations

It has just been shown how to obtain a first level realization from a 2-D transfer matrix $G(z_1, z_2)$. The next step is to obtain a realization of the 2-D input-ouput system that does not depend on either z_1 or z_2, i.e., to obtain a Roesser model that describes the system. Such model will be called a *second level realization* of $G(z_1, z_2)$.
Consider the system (3.10)-(3.11) as depending only in z_1, i.e., consider i as a constant. Rewrite the system as

$$\begin{bmatrix} x_{i+1}(z_1) \\ y_i(z_1) \end{bmatrix} = \begin{bmatrix} A(z_1) & B(z_1) \\ C(z_1) & D(z_1) \end{bmatrix} \begin{bmatrix} x_i(z_1) \\ u_i(z_1) \end{bmatrix}. \tag{3.22}$$

Thus, it can be seen as a 1-D system with input $\begin{bmatrix} x_i(z_1)^T & u_i(z_1)^T \end{bmatrix}^T$ and output $\begin{bmatrix} x_{i+1}(z_1)^T & y_i(z_1)^T \end{bmatrix}^T$ whose transfer matrix is $H(z_1) = \begin{bmatrix} A(z_1) & B(z_1) \\ C(z_1) & D(z_1) \end{bmatrix}$. If a realization for $H(z_1)$ is found, then one actually obtains a second level realization. The fact that it indeed corresponds to a Roesser model might not be obvious, but it will become clear in the following.

Suppose that a (preferably) minimal realization for $H(z_1)$ is found:

$$x_v(j+1) = \tilde{A}x_v(j) + \begin{bmatrix}\tilde{B}_1 & \tilde{B}_2\end{bmatrix}\begin{bmatrix}x_h(i,j)\\ u(i,j)\end{bmatrix} \tag{3.23}$$

$$\begin{bmatrix}x_h(i+1,j)\\ y(i,j)\end{bmatrix} = \begin{bmatrix}\tilde{C}_1\\ \tilde{C}_2\end{bmatrix}x_v(j) + \begin{bmatrix}\tilde{D}_{11} & \tilde{D}_{12}\\ \tilde{D}_{21} & \tilde{D}_{22}\end{bmatrix}\begin{bmatrix}x_h(i,j)\\ u(i,j)\end{bmatrix}, \tag{3.24}$$

where $x_v(i,j)$ is the new state vector and $x_h(i,j)$ is related with $x_i(z_1)$ by

$$x_i(z_1) = \sum_{j=0}^{\infty} x_h(i,j)z_1^{-j}. \tag{3.25}$$

Reorganizing the equations, one obtains finally:

$$\begin{bmatrix}x_h(i+1,j)\\ x_v(i,j+1)\end{bmatrix} = \begin{bmatrix}\tilde{D}_{11} & \tilde{C}_1\\ \tilde{B}_1 & \tilde{A}\end{bmatrix}\begin{bmatrix}x_h(i,j)\\ x_v(i,j)\end{bmatrix} + \begin{bmatrix}\tilde{D}_{12}\\ \tilde{B}_2\end{bmatrix}u(i,j) \tag{3.26}$$

$$y(i,j) = \begin{bmatrix}\tilde{D}_{21} & \tilde{C}_2\end{bmatrix}\begin{bmatrix}x_h(i,j)\\ x_v(i,j)\end{bmatrix} + \tilde{D}_{22}u(i,j). \tag{3.27}$$

This corresponds obviously to a Roesser model, defined in (3.1)-(3.2).

Remark Even if the realization of $H(z_1)$ defined in equations (3.23)-(3.24) is minimal, it is not guaranteed that the obtained Roesser model is a minimal realization of the initial 2-D transfer function $G(z_1,z_2)$. A clear explanation of this fact will be given later on.

3.2 Equivalent 2-D state-space models

By applying a change of basis to the state vector of a 2-D model, it is possible obtain a new model with the same input-ouput dynamics (i.e., with the same transfer function). However, for this case, the problem is not as trivial as it was for 1-D systems (recall the first results of section 2.1.1). Consider a 2-D state-space model $\Sigma = (A_{11},A_{12},A_{21},A_{22},B_1,B_2,C_1,C_2,D)$ of dimension n. From now on, such model will sometimes be abbreviately denoted by $\Sigma = (A,B,C,D)$, where

$$\begin{aligned} A &:= \begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix},\\ B := \begin{bmatrix}B_1\\ B_2\end{bmatrix},\ C &:= \begin{bmatrix}C_1 & C_2\end{bmatrix}. \end{aligned} \tag{3.28}$$

Defining

$$Z := \begin{bmatrix}z_2 I_{n_2} & 0\\ 0 & z_1 I_{n_1}\end{bmatrix}, \tag{3.29}$$

where $n_1 := dim(A_{22})$ and $n_2 := dim(A_{11})$, the transfer matrix of Σ may be expressed as

$$G(z_1,z_2) = C(Z-A)^{-1}B+D. \tag{3.30}$$

Note that equations (2.6) and (3.30) have a rather similar appearance. The only relevant difference is that z in (2.6) is a scalar, while Z in (3.30) is an $n \times n$ matrix. This will be of vital importance in the following. Consider the transfer function of the 2-D model $\overline{\Sigma} = (T^{-1}AT, T^{-1}B, CT, D)$, where T is an arbitrary nonsingular $n \times n$ real matrix:

$$\begin{aligned} \overline{G}(z_1,z_2) &= CT(Z-T^{-1}AT)^{-1}T^{-1}B+D = \\ &= C(TZT^{-1}-A)^{-1}B+D. \end{aligned} \tag{3.31}$$

Clearly, for a general T, models Σ and $\overline{\Sigma}$ are not realizations of the same input-output system, since they do not yield the same transfer function. However, if matrices Z and T commute, $G(z_1,z_2)$ and $\overline{G}(z_1,z_2)$ obviously coincide and, therefore, Σ and $\overline{\Sigma}$ define the same input-output dynamics. This is a key difference relatively to the 1-D case, in which every nonsingular matrix $T \in \mathbb{R}^{n\times n}$ would generate a model with the same transfer function.
Given the structure of Z, defined in (3.29), it is not hard to see that Z and T commute, and hence $G(z_1,z_2)$ and $\overline{G}(z_1,z_2)$ coincide, if and only if $T \in \mathscr{T}$, where:

$$\mathscr{T} = \left\{ T \in \mathbb{R}^{n\times n} : T = \begin{bmatrix} T_2 & 0 \\ 0 & T_1 \end{bmatrix},\ T_1 \in \mathbb{R}^{n_1\times n_1} \text{ and nonsingular},\ T_2 \in \mathbb{R}^{n_2\times n_2} \text{ and nonsingular} \right\}. \tag{3.32}$$

Remark In a state-space perspective, applying a transformation $T \in \mathscr{T}$ corresponds to perform a separate change of basis both for the horizontal and for the vertical state. That is, the original horizontal state x_h is transformed into a new horizontal state $\overline{x}_h = T_2^{-1}x_h$ and the original vertical state x_v is transformed into $\overline{x}_v = T_1^{-1}x_v$. Note that there is no relation between x_h and $\overline{x}_v$ neither between x_v and $\overline{x}_h$.

The result presented in this section has a crucial importance in one of the model order reduction algorithms that are presented later on. That algorithm relies upon a change of basis that transforms the original model into a *balanced* realization, in a sense that is somehow related with the 1-D definition.

3.3 Controllability and observability of 2-D models

Like in the one-dimensional case, controllability and observability of a realization are somehow related to its minimality. To state the definitions of these properties, the notion of coprimeness of polynomial matrices in two variables is needed.

Definition Two matrices $P(z_1,z_2)$ and $Q(z_1,z_2)$ over $\mathbb{R}[z_1,z_2]$, with the same number of rows, are called *left coprime* with respect to $\mathbb{C}[z_1,z_2]$ if, for every left common factor $D(z_1,z_2)$ such

that $P(z_1,z_2) = D(z_1,z_2)\tilde{P}(z_1,z_2)$ and $Q(z_1,z_2) = D(z_1,z_2)\tilde{Q}(z_1,z_2)$, where $D(z_1,z_2)$, $\tilde{P}(z_1,z_2)$ and $\tilde{Q}(z_1,z_2)$ are matrices over $\mathbb{C}[z_1,z_2]$ and $D(z_1,z_2)$ is a square matrix, one necessarily has that $det(D(z_1,z_2))$ is a complex number $d \neq 0$.

The definition of right coprimeness is analogous to this one. Note that polynomials with complex coefficients are considered in these definitions because the field where coefficients belong to must be algebraically closed.

Definition The system (3.1)-(3.2) is *controllable* if

$$\begin{bmatrix} z_2 I - A_{11} & -A_{12} \\ -A_{21} & z_1 I - A_{22} \end{bmatrix} \text{ and } \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \tag{3.33}$$

are left coprime with respect to $\mathbb{C}[z_1,z_2]$.

Definition The system (3.1)-(3.2) is *observable* if

$$[C_1\ C_2] \text{ and } \begin{bmatrix} z_2 I - A_{11} & -A_{12} \\ -A_{21} & z_1 I - A_{22} \end{bmatrix} \tag{3.34}$$

are right coprime with respect to $\mathbb{C}[z_1,z_2]$.

Contrarily to what was done for the one-dimensional case, definitions of controllability and observability given here are only formal. However, for the purpose of this thesis, the importance of these notions is simply related with the minimality of the realization, so these formal definitions, which are also used in [1], are enough. For 1-D systems, analogous definitions could have been used, because they imply the definitions that were actually previously stated (see section 2.1).

3.4 Minimality of 2-D models

As shown before, the problem of finding a minimal first level realization is equivalent to find a minimal realization for a system defined over a ring. Algorithm 3 gives a solution for this problem. However, as mentioned at the end of the previous section, for a Roesser model, it is not clear what is the minimum dimension of the state vector that is needed to define the system. It is known, though, that if a Roesser model is controllable and observable, then the matrices A_{11} and A_{22} of the model (3.1)-(3.2) will have, respectively, dimensions $q \times q$ and $k \times k$, where q is the degree in z_2 and k is the degree in z_1 of the corresponding transfer function, when it is written as a fraction of coprime polynomials. If the denominator of the transfer function is separable, i.e., if $D(z_1,z_2) = D_1(z_1)D_2(z_2) = D_2(z_2)D_1(z_1)$, it is indeed possible to construct a realization where A_{11} has dimension $q \times q$ and A_{22} has dimension $k \times k$, so it is surely a minimal realization with order $n = k + q$. However, for the more general case, it is not even known if there exists a realization with dimension $k + q$. R. Eising established in [1] upper bounds for the dimension of a minimal realization over $\mathbb{R}$. E. Sontag proved the interesting fact that, in some cases, the dimension of a

minimal realization over $\mathbb{C}$ is lower than the dimension of a minimal realization over $\mathbb{R}$ for the same input-output system (see [4]).

In section 3.1.2, it was remarked that finding a minimal realization of a transfer function $H(z_1)$ that corresponds to an also minimal first level realization does not ensure that the final 2-D model has minimal dimension. Now, that statement is going to be conveniently justified. Suppose that one has a minimal first level realization $\Sigma = (A(z_1), B(z_1), C(z_1), D(z_1))$ of a transfer matrix $G(z_1, z_2)$. A straightforward extension of Theorem 2.1.4 for systems over a ring tells that the quadruple of matrices $(\overline{A}(z_1), \overline{B}(z_1), \overline{C}(z_1), \overline{D}(z_1))$, where:

$$\overline{A}(z_1) = T(z_1)^{-1} A(z_1) T(z_1),\ \ \overline{B}(z_1) = T(z_1)^{-1} B(z_1),\ \ \overline{C}(z_1) = C(z_1) T(z_1),\ \ \overline{D}(z_1) = D(z_1), \tag{3.35}$$

for some invertible (over the field $Q(\mathbb{R}_p(z_1))$) matrix $T(z_1)$, is also a first level realization of $G(z_1, z_2)$, provided that all these four matrices are defined over $\mathbb{R}_p(z_1)$ (note that $T(z_1)$ being unimodular over $\mathbb{R}_p(z_1)$ is a sufficient but not necessary condition). Thus, a realization of

$$\overline{H}(z_1) := \begin{bmatrix} \overline{A}(z_1) & \overline{B}(z_1) \\ \overline{C}(z_1) & \overline{D}(z_1) \end{bmatrix} := \begin{bmatrix} T(z_1)^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A(z_1) & B(z_1) \\ C(z_1) & D(z_1) \end{bmatrix} \begin{bmatrix} T(z_1) & 0 \\ 0 & I \end{bmatrix} \tag{3.36}$$

is also a second level realization of $G(z_1, z_2)$. As could be expected and the following example illustrates, the McMillan degree of $\overline{H}(z_1)$ depends on the matrix $T(z_1)$ that is chosen, therefore the corresponding minimal second level realizations may have different dimensions. This is illustrated by the following example.

Example Consider the first level realization $\Sigma = (A(z_1), B(z_1), C(z_1), D(z_1))$, where

$$\begin{gathered} A(z_1) := \begin{bmatrix} 0 & 1 \\ 5 & -1/(2z_1+1) \end{bmatrix},\ \ B(z_1) := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ C(z_1) := \begin{bmatrix} 1/(2z_1+1) & 1 \end{bmatrix},\ \ D(z_1) := 0. \end{gathered} \tag{3.37}$$

Now build the corresponding transfer matrix, $H(z_1)$, as:

$$H(z_1) = \begin{bmatrix} A(z_1) & B(z_1) \\ C(z_1) & D(z_1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 5 & -1/(2z_1+1) & 1 \\ 1/(2z_1+1) & 1 & 0 \end{bmatrix}. \tag{3.38}$$

According to Theorem 2.1.6, the McMillan degree of $H(z_1)$ equals 2. A minimal realization of

$H(z_1)$ provides the 2-D model $\Sigma = (A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, D)$ where

$$\begin{aligned} A_{11} &= \begin{bmatrix} 0 & 1 \\ 5 & 0 \end{bmatrix}, \; A_{12} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \\ A_{21} &= \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \; A_{22} = \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix}, \\ B_1 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ C_1 &= \begin{bmatrix} 0 & 1 \end{bmatrix}, \; C_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \\ D &= 0. \end{aligned} \tag{3.39}$$

Clearly, the 2-D model Σ has dimension $2+2=4$.
Defining

$$T(z_1) := \begin{bmatrix} 0 & 1 \\ 1 & -1/(2z_1+1) \end{bmatrix} \tag{3.40}$$

and building $\overline{H}(z_1)$ as defined in (3.36), one obtains:

$$\overline{H}(z_1) = \begin{bmatrix} 0 & 5 & 1 \\ 1 & -1/(2z_1+1) & 0 \\ 1 & 0 & 0 \end{bmatrix}, \tag{3.41}$$

which is a proper transfer matrix whose McMillan degree is 1. A minimal realization of $\overline{H}(z_1)$ yields a 2-D model $\overline{\Sigma} = (\overline{A}_{11}, \overline{A}_{12}, \overline{A}_{21}, \overline{A}_{22}, \overline{B}_1, \overline{B}_2, \overline{C}_1, \overline{C}_2, \overline{D})$ where:

$$\begin{aligned} \overline{A}_{11} &= \begin{bmatrix} 0 & 5 \\ 1 & 0 \end{bmatrix}, \; \overline{A}_{12} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \\ \overline{A}_{21} &= \begin{bmatrix} 0 & 0.5 \end{bmatrix}, \; \overline{A}_{22} = -0.5, \\ \overline{B}_1 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; \overline{B}_2 = 0, \\ \overline{C}_1 &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \; \overline{C}_2 = 0, \\ \overline{D} &= 0. \end{aligned} \tag{3.42}$$

The 2-D model $\overline{\Sigma}$ has lower dimension $2+1=3$.

3.5 Stability

A deep discussion on the stability of 2-D systems is not the purpose of this thesis. However, just like in the one-dimensional case, stability of a 2-D system is very often a highly desirable

property. In the context of model order reduction, one usually wants to ensure that the obtained reduced order system is stable, given that the original one was also stable.

3.5.1 BIBO stability

Here, a particular notion of stability is considered: a 2-D system is said to be BIBO stable if an amplitude bounded input always originates an amplitude bounded output. The acronym BIBO stands precisely for bounded-input / bounded-output. The rigorous definition of BIBO stability for 2-D systems follows.

Definition A 2-D system with input $u(i,j)$ and output $y(i,j)$ is said to be *BIBO stable* if, for all $M < \infty$ such that $\|u(i,j)\| < M$, there exists an $N < \infty$ such that $\|y(i,j)\| < N$.

Given the impulse response of the system, the following theorem states a criterion for checking its stability. This criterion is analogous to a result that stands for 1-D systems (see, for instance, [13]).

Theorem 3.5.1. *A 2-D system with impulse response $h(i,j)$ is BIBO stable if and only if $h(i,j)$ is absolutely summable, i.e., if and only if*

$$\sum_{i=0,j=0}^{\infty,\infty} \|h(i,j)\| < \infty. \tag{3.43}$$

Proof. See [14]. □

In most cases, it is more practical to consider the transfer function instead of the impulse response. As it could be expected, it is also possible to determine whether a system is BIBO stable or not by conveniently studying its transfer function. In order to address this issue, the following auxiliary definition is crucial.

Definition Consider a 2-D rational transfer function $g(z_1,z_2) = n(z_1,z_2)/d(z_1,z_2)$, where $n(z_1,z_2)$ and $d(z_1,z_2)$ are factor coprime polynomials. If $d(z_1^*,z_2^*) = 0$, then (z_1^*,z_2^*) is said to be a *singularity* of $g(z_1,z_2)$. Furthermore, if $n(z_1^*,z_2^*) \neq 0$, then (z_1^*,z_2^*) is a *1st type singularity* or *pole*. Otherwise, (z_1^*,z_2^*) is a *2nd type singularity*.

Let

$$\bar{U} := \{(z_1,z_2) \in \mathbb{C}^2 : |z_1| \geq 1, |z_2| \geq 1\} \tag{3.44}$$

and

$$W := \{(z_1,z_2) \in \mathbb{C}^2 : |z_1| = 1, |z_2| = 1\}. \tag{3.45}$$

Then, the following result holds.

Theorem 3.5.2. *A transfer function $g(z_1,z_2)$ is BIBO stable if there are no singularities in $\bar{U}$.*

Theorem 3.5.3. *A transfer function $g(z_1,z_2)$ is BIBO stable only if there are no poles in $\bar{U}$ and no 2nd type singularities in $\bar{U}-W$.*

Proof. A proof of these theorems may be found in [15]. □

Hence, if there are no 2nd type singularities in $\bar{U}-W$, a transfer function is BIBO stable if and only if there are no singularities in $\bar{U}$.

Remark The theorems above may be easily extended to the matrix case. Just write a transfer matrix $G(z_1,z_2)$ as $G(z_1,z_2)=N(z_1,z_2)/d(z_1,z_2)$, where $N(z_1,z_2)$ is a polynomial matrix and $d(z_1,z_2)$ is a polynomial that is coprime with $N(z_1,z_2)$, i.e., a common divisor of $d(z_1,z_2)$ and all entries of $N(z_1,z_2)$ is necessarily a constant. Then, use the above criteria to check every transfer function obtained by dividing one entry of $N(z_1,z_2)$ by $d(z_1,z_2)$. The transfer matrix is stable if and only if all these transfer functions are stable.

These theorems provide rather easy tests to check the stability of a 2-D transfer function. Stability of transfer functions that do not have 2nd type singularities in $\bar{U}-W$ may also be tested using the following theorem, which is more useful when one is considering a Roesser state-space realization of that transfer function instead of the transfer function itself.

Theorem 3.5.4. *A 2-D model $\Sigma=(A,B,C,D)$ realizing a transfer matrix $G_\Sigma(z_1,z_2)$ with no 2nd type singularities in $\bar{U}-W$ is BIBO stable if $(Z-A)$ is nonsingular in $\bar{U}$.*

Proof. See [15]. □

As it has just been seen, it is much easier to check stability of transfer functions without 2nd type singularities in $\bar{U}-W$. For this reason, from now on, only this class of 2-D transfer functions will be considered.

3.5.2 Internal stability

In BIBO stability analysis, the only concern is verifying whether the output is bounded or not for a given bounded input. However, to ensure the applicability of many model order reduction algorithms, BIBO stability is not enough. In fact, the model must be *internally* stable, i.e., its state vector evolution must also be *stable* in a sense that is defined in the following.

Definition A set of initial conditions $x_h(0,j)$, $j=0,1,2,...$, and $x_v(i,0)$, $i=0,1,2,...$, is said to be *bounded* if $\exists M>0:\ \sup_j \|x_h(0,j)\|<M$ and $\sup_i \|x_v(i,0)\|<M$.

Definition A 2-D model $\Sigma=(A,B,C,D)$ is *internally stable* if, for zero input and for any bounded initial conditions $x_h(0,j)$, $j=0,1,2,...$, and $x_v(i,0)$, $i=0,1,2,...$,

$$\lim_{\|(i,j)\|\to\infty}\left\|\begin{bmatrix}x_h(i,j)\\ x_v(i,j)\end{bmatrix}\right\|=0. \tag{3.46}$$

Here, $\lim_{\|(i,j)\|\to\infty}$ denotes the limit along any path (i_n,j_n), $n\in\mathbb{N}$, such that $\lim_{n\to\infty}\sqrt{i_n^2+j_n^2}=\infty$.

The following theorem provides necessary and sufficient conditions for the internal stability of a 2-D model.

Theorem 3.5.5. *A 2-D model $\Sigma = (A,B,C,D)$ is internally stable if and only if $(Z-A)$ is nonsingular in $\bar{U}$, where Z and $\bar{U}$ are as defined in (3.29) and (3.44), respectively.*

Proof. See [3] □

From Theorems 3.5.4 and 3.5.5, one immediately obtains the following relation between internal and BIBO stability of a 2-D model Σ:

$$\Sigma \text{ is internally stable} \Rightarrow \Sigma \text{ is BIBO stable.} \tag{3.47}$$

As in the 1-D case, the converse implication is false.

Chapter 4

Model order reduction for 2-D systems

4.1 State of the art

Over the past few years, some methods for model order reduction of 2-D linear systems have been proposed. Three main approaches can be identified: reduction by finding an approximate transfer function of lower degree, low-rank approximation of the 2-D system Hankel matrix and direct reduction of the 2-D state-space realization.

A method that follows the first approach was proposed by P. Paraskevopoulos *et al.* (see [7]). It starts from the exact 2-D transfer function and transforms the problem of model order reduction into an overdetermined linear algebraic system of equations, by using orthogonal series. However, the obtained result is not a state-space realization, but rather a 2-D transfer function, whose degree is obviously smaller than the degree of the original one. For that reason, this method cannot be considered, by itself, a state-space model order reduction algorithm. Another method, developed by T. Guo *et al.* ([16]), is also based on a transfer function reduction, but uses bilinear Routh approximations instead of orthogonal series and it is only applicable to separable denominator systems, which are a particular class of 2-D systems. These systems can be thought as the result of a cascade interconnection between two 1-D systems, so they are quite easier to treat. In terms of the Roesser model, their state-space realizations have the property that either A_{12} or A_{21} are zero matrices.

The second approach generalizes the idea of the 1-D Hankel matrix approximation for 2-D systems. H. Luo *et al.* ([5]) presented a method in which the problem of model order reduction is formulated as an unconstrained optimization problem, whose objective function includes a term that depends on the sum of the discarded 2-D Hankel singular values. M. Diab *et al.* ([6]) used the singular value decomposition of the horizontal and vertical Hankel matrices of the system to obtain diagonal gramians which are used to derive a reduced order model.

The last class of methods seems to be the most popular. Some of them consider the Fornasini-Marchesini model (see, for instance, [17]), but the Roesser model is indeed the most widely used. The majority of the algorithms using Roesser state-space model try to generalize the notion of balanced realization for 2-D systems, which, as we have seen for the one-dimensional case, is

very useful for model order reduction. P. Misra *et al.* ([18]) followed this idea and presented a model order reduction algorithm for separable denominator systems. K. Zhou *et al.* ([19]) used the concept of a pseudo-balanced realization for 2-D systems, introduced in [20], to develop a method which is applicable to any 2-D realization described by a Roesser model, but conditions for the stability of the reduced order system are only given for separable denominator systems. C. Xiao *et al.* ([21]) proposed a quite simple algorithm, also starting from a Roesser state-space representation. The biggest computation effort required by this algorithm is the resolution of some Lyapunov equations, what makes its complexity very similar to the complexity of the 1-D case. The asymptotic stability of the obtained reduced order system is also guaranteed if the original one is also asymptotically stable. However, once again, it has the drawback of being only applicable to separable denominator systems. Recently, C. Wang *et al.* ([22]) came up with an algorithm for model order reduction of positive 2-D systems where the error system (i.e., the system defined by the difference of the original and reduced order transfer functions) satisfies a prescribed constraint for its $\mathscr{H}_\infty$ norm. A positive system is, however, a very particular class of systems, where the state components and the output are nonnegative whenever the initial conditions and input are nonnegative.

As just pointed out, a great part of the existing model order reduction methods for 2-D systems can only be applied to particular classes of these systems, namely to separable denominator ones. However, the relevance of these systems is far from being negligible, since any 2-D FIR filter is indeed a separable denominator system and any stable IIR filter can be approximated by a FIR ([21]). It is also interesting to observe that none of the methods found compute the reduced order model by using the first level realization. This approach would imply the usage of a model order reduction method for 1-D systems defined over a ring (more precisely, over a principal ideal domain). Apparently, although some investigation about minimality of this kind of systems has already been done (see, for instance, [1], [23] and [24]), such method does not exist yet. In this chapter, the main difficulties of this problem will be highlighted and some new results will be presented.

There is a particular perspective of multidimensional systems which was found to be useful in the development of model order reduction algorithms. In fact, a 2-D system may be modelled as a constant gain system plus a structured uncertainty, which represents the shift operators. Uncertain systems arise very commonly in robust control problems and some algorithms for model order reduction of these kind of systems already exist (see [8]). One of these algorithms will be presented in great detail later on. When applied to multidimensional systems, that algorithm may be included in the last class of algorithms described above, since it also uses an extension of the concept of balanced realizations.

4.2 Order reduction of first level realizations – preliminary results

The focus of this section is on balancing first level realizations. The notion of controllability and observability gramians will be extended for these kind of realizations and some effort developed

to indeed obtain a *rational balanced realization* will be presented.

4.2.1 Generalized gramians

For a matter of simplicity, the reasoning that leads to the definitions of the generalized gramians will assume that the considered model is a SISO one.
Recall that the state-space model corresponding to a first level realization $\Sigma = (A(z_1), B(z_1), C(z_1), D(z_1))$ is given by:

$$x_{i+1}(z_1) = A(z_1)x_i(z_1) + B(z_1)u_i(z_1) \tag{4.1}$$

$$y_i(z_1) = C(z_1)x_i(z_1) + D(z_1)u_i(z_1), \tag{4.2}$$

where

$$u_i(z_1) = \sum_{j=0}^{\infty} u(i,j)z_1^{-j}, \tag{4.3}$$

$$y_i(z_1) = \sum_{j=0}^{\infty} y(i,j)z_1^{-j}, \tag{4.4}$$

and $A(z_1)$, $B(z_1)$, $C(z_1)$ and $D(z_1)$ are matrices over $\mathbb{R}_p(z_1)$.
The energy of the signal $u(i,j)$ is defined analogously to the 1-D case:

$$E_u := \sum_{i=0,j=0}^{\infty,\infty} u(i,j)^2. \tag{4.5}$$

By using Parseval's identity (see, for instance, [25]), one may write the above sum as

$$E_u = \sum_{i=0}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} u_i(e^{-j\omega})u_i(e^{j\omega})d\omega, \tag{4.6}$$

which can also be expressed with a path integral in the complex plane:

$$E_u = \sum_{i=0}^{\infty} \frac{1}{2\pi j} \oint_{|z_1|=1} u_i(z_1^{-1})u_i(z_1)\frac{dz_1}{z_1}. \tag{4.7}$$

If both the integral and the series converge, their order may be interchanged:

$$E_u = \frac{1}{2\pi j} \oint_{|z_1|=1} \sum_{i=0}^{\infty} u_i(z_1^{-1})u_i(z_1)\frac{dz_1}{z_1}. \tag{4.8}$$

The state of the system (4.1)-(4.2) after N samples is given by:

$$x_N(z_1) = \mathscr{C}_N(z_1)\mathscr{U}_N(z_1), \tag{4.9}$$

where

$$\mathscr{C}_N(z_1) := \begin{bmatrix} B(z_1) & A(z_1)B(z_1) & \cdots & A(z_1)^N B(z_1) \end{bmatrix} \tag{4.10}$$

and

$$\mathscr{U}_N(z_1) := \begin{bmatrix} u_{N-1}(z_1) \\ u_{N-2}(z_1) \\ \cdots \\ u_0(z_1) \end{bmatrix}. \tag{4.11}$$

For a given state $x_N(z_1)$, the solution of (4.9) with minimum 2-norm for each ω such that $z_1 = e^{j\omega}$ is

$$\mathscr{U}_N(z_1) = \mathscr{C}_N(z_1)^*(\mathscr{C}_N(z_1)\mathscr{C}_N(z_1)^*)^{-1}x_N(z_1), \tag{4.12}$$

where the operator * applied over a rational matrix $M(z)$ denotes the operation defined by $M(z)^* := M(z^{-1})^T$.

Let $\mathscr{U}_\infty(z_1)$ and $\mathscr{C}_\infty(z_1)$ denote, respectively, $\mathscr{U}_N(z_1)$ and $\mathscr{C}_N(z_1)$ when N is extended to infinity. The minimum control energy required to transfer the system from the zero initial state to a given final state $x^f_\infty(z_1)$ in an arbitrarily large number of samples is

$$\begin{aligned} E_u|_{min} &= \frac{1}{2\pi j}\oint_{|z_1|=1} \mathscr{U}_\infty(z_1^{-1})^T\mathscr{U}_\infty(z_1)\frac{dz_1}{z_1} = \\ &= \frac{1}{2\pi j}\oint_{|z_1|=1} x^f_\infty(z_1)^*(\mathscr{C}_\infty(z_1)\mathscr{C}_\infty(z_1)^*)^{-1}x^f_\infty(z_1)\frac{dz_1}{z_1}. \end{aligned} \tag{4.13}$$

Comparing equations (2.1.2) and (4.13) immediately leads to the following definition.

Definition The *controllability gramian* of a first level realization $\Sigma = (A(z), B(z), C(z), D(z))$ is defined as $P(z) := \mathscr{C}_\infty(z)\mathscr{C}_\infty(z)^*$, where $\mathscr{C}_\infty(z)$ is the controllability matrix extended to infinity.

An analogous reasoning motivates the definition of observability gramian of first level realizations. Start by considering the energy of the output signal:

$$E_y := \sum_{i=0,j=0}^{\infty,\infty} y(i,j)^2 = \sum_{i=0}^{\infty}\frac{1}{2\pi j}\oint_{|z_1|=1} y_i(z_1^{-1})y_i(z_1)\frac{dz_1}{z_1}. \tag{4.14}$$

Defining

$$\mathscr{Y}_N(z_1) := \begin{bmatrix} y_0(z_1) \\ y_1(z_1) \\ \vdots \\ y_{N-1}(z_1) \end{bmatrix} \tag{4.15}$$

and

$$\mathscr{O}_N(z_1) := \begin{bmatrix} C(z_1) \\ C(z_1)A(z_1) \\ \vdots \\ C(z_1)A(z_1)^{N-1} \end{bmatrix}, \tag{4.16}$$

it is easy to verify that, given an initial state $x_z(0)$ and assuming zero input, the sequence of outputs $\mathscr{Y}_N(z_1)$ can be determined through:

$$\mathscr{Y}_N(z_1) = \mathscr{O}_N(z_1)x_0(z_1). \tag{4.17}$$

The energy of the output signal of the system with initial state $x_0(z_1)$ and zero input is, therefore:

$$\begin{aligned} E_y &= \frac{1}{2\pi j}\oint_{|z_1|=1} \mathscr{Y}_\infty(z_1^{-1})^T\mathscr{Y}_\infty(z_1)\frac{dz_1}{z_1} = \\ &= \frac{1}{2\pi j}\oint_{|z_1|=1} (\mathscr{O}_\infty(z_1)x_0(z_1)^T\mathscr{O}_\infty(z_1^{-1})x_0(z_1^{-1})\frac{dz_1}{z_1} = \\ &= \frac{1}{2\pi j}\oint_{|z_1|=1} x_0(z_1^{-1})^*\mathscr{O}_\infty(z_1^{-1})^*\mathscr{O}_\infty(z_1^{-1})x_0(z_1^{-1})\frac{dz_1}{z_1}, \end{aligned} \tag{4.18}$$

where $\mathscr{Y}_\infty(z_1)$ and $\mathscr{O}_\infty(z_1)$ are $\mathscr{Y}_N(z_1)$ and $\mathscr{O}_N(z_1)$, respectively, when N is extended to infinity. The definition of observability gramian of a first level realization is now obvious.

Definition The *observability gramian* of a first level realization $\Sigma = (A(z),B(z),C(z),D(z))$ is defined as $Q(z) := \mathscr{O}_\infty(z^{-1})^*\mathscr{O}_\infty(z^{-1})$, where $\mathscr{O}_\infty(z)$ is the observability matrix extended to infinity.

Remark In the previous definitions the sub-index of z was deliberately omitted. From now on, this option will be kept whenever it does not arise any confusion for the reader. Note that it was always being assumed that the first level realizations depended on z_1, but if z_2 is chosen the same results hold.

Observe that the previous definitions are only formal, i.e., no assumptions or conditions about convergence were stated. However, they are meaningless definitions if $P(e^{j\omega})$ and $Q(e^{j\omega})$ are not well defined for all $\omega \in [-\pi,\pi]$. The following theorem states conditions for that convergence to hold.

Theorem 4.2.1. *If $A(z)$ is such that $\max_{\omega\in[0,\pi]}\rho(A(e^{j\omega})) < 1$, where $\rho()$ denotes the spectral radius, then the series defining $P(e^{j\omega})$ and $Q(e^{j\omega})$ converge for all $\omega \in \mathbb{R}$.*

Proof. Only the case of $P(z)$ is considered, since the proof for $Q(z)$ is completely analogous.
Note that the entries of $A(z)$ and $B(z)$ are defined over $\mathbb{R}_p(z)$, and so these matrices have the property that $A(e^{j\omega})^* = \bar{A}(e^{j\omega})^T$, where $\bar{A}(e^{j\omega})$ denotes the matrix $A(e^{j\omega})$ with conjugate entries. Hence, $\rho(A(e^{j\omega})) = \rho(A(e^{j\omega})^*)$, so, if $\max_{\omega\in[0,\pi]}\rho(A(e^{j\omega})) < 1$ the series

$$\sum_{i=0}^{\infty} A(e^{j\omega})^i B(e^{j\omega})B(e^{j\omega})^*(A(e^{j\omega})^*)^i = \mathscr{C}_\infty(e^{j\omega})\mathscr{C}_\infty(e^{j\omega})^* \tag{4.19}$$

converges for all $\omega \in [-\pi,\pi]$. Due to the periodicity of $A(e^{j\omega})$ and $B(e^{j\omega})$, the result is valid for all $\omega \in \mathbb{R}$. □

Corollary 4.2.2. *If $A(z)$ satisfies the conditions of Theorem 4.2.1, then $P(e^{j\omega}) \geq 0$ and $Q(e^{j\omega}) \geq 0$ for all $\omega \in \mathbb{R}$.*

Although conditions for the convergence of $P(e^{j\omega})$ and $Q(e^{j\omega})$ have been given, the only way one has to compute these matrices is through the series that define them, which is a rather inconvenient way. Moreover, rationality of $P(z)$ and $Q(z)$ is not ensured yet. The following results solve both problems.

Theorem 4.2.3. *If $A(z)$ satisfies the conditions of Theorem 4.2.1, then:*

1. *$(I - A(z^{-1}) \otimes A(z))$, where $\otimes$ denotes the Kronecker product, is a square nonsingular rational matrix;*

2. *The unique solution $X(z)$ of the Lyapunov equation*

$$X(z) = A(z)X(z)A(z)^* + B(z)B(z)^* \tag{4.20}$$

   is a rational matrix such that

$$vec(X(z)) = (I - A(z^{-1}) \otimes A(z))^{-1} vec(B(z)B(z)^*), \tag{4.21}$$

   where $vec()$ denotes de vectorization operation, which transforms a matrix into a vector by stacking its columns;

3. *$X(z)$ coincides with $P(z)$ for $z = e^{j\omega}$.*

Proof.

1. Clearly, $(I - A(z^{-1}) \otimes A(z))$ is a square rational matrix, since $A(z)$ is a square rational matrix. The proof of the nonsingularity follows. Consider a point $z_0 = e^{j\omega_0}$. If $(I - A(z_0^{-1}) \otimes A(z_0))^{-1}$ exists, then $(I - A(z^{-1}) \otimes A(z))$ is nonsingular as a rational matrix. Moreover, $(I - A(z_0^{-1}) \otimes A(z_0))^{-1}$ exists if and only if $A(z_0^{-1}) \otimes A(z_0)$ has no eigenvalues equal to one. Since $A(z)$ satisfies the conditions of Theorem 4.2.1, $\rho(A(z_0)) = \rho(A(z_0)^{-1}) < 1$. It is possible to prove that the spectral radius of a Kronecker product equals the product of the spectral radius of each matrix (see [11]), thus $\rho(A(z_0^{-1}) \otimes A(z_0)) = \rho(A(z_0^{-1}))\rho(A(z_0)) < 1$. Consequently, $(I - A(z_0^{-1}) \otimes A(z_0))^{-1}$ exists and, therefore, $(I - A(z^{-1}) \otimes A(z))$ is a square nonsingular rational matrix.

2. Using the fact that $vec(MNR) = R^T \otimes M vec(N)$ (see [11]), (4.20) is equivalent to

$$(I - A(z^{-1}) \otimes A(z))vec(X(z)) = vec(B(z)B(z)^*). \tag{4.22}$$

   Since $(I - A(z^{-1}) \otimes A(z))$ is an invertible rational matrix, the solution $vec(X(z)$ is uniquely given by (4.21), proving that the solution $X(z)$ of (4.20) is unique.

3. It is enough to recall the proof of 2.1.8, since this one follows the same reasoning.

□

For the observability gramian analogous results hold.

Theorem 4.2.4. *If $A(z)$ satisfies the conditions of Theorem 4.2.1, then:*

1. *$(I - A(z^{-1}) \otimes A(z))$ is a square nonsingular rational matrix;*
2. *The unique solution $Y(z)$ of the Lyapunov equation*

$$Y(z) = A(z^{-1})^* Y(z) A(z^{-1}) + C(z^{-1})^* C(z^{-1}) \tag{4.23}$$

is a rational matrix such that

$$vec(Y(z)) = (I - A(z)^* \otimes A(z^{-1})^*) vec(C(z^{-1})^* C(z^{-1})); \tag{4.24}$$

3. *$Y(z)$ coincides with $Q(z)$ for $z = e^{j\omega}$.*

Proof. Analogous to the proof of 4.2.3. □

Remark In fact, the solutions $X(z)$ and $Y(z)$ of these Lyapunov equations coincide with the respective gramians wherever $P(z)$ and $Q(z)$ are well defined. Thus, $X(z)$ and $Y(z)$ can be seen as analytic continuations of $P(z)$ and $Q(z)$, respectively. For this reason and also because $z = e^{j\omega}$ is actually the only region where these gramians have some meaning, from now on, $P(z)$ and $Q(z)$ are redefined as the solutions $X(z)$ of (4.20) and $Y(z)$ of (4.23) themselves. The assumption that $A(z)$ satisfies the conditions of Theorem 4.2.1 and thus these gramians are well defined in $z = e^{j\omega}$ will hold for the remaining of the discussion.

The notions of controllability and observability gramians are now clear for first level realizations. The next step would be performing a change of basis such that these gramians become equal and diagonal. For this purpose, it is necessary to understand how rational matrices can be factorized. This problem motivates the discussion of the next subsection.

4.2.2 Factorization of rational matrices

In this subsection, two types of factorization for rational matrices will be presented: spectral factorization and a pseudo singular value decomposition. Conditions for the existence of a spectral factorization and an algorithm to obtain such factorization will be given.
Before defining both factorizations, the notions of *pole* and *zero* of a rational matrix must be clarified.

Definition A rational matrix $G(z)$ has a *pole* at $z_0 \in \mathbb{C}$ if at least one of its entries has a pole at z_0.

Definition A rational matrix $G(z)$ has a *zero* at $z_0 \in \mathbb{C}$ if the rank of $G(z_0)$ over $\mathbb{C}$ is strictly less than the rank of $G(z)$ over $\mathbb{R}(z)$.

Now, two particularly relevant types of rational matrices may be defined.

Definition A square matrix $B(z)$ over $\mathbb{R}(z)$ is said to be *inner* if $B(z)B(z)^* = I$.

Definition A full column rank matrix $M(z)$ over $\mathbb{R}(z)$ is said to be *outer* if its poles and zeros are all located inside the closed unit circle (i.e., boundary included).

4.2.2.1 Spectral factorization

If one takes any rational matrix $M(z)$, the matrix $\Phi(z)$ defined by $\Phi(z) = M(z)M(z)^*$ is a rational matrix satisfying the following two properties:

$$\Phi(z) = \Phi(z)^*, \tag{4.25}$$

$$\Phi(e^{j\omega}) \geq 0 \; \forall \omega \in \mathbb{R}. \tag{4.26}$$

The spectral factorization problem concerns with the reversal of this construction: given a rational matrix $\Phi(z)$ satisfying (4.25) and (4.26), find an $M(z)$ such that $\Phi(z) = M(z)M(z)^*$.
Before treating the matrix case, it will be useful to solve this problem for scalar rational functions.

Theorem 4.2.5. *For all rational function $g(z)$ that satisfies $g(z) = g(z^{-1})$ for all $z \in \mathbb{C}$ and $g(e^{j\omega}) \geq 0$ for all $\omega \in \mathbb{R}$, there exists a rational function $f(z)$, having all its poles and zeros inside the closed unit circle, such that $g(z) = f(z)f(z^{-1})$.*

Proof. The proof will be given through the next algorithm. □

The following algorithm describes how to compute the spectral factorization of a function that satisfies the conditions established in Theorem 4.2.5.

Algorithm 4. *Spectral factorization of a rational function*
Input data: rational function $g(z)$ satisfying the conditions of Theorem 4.2.5.

1. *Write a factorization of $g(z)$ into polynomials*

$$g(z) = \gamma z^l \frac{p_i(z)p_c(z)p_o(z)}{q_i(z)q_c(z)q_o(z)}, \tag{4.27}$$

where γ is a real constant, $l \in \mathbb{Z}$, $p_i(z)$ and $q_i(z)$ are polynomials that have all their zeros inside the open unit circle except the origin, $p_o(z)$ and $q_o(z)$ have all their zeros outside the closed unit circle and $p_c(z)$ and $q_c(z)$ have all their zeros on the boundary of the unit circle. Note that, due to the positivity condition, γ must be nonnegative and all zeros of $p_c(z)$ and $q_c(z)$ must have even multiplicity. Furthermore, the condition $g(z) = g(z^{-1})$ implies that $p_i(\alpha) = 0$ if and only if $p_o(\alpha^{-1}) = 0$, where α is any nonzero complex number. The same holds for $q_i(\alpha)$ and $q_o(\alpha)$.

Thus, $p_i(z)$, $p_c(z)$ and $p_o(z)$ can be written as

$$p_i(z) = \prod_{j=0}^{K-1} (a_j z + 1)^{k_j}, \ |a_j| > 1, \tag{4.28}$$

$$p_c(z) = \prod_{j=0}^{W-1} (z + c_j)^{2w_j}, \ |c_j| = 1, \tag{4.29}$$

$$p_o(z) = \prod_{j=0}^{K-1} (z + a_j)^{k_j}, \tag{4.30}$$

and the analogous holds for $q_i(z)$, $q_c(z)$ and $q_o(z)$.

2. *Define $\tilde{p}_c(z)$ as*

$$\tilde{p}_c(z) = \prod_{j=0}^{W-1} (z + c_j)^{w_j} \tag{4.31}$$

and $\tilde{q}_c(z)$ analogously.

3. *Build $f(z)$ as*

$$f(z) = \sqrt{\gamma} z^l \frac{p_i(z)\tilde{p}_c(z)}{q_i(z)\tilde{q}_c(z)}. \tag{4.32}$$

One has $g(z) = f(z)f(z^{-1})$.

Now, one can turn to the matrix case.

Theorem 4.2.6. *Let $\Phi(z)$ be an $n \times n$ rational matrix of rank k and suppose that it satisfies (4.25) and (4.26). Under these conditions, there exists an outer rational matrix $M(z)$ of size $n \times k$ such that $\Phi(z) = M(z)M(z)^*$.*

Proof. Once again, the sketch of the proof will be given through the following algorithm. □

Algorithm 5. *Spectral factorization of a rational matrix*
Input data: rational matrix $\Phi(z)$ satisfying the conditions of Theorem 4.2.6.

1. *If $\Phi(z)$ is identically zero, the problem is already solved. If it is not, take a nonzero element of the diagonal of $\Phi(z)$ and, by row and column permutations, place it in the position $(1,1)$. These permutations correspond to building a matrix $\bar{\Phi}(z)$ such that*

$$\bar{\Phi}(z) = U\Phi(z)U, \tag{4.33}$$

where U is a constant and orthogonal matrix.
The existence of a nonzero element in the diagonal of $\Phi(z)$ is ensured because $\Phi(e^{j\omega})$ is a nonnegative hermitian matrix.

2. *Compute the spectral factorization of the first element, $\bar{\phi}_{11}(z)$, of $\bar{\Phi}(z)$. Say, $\bar{\phi}_{11}(z) = f_{11}(z)f_{11}(z^{-1})$.*

3. *Build the matrix* $\bar{\bar{\Phi}}(z)$ *as*

$$\bar{\bar{\Phi}}(z) := \begin{bmatrix} f_{11}(z) & 0 \\ 0 & I \end{bmatrix}^{-1} \bar{\Phi}(z) \begin{bmatrix} f_{11}(z^{-1}) & 0 \\ 0 & I \end{bmatrix}^{-1}. \tag{4.34}$$

Clearly, the first element of $\bar{\bar{\Phi}}(z)$ *is equal to 1.*

4. *Define* $R(z)$ *as*

$$R(z) := \begin{bmatrix} 1 & 0 \\ -\bar{\bar{\Phi}}_{21}(z) & I \end{bmatrix}, \tag{4.35}$$

where $\bar{\bar{\Phi}}_{21}(z) := \begin{bmatrix} \bar{\bar{\phi}}_{21}(z) \\ \vdots \\ \bar{\bar{\phi}}_{n1}(z) \end{bmatrix}$.

5. *Compute* $R(z)\bar{\bar{\Phi}}(z)R(z)^*$ *to obtain:*

$$R(z)\bar{\bar{\Phi}}(z)R(z)^* = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{\Phi}_{22}(z) \end{bmatrix}. \tag{4.36}$$

6. *Repeat steps 1-5 over* $\tilde{\Phi}_{22}(z)$ *and so on until getting the identity matrix. Then, the inverse transformations must be performed to obtain the spectral factor.*

Both algorithms presented in this section were adapted from proofs of theorems in [12]. There, a different form of spectral factorization is considered, which is more suitable for continuous time transfer matrices. An example of a spectral factorization of a rational matrix will be given later on.

4.2.2.2 Pseudo Singular Value Decomposition

In this subsection, another form of factorization for rational matrices is presented. This was firstly proposed by Tsai *et al.* in [26].

Definition A *pseudo singular value decomposition* (PSVD) of a matrix $\Phi(z) \in \mathbb{R}(z)^{k\times n}$, $k \geq n$, is a factorization

$$\Phi(z) = W(z) \begin{bmatrix} S(z) \\ 0 \end{bmatrix} V(z)^* \tag{4.37}$$

, where $W(z)$, $S(z)$ and $V(z)$ are defined over $\mathbb{R}(z)$ and have dimensions $k \times k$, $n \times n$ and $n \times n$, respectively. Furthermore, $W(z)$ and $V(z)$ are inner, $S(z) = diag(\sigma_1(z), \cdots, \sigma_n(z))$ and $|\sigma_1(e^{j\omega})| \geq \cdots \geq |\sigma_n(e^{j\omega})|$, $\forall \omega \in \mathbb{R}$.

In the following example, a matrix that has a PSVD is given.

Example Consider the following matrix:

$$\Phi(z) = \begin{bmatrix} \frac{2+32z}{-3-6z+24z^2} & \frac{2(1+2z)(5+8z)}{3(2+z)(4+z)(-1+2z)} \\ \frac{2(2+z)(4+z)(5+8z)}{3(1+4z)^2(-1+4z^2)} & \frac{2+32z}{-3-6z+24z^2} \end{bmatrix}. \tag{4.38}$$

This matrix can be factored as:

$$\Phi(z) := W(z)S(z)W(z)^*, \tag{4.39}$$

where

$$W(z) = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1+2z}{2+z} & -\frac{(1+2z)(1+3z)(1+4z)}{(2+z)(3+z)(4+z)} \\ \frac{4+z}{1+4z} & \frac{1+3z}{3+z} \end{bmatrix}, \tag{4.40}$$

$$S(z) = \begin{bmatrix} \frac{4}{-1+2z} & 0 \\ 0 & \frac{8}{3+12z} \end{bmatrix}. \tag{4.41}$$

It is straightforward to check that $W(z)W(z)^* = I$. Furthermore, it is also not hard to verify that

$$\left| \frac{4}{-1+2e^{j\omega}} \right| \geq \left| \frac{8}{3+12e^{j\omega}} \right|, \ \forall \omega \in \mathbb{R}. \tag{4.42}$$

Thus, this factorization is a PSVD of $\Phi(z)$.

This is a natural extension to rational matrices of the traditional singular value decomposition. However, unfortunately, such factorization is not always possible, i.e., an arbitrary $\Phi(z) \in \mathbb{R}(z)$ may not have a PSVD. The structure of the class of rational matrices that have such factorization is also not easy to identify, therefore there is no algorithm to obtain the PSVD of a given rational matrix, if it exists. These problems make the task of finding a rational balanced realization very hard to solve.

4.2.3 Towards balancing first level realizations

The expression *rational balanced realization* was already used a couple of times previously. Although the definition might be rather intuitive, it was not conveniently stated yet.

Definition A minimal first level model $\Sigma = (A(z), B(z), C(z), D(z))$, of dimension n_{min}, is a *rational balanced realization* if

$$P(z) = Q(z) = S(z) := diag(\sigma_1(z), \cdots, \sigma_{n_{min}}(z)), \tag{4.43}$$

$$|\sigma_1(e^{j\omega})| \geq \cdots \geq |\sigma_{n_{min}}(e^{j\omega})|, \ \forall \omega \in \mathbb{R}, \tag{4.44}$$

where $P(z)$ and $Q(z)$ are, respectively, the controllability and observability gramians of Σ.

Now, it is clear that if one can find a rational balanced realization, then one can find a reduced order first level realization by simply truncating matrices $A(z)$, $B(z)$ and $C(z)$ of such realization.

Another possibility would be truncating the matrices of a realization in which one of the gramians was the identity matrix and the other was in the form of the $S(z)$ defined above.
From what was stated in the previous sections, it should be clear that $P(z)$ and $Q(z)$ always have a spectral factorization. Moreover, it is easy to show that when a transformation $T(z) \in \mathbb{R}(z)^{n\times n}$ is applied over a realization $\Sigma = (A(z),B(z),C(z),D(z))$ as defined in (3.35), the gramians $\overline{P}(z)$ and $\overline{Q}(z)$ of the obtained realization will be related with the original ones, $P(z)$ and $Q(z)$, by:

$$\overline{P}(z) = T(z)^{-1}P(z)(T(z)^*)^{-1}, \tag{4.45}$$

$$\overline{Q}(z) = T(z)^*Q(z)T(z). \tag{4.46}$$

Now suppose that a spectral factorization of $P(z)$ is computed, say $P(z) = M(z)M(z)^*$. Defining $T(z) := M(z)$ clearly yields a $\overline{P}(z) = I$. The problem here is that the quadruple of matrices $(\overline{A}(z),\overline{B}(z),\overline{C}(z),D(z))$, as defined in (3.35), must be indeed a realization, i.e., all matrices must be defined over $\mathbb{R}_p(z)$. If $M(z)$ is unimodular as an element of $\mathbb{R}_p^{n\times n}(z)$, such requirement is obviously satisfied. For a particular class of models, and considering an approximation of these gramians, it is possible to show that $M(z)$ is indeed a proper rational matrix whose inverse is also proper.
Consider the following class of first level realizations $\Sigma = (A(z),B(z),C(z),D(z))$:

$$A(z) = \begin{bmatrix} 0 & \cdots & 0 & -\alpha_0(z) \\ 1 & \cdots & 0 & -\alpha_1(z) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & -\alpha_{n-1} \end{bmatrix}, \quad B(z) = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{4.47}$$

and $C(z)$ and $D(z)$ are arbitrary matrices over $\mathbb{R}_p(z)$ with adequate dimensions. Note that the controllability matrix of this model equals the identity matrix, thus a realization of this type is always controllable. If the system is SISO, it is always possible to find a minimal realization with this structure. Therefore, from now on, the discussion is restricted to SISO systems. Moreover, it is assumed that that $A(z)$ satisfies the conditions of Theorem 4.2.1 and that the $\alpha_i(z)$ are strictly proper rational functions with no zeros or poles at the origin. This class of models Σ will be denoted by Ψ. Starting from a model of this type, it will be shown how to obtain a realization whose *truncated* controllability gramian is the identity matrix. The same can be done for the *truncated* observability gramian, but a different structure has to be considered for the initial model. For a better understanding of the proofs of the following results, the reader is referred to Appendix B, where some algebraic properties on the relative degree of rational functions are enumerated.

Lemma 4.2.7. *For every $N \geq n$, the truncated controllability gramian $P_N(z) := \mathscr{C}_N(z)\mathscr{C}_N(z)^*$ of a realization that belongs to the class Ψ is a proper rational matrix with rank n.*

Proof. If a proper rational function $g(z)$ has no poles or zeros at the origin, then it is easy to verify that $g(z^{-1})$ is a biproper rational function (see property 4. in Appendix B). Thus, all entries of $A(z^{-1})$ are constant or biproper rational functions.

$A(z)$ and $B(z)$ as defined in (4.47) yield

$$\mathscr{C}_N(z) = \begin{bmatrix} I & A(z)^n & A(z)^{2n} & \cdots & A(z)^{N-n} \end{bmatrix}, \tag{4.48}$$

where it is assumed that N is a multiple of n. If this is not true, $\mathscr{C}_N(z)$ has an identical structure, but the last block, $A(z)^{N-n}$, is truncated. In any case, the same results hold.
Since the sum and the product of proper rational functions is a proper rational function (see properties 1. and 2. in Appendix B), every entry of $\mathscr{C}_N(z)$ and $\mathscr{C}_N(z)^*$ is proper. Hence, $P_N(z)$ is itself a proper rational matrix. Finally, $P_N(z)$ has rank n due to the controllability of the realization. □

Lemma 4.2.8. *The truncated controllability gramian $P_N(z)$, $N \geq n$, of a realization in Ψ is unimodular as an element of $\mathbb{R}_p(z)^{n\times n}$.*

Proof. Note that every entry of $A(z)^k$ is either zero or a strictly proper rational function for all integer $k > 1$. Moreover, $P_N(z) = I + \sum_{i=1}^{(N-n)/n} A(z)^{ni} A(z)^{*ni}$. Therefore, the entries in the diagonal of $P_N(z)$ are of the form $d_j(z) = 1 + g_j(z)$, $j = 1, \cdots, n$ where the $g_j(z)$ are strictly proper rational functions. Thus, according to property 2. in Appendix B, every $d_j(z)$ is a biproper rational function.
On the other hand, again because $A(z)^k$, $k > 1$, is strictly proper, every entry of $P_N(z)$ that does not belong to its diagonal is either strictly proper or zero.
Hence, $det(P_N(z))$ can be written as a sum of products where each of them is strictly proper except one, which is $\prod_{j=1}^{n} d_j$, that is clearly biproper (property 1. in Appendix B). Consequently, $det(P_N(z))$ is a biproper rational function, so $P_N(z)^{-1} = \frac{1}{det(P_N(z))} P_N(z)^{adj}$ is necessarily a proper rational matrix. □

Note that, although proving the unimodularity of a $P_N(z)$ is a relevant step, the goal is to prove the unimodularity of a matrix $M(z)$ such that $P_N(z) = M(z)M(z)^*$. The following lemma will be very useful to complete that task.

Lemma 4.2.9. *Let $g(z)$ be a biproper rational function that, in addition, satisfies the conditions of Theorem 4.2.5. Then, there exists a biproper rational function $f(z)$ such that $g(z) = f(z)f(z^{-1})$.*

Proof. If $g(z)$ is constant the problem is trivial, so assume that is not the case. Recall the factorization of $g(z)$ suggested in the step 1 of Algorithm 4:

$$g(z) = \gamma z^l \frac{p_i(z)p_c(z)p_o(z)}{q_i(z)q_c(z)q_o(z)}, \tag{4.49}$$

where γ, l, $p_i(z)$, $p_c(z)$, $p_o(z)$, $q_i(z)$, $q_c(z)$ and $q_o(z)$ are as defined there. Recall that $deg(p_c(z))$ and $deg(q_c(z))$ are always even and let $n_i := deg(p_i(z))$, $2n_c := deg(p_c(z))$, $m_i := deg(q_i(z))$ and $2m_c := deg(p_c(z))$. Then, $deg(p_o(z)) = n_i$ and $deg(q_o(z)) = m_i$. The biproperness of $g(z)$ yields

$$l + 2n_i + 2n_c = 2m_i + 2m_c. \tag{4.50}$$

Consider the $f(z)$ defined in the step 3 of that algorithm:

$$f(z) = \sqrt{\gamma} z^l \frac{p_i(z)\tilde{p}_c(z)}{q_i(z)\tilde{q}_c(z)}, \tag{4.51}$$

where $\tilde{p}_c(z)$ and $\tilde{q}_c(z)$ are as defined there. Note that $f(z)$ is biproper if and only if

$$l + n_i + n_c = m_i + m_c. \tag{4.52}$$

Since the term z^l represents all poles or zeros of $g(z)$ at the origin, $p_i(z^{-1})$, $\tilde{p}_c(z^{-1})$, $q_i(z^{-1})$ and $\tilde{q}_c(z^{-1})$ are biproper rational functions. Thus, the difference between the degree of the numerator and the degree of the denominator in $f(z^{-1})$ equals $-l$.
Given that $g(z) = f(z)f(z^{-1})$ and using again the fact that $g(z)$ is biproper, one concludes that the following equality must hold:

$$(l + n_i + n_c) - (m_i + m_c) + (-l) = 0. \tag{4.53}$$

Equations (4.50) and (4.53) imply that $n_i + n_c = m_i + m_c$ and $l = 0$, thus equality (4.52) is verified. □

Now, the main result of this section may be presented.

Theorem 4.2.10. *The truncated controllability gramian $P_N(z)$, $N \geq n$, of a realization in Ψ may be factored as $P_N(z) = M(z)M(z)^*$, where $M(z)$ is unimodular as an element of $\mathbb{R}_p(z)^{n\times n}$.*

Proof. It is straightforward to check that $P_N(z) = P_N(z)^*$. Since $P_N(e^{j\omega}) := \mathscr{C}_N(e^{j\omega})\mathscr{C}_N(e^{j\omega})^*$ and $\mathscr{C}_N(e^{j\omega})^*$ coincides with the hermitian transpose of $\mathscr{C}_N(e^{j\omega})$, one concludes that $P_N(e^{j\omega}) \geq 0$. Hence, $P_N(z)$ has a spectral factorization $P_N(z) = M(z)M(z)^*$. The spectral factor $M(z)$ belongs to $\mathbb{R}_p^{n\times n}(z)$, because $rank(P_N(z)) = n$.
Lemma 4.2.8 ensures that $P_N(z)$ is unimodular as an element of $\mathbb{R}_p^{n\times n}(z)$, thus $det(P_N(z))$ is a biproper rational function. Note that

$$\begin{aligned} det(P_N(z)) &= det(M(z)M(z)^*) \\ &= det(M(z))det(M(z^{-1})). \end{aligned} \tag{4.54}$$

Lemma 4.2.9 ensures the biproperness of $det(M(z))$, thus $M(z)$ is unimodular as an element of $\mathbb{R}_p(z)^{n\times n}$. □

Conjecture Lemmas 4.2.7 and 4.2.8 and Theorem 4.2.10 hold if one uses the actual gramian, $P(z)$, instead of the truncated one. This is supported by the following example.

Example Consider a realization $\Sigma = (A(z), B(z), C(z), D(z)) \in \Psi$, where $A(z)$ is:

$$A(z) := \begin{bmatrix} 0 & \frac{1}{3z+1} \\ 1 & \frac{1}{4z+2} \end{bmatrix}. \tag{4.55}$$

It is shown next how to compute a spectral factorization for the controllability gramian, $P(z)$, of this realization and the unimodularity of such spectral factor is verified.
It is possible to check that $\rho(A(e^{j\omega})) < 1$, $\forall \omega \in \mathbb{R}$. Therefore, one can obtain $P(z)$ by solving equation (4.20). This yields:

$$P(z) = \begin{bmatrix} \frac{144+1602z+7208z^2+16927z^3+22398z^4+16927z^5+7208z^6+1602z^7+144z^8}{144+1554z+6824z^2+15763z^3+20730z^4+15763z^5+6824z^6+1554z^7+144z^8} & \frac{60z+446z^2+1204z^3+1482z^4+868z^5+236z^6+24z^7}{144+1554z+6824z^2+15763z^3+20730z^4+15763z^5+6824z^6+1554z^7+144z^8} \\ \frac{24z+236z^2+868z^3+1482z^4+1204z^5+446z^6+60z^7}{144+1554z+6824z^2+15763z^3+20730z^4+15763z^5+6824z^6+1554z^7+144z^8} & \frac{144+1632z+7476z^2+17796z^3+23664z^4+17796z^5+7476z^6+1632z^7+144z^8}{144+1554z+6824z^2+15763z^3+20730z^4+15763z^5+6824z^6+1554z^7+144z^8} \end{bmatrix}. \quad (4.56)$$

The step 1 of Algorithm 5 may be skipped because the first entry, $p_{11}(z)$, of $P(z)$ is nonzero. Thus, the first step to obtain the spectral factorization of this $P(z)$ is computing the spectral factorization of $p_{11}(z)$. That entry has zeros at

$$\begin{aligned} \{-1/3,\ -0.353972,\ -0.528437 - j0.134108,\ -0.528437 + j0.134108, \\ -1.77787 - j0.451191,\ -1.77787 + j0.451191,\ -3\} \end{aligned} \quad (4.57)$$

and poles at

$$\begin{aligned} \{-0.355733 - j0.0240895,\ -0.355733 + j0.0240895, \\ -0.546348 - j0.15408,\ -0.546348 + j0.15408, \\ -1.69549 - j0.478158,\ -1.69549 + j0.478158, \\ -2.79827 - j0.189493,\ -2.79827 + j0.189493\}. \end{aligned} \quad (4.58)$$

Note that, since the numerator and the denominator of $p_{11}(z)$ have degree 8, numerical approximations had to be made in the computation of their roots.
Applying Algorithm 4, one immediately obtains the spectral factor $f_1(z)$ of $p_{11}(z)$:

$$f_1(z) = \sqrt{0.856114}\frac{(1+3.55574z+3.36439z^2)(1+2.282508z)(1+3z)}{(1+3.39097z+3.10331z^2)(1+5.59653z+7.8662z^2)}. \quad (4.59)$$

Now, define $S_1(z)$ as

$$S_1(z) := \begin{bmatrix} f_1(z) & 0 \\ 0 & 1 \end{bmatrix}. \quad (4.60)$$

Building $\bar{\bar{\Phi}}(z)$ as in the step 3 of Algorithm 4, gives:

$$\bar{\bar{\Phi}}(z) := S_1(z)^{-1} P(z) S_1(z)^{*-1} = \begin{bmatrix} 1 & \bar{\bar{\phi}}_{21}(z^{-1}) \\ \bar{\bar{\phi}}_{21}(z) & p_{22}(z) \end{bmatrix}, \quad (4.61)$$

where $p_{22}(z)$ is the last entry of $P(z)$ and

$$\bar{\bar{\phi}}_{21}(z) = \frac{1.46572z+16.5688z^2+75.2888+z^3+179.272+z^4+247.463+z^5+208.552+z^6+108.988+z^7+34.476z^8+6.0437z^9+0.450322z^{10}}{9.50468+115.981z+601.509z^2+1745.76+z^3+3149.34z^4+3716.73z^5+2946.34z^6+1580.48z^7+566.062z^8+129.658z^9+17.1725z^{10}+z^{11}}. \quad (4.62)$$

Define $R(z)$ as in the step 4:

$$R(z) := \begin{bmatrix} 1 & 0 \\ -\bar{\bar{\phi}}_{21}(z) & 1 \end{bmatrix} \quad (4.63)$$

and compute $\bar{\bar{\Phi}}(z)$ as:

$$\bar{\bar{\Phi}}(z) := R(z)\bar{\bar{\Phi}}(z)R(z)^* = \begin{bmatrix} 1 & 0 \\ 0 & p_{22}(z) \end{bmatrix}. \tag{4.64}$$

Now, one has to compute the spectral factorization of $p_{22}(z)$. The zeros of this function are

$$\{1/3,\ 1/2,\ 1/2,\ 0.381966,\ 2,\ 2,\ 2.61803,\ 3\} \tag{4.65}$$

and the poles coincide with the poles of $p_{11}(z)$, given in (4.58). Therefore, the spectral factor of $p_{22}(z)$ is:

$$f_2(z) = \sqrt{0.77702}\frac{(2z+1)^2(3z+1)(2.61803z+1)}{(1+3.39097z+3.10331z^2)(1+5.59653z+7.8662z^2)}. \tag{4.66}$$

Defining

$$S_2(z) := \begin{bmatrix} 1 & 0 \\ 0 & f_2(z) \end{bmatrix}, \tag{4.67}$$

the following equality obviously holds:

$$S_2(z)^{-1}\bar{\bar{\Phi}}(z)(S_2(z)^*)^{-1} = I. \tag{4.68}$$

The work to obtain the spectral factor of $P(z)$ is essentially done. From (4.68) and from the definition of $\bar{\bar{\Phi}}(z)$, in (4.64), one has:

$$S_2(z)S_2(z)^* = R(z)\bar{\bar{\Phi}}(z)R(z)^*. \tag{4.69}$$

Thus,

$$\bar{\bar{\Phi}}(z) = R(z)^{-1}S_2(z)S_2(z)^*(R(z)^*)^{-1}. \tag{4.70}$$

Replacing this equality in the definition of $\bar{\bar{\Phi}}(z)$, in (4.61), and doing some straightforward manipulation yields:

$$P(z) = S_1(z)R(z)^{-1}S_2(z)(S_1(z)R(z)^{-1}S_2(z))^*. \tag{4.71}$$

Thus, $M(z)$, the spectral factor of $P(z)$, is:

$$M(z) = S_1(z)R(z)^{-1}S_2(z) \tag{4.72}$$

and its entries $m_{ij}(z)$ are:

$$\begin{aligned}
m_{11}(z) &= \frac{0.925264+8.67974z+30.1193z^2+46.0167z^3+26.383z^4}{1+8.98751z+29.9472z^2+44.0418z^3+24.4113z^4}, \\
m_{12}(z) &= 0, \\
m_{21}(z) &= \frac{4.39717z+51.172z^2+242.435z^3+613.106z^4+921.662z^5+873.119z^6+535.517z^7+212.416z^8+52.607z^9+7.39466z^{10}+0.450322z^{11}}{28.514+357.448z+1920.51z^2+5838.79z^3+11193.8z^4+14299.5z^5+12555.8z^6+7687.77z^7+3278.66z^8+955.037z^9+181.176z^{10}+20.1725z^{11}+z^{12}}, \\
m_{22}(z) &= \frac{0.881489+8.47819z+30.2582z^2+47.5021z^3+27.6932z^4}{1.+8.98751z+29.9472z^2+44.0418z^3+24.4113z^4}.
\end{aligned} \tag{4.73}$$

Clearly, $M(z)$ is a proper rational matrix. The unimodularity of $M(z)$ as an element of $\mathbb{R}_p^{2\times 2}(z)$ can

be checked by calculating its determinant:

$$det(M(z)) = \frac{0.81561 + 15.4957z + 128.135z^2 + 602.505z^3 + 1762.68z^4 + 3287.16z^5 + 3818.29z^6 + 2527.6z^7 + 730.63z^8}{1 + 17.975z + 140.67z^2 + 626.385z^3 + 1737.31z^4 + 3076.65z^5 + 3401.78z^6 + 2150.23z^7 + 595.909z^8}, \quad (4.74)$$

which is a biproper rational function. Hence, $M(z)$ has a proper inverse.

Remark Clearly, the computations involved in the previous example are rather laborious and they cannot be effectively performed in most of the numerical software applications, since they involve operations with symbolic variables. However, Wolfram *Mathematica* 9, for instance, provides all calculation resources required for this task.

As a consequence of Theorem 4.2.10, one is able to obtain a first level realization whose truncated controllability gramian equals the identity matrix, when the original model is in Ψ. Unfortunately, this is the closest to a balanced realization one could get until now. Ensuring the existence of a PSVD for the gramians of a known class of realizations would solve the problem. In fact, in that case, one could immediately extend Algorithm 1 for realizations defined over the ring $\mathbb{R}_p(z)$. Recall that, in that algorithm, one applies successively two similarity transformations, $T_1 := US_q^{-1/2}$ and $T_2 = VS_p^{1/4}$. If one applies only T_1, what one gets is precisely a realization where one of the gramians is the identity matrix. Therefore, in this section, that first similarity transformation T_1 was extended to (a particular class of) realizations defined over $\mathbb{R}_p(z)$. Ensuring the existence and being able to compute the PSVD of the realization that is obtained after applying $T_1(z)$, would allow the construction of a $T_2(z)$ which would have here exactly the same role that T_2 has in Algorithm 1. Hence, obtaining a rational balanced realization would be possible.
Furthermore, even if one could find a rational balanced realization and, therefore, perform the order reduction of the first level realization, the dimension of the resulting second level realization could be very big. As shown before, the order of the second level realization varies with the basis where the corresponding first level realization is expressed, i.e., the McMillan degree of the transfer matrix

$$\begin{bmatrix} T(z_1)^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A(z_1) & B(z_1) \\ C(z_1) & D(z_1) \end{bmatrix} \begin{bmatrix} T(z_1) & 0 \\ 0 & I \end{bmatrix} \quad (4.75)$$

depends on the matrix $T(z)$ that is chosen. Thus, performing the changes of basis that would be necessary to obtain the rational balanced realization could considerably increase the McMillan degree of that transfer matrix. However, if the purpose was actually reducing the order of the first level realization, neglecting the dimension of the resulting 2-D model, this problem would be irrelevant.

4.3 Order reduction of second level realizations

In this section, two algorithms for 2-D model order reduction will be presented. Both use the concept of balanced realization, but while the first one tries to extend that definition to 2-D systems, the second one simply uses the conventional 1-D notion.

4.3.1 Beck-Doyle algorithm

The algorithm presented in this subsection is proposed by C.Beck and J.Doyle in [8]. There, the subject is exposed in a more general framework, since it applies to both multidimensional and uncertain systems. Here, the focus will be exclusively on two-dimensional systems, but the generalization to the n-D case is straightforward. The case of uncertain systems can be treated within the framework of n-D systems, as it will be illustrated later on.
The following definition plays a major role in this approach.

Definition A 2-D model $\Sigma = (A,B,C,D)$ is a *balanced realization* if there exist $P \geq 0$ and $Q \geq 0$ that satisfy the linear matrix inequalities (LMI)

$$P \geq APA^T + BB^T, \tag{4.76}$$

$$Q \geq A^T QA + C^T C, \tag{4.77}$$

and such that

$$P = Q = S := diag(S_2, S_1), \tag{4.78}$$

where the $S_i \geq 0$ are diagonal and $dim(S_i) = n_i$.

The only relevant difference of this definition relatively to the 1-D case is the fact that inequalities are used instead of equalities. As a consequence, for a given system, neither the balanced realization nor the balanced S are unique. The following theorem gives conditions for the existence of a balanced realization.

Theorem 4.3.1. *[8] Let $\Sigma = (A,B,C,D)$ be a 2-D model and define Z as in (3.29), $\mathscr{T}$ as in (3.32) and $\bar{U}$ as in (3.44). There exist $P \geq 0$ and $Q \geq 0$, both in $\mathscr{T}$, satisfying (4.76) and (4.77), respectively, if and only if $(Z-A)$ is nonsingular in $\bar{U}$.*

This theorem guarantees the existence of balanced realizations for any internally stable system. In fact, if one finds solutions P and Q for the LMI above that belong to $\mathscr{T}$, one can find an admissible similarity transformation T (i.e., a $T \in \mathscr{T}$) that transforms the original realization into a balanced one. This procedure will be explained in detail later on.
Given a balanced realization, the way of finding the reduced order model is analogous to the 1-D case. Admit that $\Sigma = (A_{11},A_{12},A_{21},A_{22},B_1,B_2,C_1,C_2,D)$ is a balanced realization. One has $P = Q = S = diag(S_2,S_1)$, where

$$S_i = diag(\sigma_{1i}, ..., \sigma_{r_i i}, \sigma_{(r_i+1)i}, ..., \sigma_{n_i i}). \tag{4.79}$$

Suppose that one desires an approximate model with dimension r_1 in the vertical state and r_2 in the horizontal state. Then, one can build matrices $\hat{A}_{11}$ and $\hat{A}_{22}$ by extracting the upper $r_2 \times r_2$ and $r_1 \times r_1$ blocks of matrices A_{11} and A_{22}, respectively. Matrices $\hat{B}_1$, $\hat{B}_2$, $\hat{C}_1$ and $\hat{C}_2$ are built by truncating B_1, B_2, C_1 and C_2 accordingly. The model $\Sigma_r = (\hat{A}_{11},\hat{A}_{12},\hat{A}_{21},\hat{A}_{22},\hat{B}_1,\hat{B}_2,\hat{C}_1,\hat{C}_2,D)$

will be that desired approximate model. Moreover, it has the properties stated in the following theorem.

Theorem 4.3.2. *Suppose that* $\Sigma = (A_{11},A_{12},A_{21},A_{22},B_1,B_2,C_1,C_2,D)$ *is a balanced realization, with transfer matrix* $G(z_1,z_2)$*, whose* $S_i = diag(\sigma_{1i},...,\sigma_{r_i i},\sigma_{(r_i+1)i},...,\sigma_{n_i i})$*. Then, the model* $\Sigma_r = (\hat{A}_{11},\hat{A}_{12},\hat{A}_{21},\hat{A}_{22},\hat{B}_1,\hat{B}_2,\hat{C}_1,\hat{C}_2,D)$*, where matrices are as defined above, with transfer matrix* $G_r(z_1,z_2)$*, has the following properties:*

1. $(\hat{Z}-\hat{A})$*, where* $\hat{Z} = \begin{bmatrix} z_2 I_{r_2} & 0 \\ 0 & z_1 I_{r_1} \end{bmatrix}$ *and* $\hat{A} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}$*, is nonsingular in* $\bar{U}$*;*
2. $\|G(z_1,z_2) - G_r(z_1,z_2)\|_{\mathscr{H}_\infty} := \max_{\forall(\omega_1,\omega_2)\in\mathbb{R}^2} \|G(e^{j\omega_1},e^{j\omega_2}) - G_r(e^{j\omega_1},e^{j\omega_2})\|_2 \leq 2(\sigma_{r_1+1} + \cdots + \sigma_{n_1}) + 2(\sigma_{r_2+1} + \cdots + \sigma_{n_2})$.

Proof. See [8]. □

Note that Theorem 4.3.2 shows that if one finds a solution for (4.76) or (4.77) that is singular, then one can reduce the order of the model with no error. This fact introduces the following remarkable result about minimality of 2-D models.

Theorem 4.3.3. *Given a model* $\Sigma = (A,B,C,D)$*, with transfer function* $G(z_1,z_2)$ *such that* $(Z-A)$ *is invertible in* $\bar{U}$*, there exists a lower order model* $\Sigma_r = (\hat{A},\hat{B},\hat{C},D)$*, with transfer function* $G_r(z_1,z_2)$*, such that* $G_r(z_1,z_2) = G(z_1,z_2)$ *if and only if there exist singular* $P \geq 0$ *or* $Q \geq 0$*, both in* $\mathscr{T}$*, satisfying (4.76) and (4.77). Furthermore,* $\max(dim(A) - dim(\hat{A}))$ *equals the number of zero-valued eigenvalues of the product* PQ*.*

Proof. See [27]. □

For a matter of convenience, the Beck-Doyle model order reduction algorithm is now presented in detail and step by step.

Algorithm 6. *Beck-Doyle algorithm*
Input data: a 2-D model $\Sigma = (A_{11},A_{12},A_{21},A_{22},B_1,B_2,C_1,C_2,D)$

1. *Find solutions* $P \geq 0$ *and* $Q \geq 0$*, both in* $\mathscr{T}$*, for the Lyapunov inequalities*

$$P \geq APA^T + BB^T, \tag{4.80}$$

$$Q \geq A^T QA + C^T C. \tag{4.81}$$

   Say $P = diag(P_2,P_1)$ *and* $Q = diag(Q_2,Q_1)$.

2. *Find invertible matrices* T_1 *and* T_2 *such that*

$$T_i^{-1} P_i (T_i^{-1})^T = T_i^T Q_i T_i = S_i = diag(\sigma_{1i},...,\sigma_{n_i i}),\ i = 1,2. \tag{4.82}$$

   Algorithm 1 (section 2.1.2) may be used for this purpose.

3. *Build the similarity transformation* $T := diag(T_2, T_1)$ *and apply it over* Σ*, obtaining a new realization*

$$\Sigma_b = (A_b, B_b, C_b, D) = (T^{-1}AT, T^{-1}B, CT, D). \tag{4.83}$$

This realization is a balanced one.

4. *Choose the dimensions of the reduced order horizontal and vertical states, say* r_2 *and* r_1*, respectively. Build matrices* $\hat{A}_{11}$ *and* $\hat{A}_{22}$ *by extracting the upper* $r_2 \times r_2$ *and* $r_1 \times r_1$ *blocks of matrices* A_{11_b} *and* A_{22_b} *and matrices* $\hat{B}_1$*,* $\hat{B}_2$*,* $\hat{C}_1$ *and* $\hat{C}_2$ *by truncating* B_{1_b}*,* B_{2_b}*,* C_{1_b} *and* C_{2_b} *accordingly.*
The model $\Sigma_r = (\hat{A}, \hat{B}, \hat{C}, D)$ *has reduced order* $r_1 + r_2$ *and satisfies the properties listed in Theorem 4.3.2.*

Remark The step 1. of the previous algorithm is actually far from being trivial. If the admissible solutions of (4.80) and (4.81) were unstructured, they could be easily found by adding an arbitrary positive matrix to the right member of the inequalities and transforming them into equalities. However, the constraint that both solutions belong to $\mathscr{T}$ requires more advanced methods that will not be discussed here. Fortunately, some numerical software applications (such as MATLAB) provide powerful LMI toolboxes to solve this problem.

When comparing the previous algorithm with the one for 1-D systems (recall section 2.2.1), one may wonder why equalities instead of inequalities are used in the latter. In fact, inequalities could be used in the one-dimensional case also. However, when equalities are used, the singular values of P and Q tend to be smaller, thus the error bound of the reduced order system is also smaller.

4.3.2 An alternative approach

Recall the 2-D state-space equations considered in this thesis:

$$\begin{bmatrix} x_h(i+1,j) \\ x_v(i,j+1) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ x_v(i,j) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(i,j) \tag{4.84}$$

$$y(i,j) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ x_v(i,j) \end{bmatrix} + Du(i,j). \tag{4.85}$$

One can reorganize these equations in two particularly interesting forms:

$$x_h(i+1,j) = A_{11}x_h(i,j) + \begin{bmatrix} A_{12} & B_1 \end{bmatrix} \begin{bmatrix} x_v(i,j) \\ u(i,j) \end{bmatrix} \tag{4.86}$$

$$\begin{bmatrix} x_v(i,j+1) \\ y(i,j) \end{bmatrix} = \begin{bmatrix} A_{21} \\ C_1 \end{bmatrix} x_h(i,j) + \begin{bmatrix} A_{22} & B_2 \\ C_2 & D \end{bmatrix} \begin{bmatrix} x_v(i,j) \\ u(i,j) \end{bmatrix} \tag{4.87}$$

and

$$x_v(i,j+1) = A_{22}x_v(i,j) + \begin{bmatrix} A_{21} & B_2 \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ u(i,j) \end{bmatrix} \tag{4.88}$$

$$\begin{bmatrix} x_h(i+1,j) \\ y(i,j) \end{bmatrix} = \begin{bmatrix} A_{12} \\ C_2 \end{bmatrix} x_v(i,j) + \begin{bmatrix} A_{11} & B_1 \\ C_1 & D \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ u(i,j) \end{bmatrix}. \tag{4.89}$$

Clearly, equations (4.86)-(4.87) represent a 1-D system with input $\begin{bmatrix} x_v(i,j)^T & u(i,j)^T \end{bmatrix}^T$ and output $\begin{bmatrix} x_v(i,j+1)^T & y(i,j)^T \end{bmatrix}^T$. Analogously, equations (4.88)-(4.89) represent a 1-D system with input $\begin{bmatrix} x_h(i,j)^T & u(i,j)^T \end{bmatrix}^T$ and output $\begin{bmatrix} x_h(i+1,j)^T & y(i,j)^T \end{bmatrix}^T$. It is assumed that systems (4.88)-(4.89) and (4.86)-(4.87) are controllable and observable (in the 1-D sense). If this is not the case, one can discard the uncontrollable or unobservable subspaces obtaining exact lower order realizations for the original systems. Hence, one can apply the conventional 1-D model order reduction algorithm, described in section 2.2.1, considering one of these systems. This will reduce the dimension of either the horizontal state (if form (4.86)-(4.87) is chosen) or the vertical state (if form (4.88)-(4.89) is chosen). The dimension of the other state vector can of course be reduced by taking the reduced order system obtained in the previous step, rewriting it in the other form and applying the algorithm again.
Note, however, that the applicability of this algorithm requires that A_{11} and A_{22} are stable matrices (in the 1-D sense). The next theorem gives conditions for this requirement to hold.

Theorem 4.3.4. *Consider a 2-D model $\Sigma = (A,B,C,D)$ where $(Z-A)$ is nonsingular in $\bar{U}$. Then, $\rho(A_{11}) < 1$ and $\rho(A_{22}) < 1$.*

Proof. Assume that $(Z-A)$ is nonsingular in $\bar{U}$. Then, by Theorem 3.5.5, Σ is internally stable. Suppose that the initial conditions are

$$x_h(0,0) = x_0,\ \ x_h(0,j) = 0_{n_2},\ \forall j \geq 1,\ \text{and}\ x_v(i,0) = 0_{n_1},\ \forall i \geq 0, \tag{4.90}$$

and zero inputs are applied. It is straightforward to check that

$$x_h(k,0) = {A_{11}}^k x_0,\ k = 1,2,3... \tag{4.91}$$

Since the 2-D model is internally stable, the infinite sequence $x_h(0,0), x_h(1,0), x_h(2,0), ...$ converges to the zero state for any arbitrary x_0. Therefore, A_{11} must be a stable matrix, i.e., $\rho(A_{11}) < 1$. To prove that $\rho(A_{22}) < 1$, one can repeat this proof interchanging the roles of x_h and x_v. □

Like in any other model reduction algorithm, having an upper bound for the error of the approximation is highly desirable. For this case, which is based in the 1-D balanced truncation algorithm, it is straightforward to get one. Recall that, according to Theorem 2.2.1, the reduced order system

transfer function, $G_r(z)$, satisfies:

$$\|G(z) - G_r(z)\|_{\mathscr{H}_\infty} \leq 2(\sigma_{r+1} + \cdots + \sigma_n), \tag{4.92}$$

where $G(z)$ is the original transfer function and the σ_i are the discarded singular values of that system. Consequently, if one chooses form (4.88)-(4.89) and defines

$$H^v(z_1) := \begin{bmatrix} A(z_1) & B(z_1) \\ C(z_1) & D(z_1) \end{bmatrix} \text{ and } H_r^v(z_1) := \begin{bmatrix} A_r(z_1) & B_r(z_1) \\ C_r(z_1) & D_r(z_1) \end{bmatrix}, \tag{4.93}$$

where

$$A(z_1) := A_{12}(z_1 I - A_{22})^{-1} A_{21} + A_{11}, \tag{4.94}$$
$$B(z_1) := A_{12}(z_1 I - A_{22})^{-1} B_2 + B_1, \tag{4.95}$$
$$C(z_1) := C_2(z_1 I - A_{22})^{-1} A_{21} + C_1, \tag{4.96}$$
$$D(z_1) := C_2(z_1 I - A_{22})^{-1} B_2 + D \tag{4.97}$$

and $A_r(z_1)$, $B_r(z_1)$, $C_r(z_1)$ and $D_r(z_1)$ are defined analogously using the truncated matrices, then:

$$\|H^v(z_1) - H_r^v(z_1)\|_{\mathscr{H}_\infty} \leq 2(\sigma_{r+1_1}^v + \cdots + \sigma_{n_1}^v), \tag{4.98}$$

where the $\sigma_{i_1}^v$ are the discarded singular values of the model (4.88)-(4.89) when it is seen as 1-D system. Note that the quadruple of matrices $(A(z_1), B(z_1), C(z_1), D(z_1))$ is a first level realization corresponding to the 2-D model $\Sigma = (A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, D)$, while the quadruple $(A_r(z_1), B_r(z_1), C_r(z_1), D_r(z_1))$ is a first level realization corresponding to the obtained reduced order 2-D model. For the form (4.86)-(4.87), if one defines $H^h(z_2)$ and $H_r^h(z_2)$ in the same way as $H^v(z_1)$ and $H_r^v(z_1)$, respectively, the following error bound holds:

$$\|H^h(z_2) - H_r^h(z_2)\|_{\mathscr{H}_\infty} \leq 2(\sigma_{r+1}^h + \cdots + \sigma_n^h), \tag{4.99}$$

where the σ_{i_2} are the discarded singular values of the model (4.86)-(4.87) when it is seen like a 1-D system.

4.3.3 Some academic application examples

In this section, the two methods just presented are applied to some given 2-D state-space models in order to demonstrate the effectiveness of both and also to highlight the main advantages and drawbacks of each of them.

Before presenting the first example, it is important to discuss some options that were taken in the computational implementation of both methods. The most relevant one is related with the first step of Beck-Doyle algorithm (recall Algortithm 6 in section 4.3.1), i.e., with the computation of the solutions of the LMI (4.80) and (4.81). As pointed out before, MATLAB is able to compute

solutions for those inequalities. However, there are infinitely many admissible solutions. Thus, an optimization criterion should be defined. Here and also in [28], the following heuristic is used:

$$\min_P Trace(P) : P \geq APA^T + BB^T \tag{4.100}$$

and

$$\min_Q Trace(Q) : Q \geq A^T QA + C^T C. \tag{4.101}$$

Note that P and Q are hermitian matrices and, for any hermitian matrix X, one has $Trace(X) = \sum_i \sigma_i(X)$, where the $\sigma_i(X)$ are the singular values of X.
Recall that the error of the Beck-Doyle reduced order system is bounded by the sum of the square roots of the discarded singular values of the product PQ, which are invariant under any change of basis. Given that

$$Trace(PQ) \leq Trace(P)Trace(Q) \tag{4.102}$$

(see [28]), separately minimizing the trace of P and Q leads to solutions for which the bound for the value of $\sum_i \sigma_i(PQ)$ is lowered. Consequently, the error of the reduced order is potentially smaller. Furthermore, this criterion can be easily implemented using the MATLAB LMI Control Toolbox. For a complete explanation of how to use this toolbox, see [29].
As for the implementation of the alternative approach, there are not many relevant aspects to highlight. In the examples presented here, the order reduction is always made by firstly reducing the dimension of the horizontal state and then reducing the dimension of the vertical one. This was simply an arbitrary choice, though.
In Appendix C, the MATLAB code that implements both algorithms is presented.
In order to compare the performance of each of the algorithms for the given examples, the sum of the discarded singular values may appear to be a reasonable measure. However, for the second method, it does not provide a rigorous error bound, unless the order reduction is made exclusively for one of the state vectors (either horizontal or vertical). Hence, also the following dimensionless quantity, known as the *peak signal-to-noise ratio* (PSNR), is considered:

$$\text{PSNR} := 10\log_{10}\left(\frac{\max_{i,j} h(i,j)^2}{\text{MSE}}\right), \tag{4.103}$$

where MSE denotes the *mean squared error*:

$$\text{MSE} := \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}(h(i,j) - h_r(i,j))^2. \tag{4.104}$$

Here, $h(i,j)$ and $h_r(i,j)$ denote the impulse responses of the original and of the reduced order single-output systems, respectively. If the system has multiple outputs, this measure can still be used for each output separately. It is common to express the value of the PSNR in the dimensionless unit dB. Since the impulse responses considered here have typically infinite length, for the evaluation of the PSNR it is necessary to truncate them. Here, $M = N = 50$ is used.

In the sequel, three numerical examples are presented. Note that all the following given models are internally stable.

Example 1. *Consider the following 2-D model* $\Sigma = (A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, D)$*, also presented as an example in [16]*

$$A_{11} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -0.28214 & 0.5512 & -0.8756 & 1.3618 \end{bmatrix}, \quad A_{12} = 0_{4\times 4},$$

$$A_{21} = \begin{bmatrix} -0.020056 & 0.11489 & -0.16782 & 0.2271 \\ 0.01159 & -0.054709 & 0.12089 & -0.15035 \\ 0.026037 & 0.03553 & -0.0348 & 0.047409 \\ 0.032325 & -0.055488 & 0.045669 & 0.07734 \end{bmatrix},$$

$$A_{22} = \begin{bmatrix} 0 & 0 & 0 & -0.090322 \\ 1 & 0 & 0 & 0.19922 \\ 0 & 1 & 0 & -0.39063 \\ 0 & 0 & 1 & 0.75 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \tag{4.105}$$

$$B_2 = \begin{bmatrix} 0.3681 \\ -0.3151 \\ 0.2273 \\ -0.0730 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 0.2610 & -0.1276 & -0.0339 & 0.4243 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}, \quad D = 0.6520.$$

Note that, since A_{12} *is a zero matrix, this is a separable denominator system, i.e., the denominator of the corresponding transfer function can be factored as the product of a polynomial in* z_1 *by a polynomial in* z_2*. This can easily be checked by applying (3.7).*
The impulse response and the magnitude of the frequency response of this system are represented in Figure 4.1.
In the following, a reduced order model in which the vertical state has dimension $n_1 = 3$ *and the horizontal state has dimension* $n_2 = 3$ *will be obtained using both methods.*

Beck-Doyle algorithm
Computing P and Q according to (4.100) and (4.101) and diagonalizing these matrices yields:

$$P_{bal} = Q_{bal} = S = diag(4.2713, 1.7611, 0.8225, 0.7254,\ 3.7842, 2.0321, 1.8910, 1.8092) \tag{4.106}$$

Hence, the discarded singular values will be $\sigma_{4_1} = 1.8092$ *and* $\sigma_{4_2} = 0.7254$*. Consequently, the transfer function of the reduced order system,* $G_r(z_1, z_2)$*, satisfies*

$$\|G(z_1, z_2) - G_r(z_1, z_2)\|_{\mathscr{H}_\infty} \leq 2(\sigma_{4_1} + \sigma_{4_2}) = 5.0692, \tag{4.107}$$

where $G(z_1,z_2)$ is the original system transfer function.
The reduced order model $\Sigma_r = (\hat{A},\hat{B},\hat{C},D)$ is obtained by simply truncating the balanced realization (which was left out for a matter of brevity):

$$\hat{A}_{11} = \begin{bmatrix} 0.8013 & 0.2718 & -0.0885 \\ -0.2712 & 0.7330 & 0.1583 \\ -0.1193 & -0.3117 & 0.1321 \end{bmatrix}, \ \hat{A}_{12} = 0_{3\times 3}$$

$$\hat{A}_{21} = \begin{bmatrix} -0.1607 & 0.0666 & -0.0100 \\ 0.1016 & 0.0069 & 0.0504 \\ -0.0628 & -0.0753 & -0.0206 \end{bmatrix},$$

$$\hat{A}_{22} = \begin{bmatrix} 0.7082 & -0.2124 & -0.2849 \\ 0.2707 & -0.5090 & 0.3116 \\ -0.1288 & -0.6228 & 0.0268 \end{bmatrix}, \ \hat{B}_1 = \begin{bmatrix} 0.9645 \\ 0.6208 \\ 0.4781 \end{bmatrix}, \tag{4.108}$$

$$\hat{B}_2 = \begin{bmatrix} -0.0983 \\ 0.2799 \\ -0.1012 \end{bmatrix}, \ \hat{C}_1 = \begin{bmatrix} 0.5360 & -0.2393 & 0.2243 \end{bmatrix},$$

$$\hat{C}_2 = \begin{bmatrix} -0.9548 & -0.5252 & -0.6024 \end{bmatrix}.$$

The sum of the discarded singular values is 2.5345 and the PSNR of the impulse of this reduced order model is 36.665dB. The computation time was $\approx 0.47s$.
Figure 4.2 represents the impulse response and the frequency response of this reduced order model. In order to compare the reduced order model with the original one, the plots of the errors of the impulse response and of the frequency response are presented in Figure 4.3.

Alternative approach
Start by finding a balanced realization (in the 1-D sense) for the system in the form (4.86)-(4.87). The corresponding gramians are:

$$P^h_{bal} = Q^h_{bal} = S^h = diag(1.7889, 0.6513, 0.3442, 0.3221). \tag{4.109}$$

Reduce the order of the horizontal state by truncating the balanced realization. This corresponds to discarding the singular value $\sigma^h_4 = 0.3221$.
After writing the reduced order model in the form (4.88)-(4.89) and balancing (in the 1-D sense) one obtains a realization with gramians:

$$P^v_{bal} = Q^v_{bal} = S^v = diag(0.6566, 0.4947, 0.4344, 0.3993). \tag{4.110}$$

The singular value $\sigma^v_4 = 0.3993$ is discarded by truncating the matrices of the realization. The

final reduced order model $\Sigma_r = (\hat{A}, \hat{B}, \hat{C}, D)$ *has:*

$$\hat{A}_{11} = \begin{bmatrix} 0.8026 & 0.2452 & -0.0981 \\ -0.2640 & 0.7499 & 0.1930 \\ -0.1604 & -0.3322 & 0.1053 \end{bmatrix}, \ \hat{A}_{12} = 0_{3\times 3},$$

$$\hat{A}_{21} = \begin{bmatrix} 0.3619 & -0.1697 & -0.0311 \\ -0.1584 & 0.0619 & -0.0804 \\ -0.1964 & -0.1712 & -0.0699 \end{bmatrix},$$

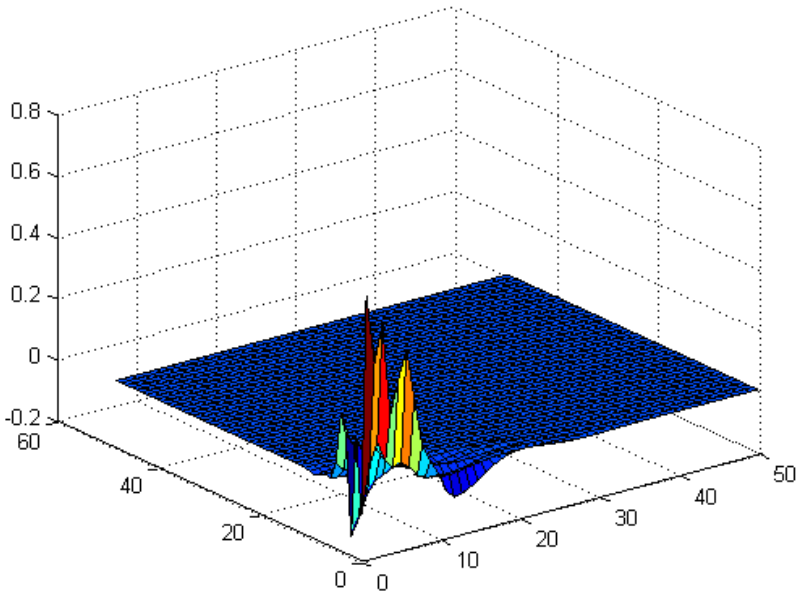
$$\hat{A}_{22} = \begin{bmatrix} 0.3619 & -0.1697 & -0.0311 \\ -0.1584 & 0.0619 & -0.0804 \\ -0.1964 & -0.1712 & -0.0699 \end{bmatrix}, \ \hat{B}_1 = \begin{bmatrix} -0.7667 \\ -0.3603 \\ -0.3409 \end{bmatrix}, \tag{4.111}$$

$$\hat{B}_2 = \begin{bmatrix} -0.1384 \\ 0.3738 \\ 0.2518 \end{bmatrix}, \ \hat{C}_1 = \begin{bmatrix} -0.6866 & 0.3519 & -0.1938 \end{bmatrix},$$

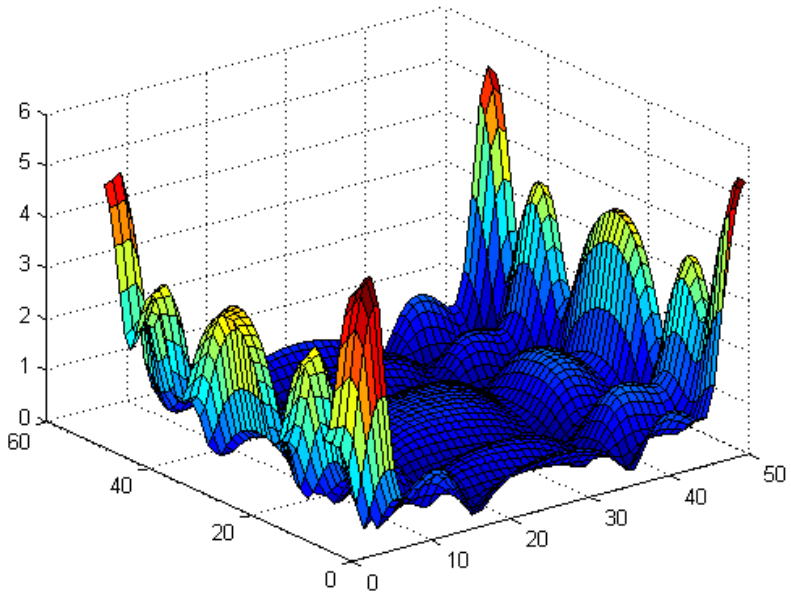
$$\hat{C}_2 = \begin{bmatrix} -0.5344 & -0.4018 & 0.2420 \end{bmatrix}.$$

Moreover, this 2-D system is BIBO stable.
The sum of the discarded singular values is 0.7213 and the impulse response of this model yields a $PSNR = 36.950dB$. *The computation time was* $\approx 0.048s$. *The plots for this reduced order model and for the corresponding error are displayed in figures 4.4 and 4.5.*

As the results show, the alternative approach performs slightly better for this system. Since that method is also less complex than the Beck-Doyle algorithm, in a computational point of view, it is clearly preferable in this case. In fact, the computation time for the alternative approach was almost ten times smaller than the computation time for the Beck-Doyle algorithm.

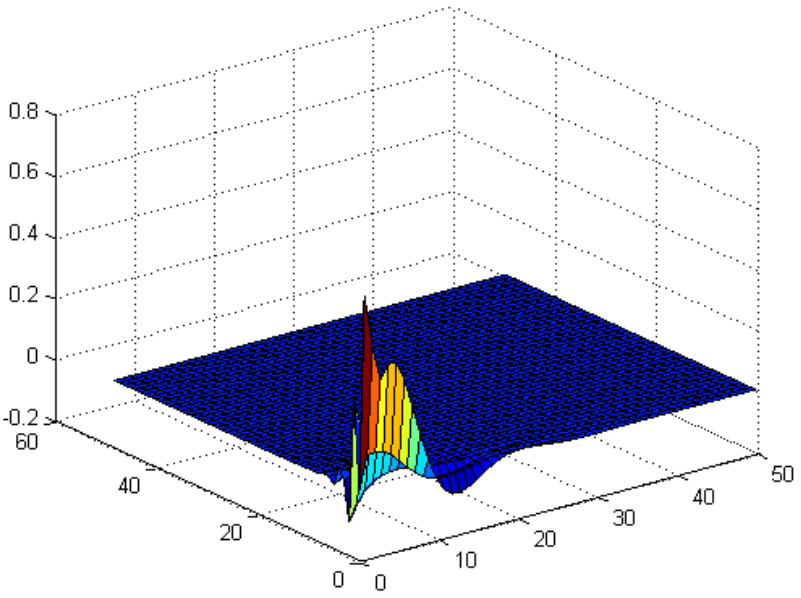


(a) Impulse response

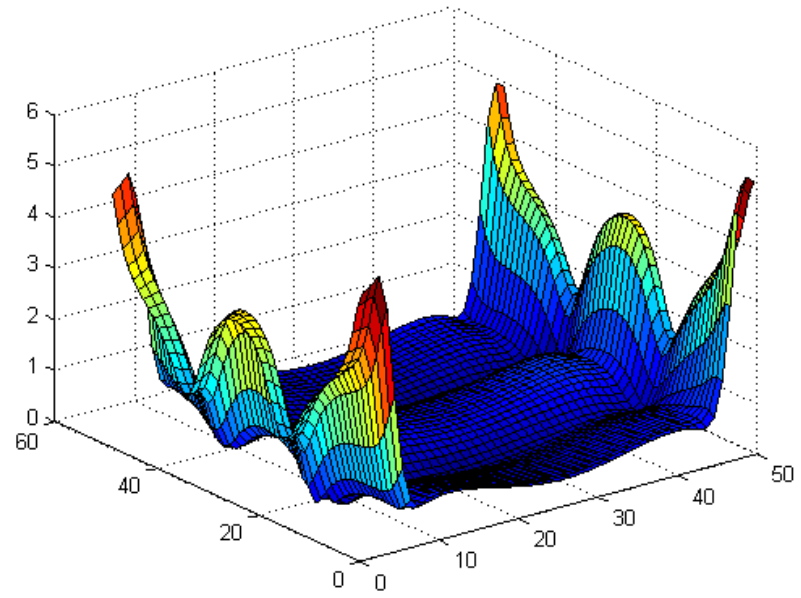


(b) Frequency response (magnitude)

Figure 4.1: Original system – Example 1

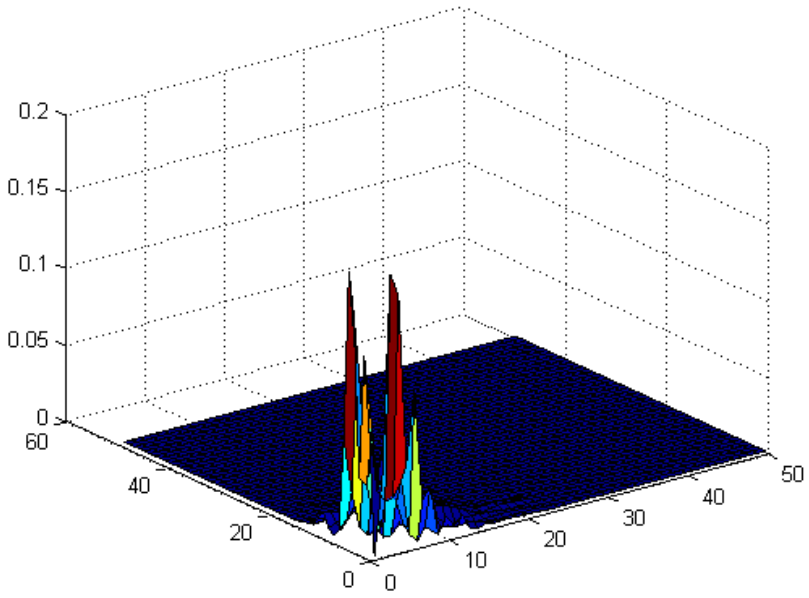


(a) Impulse response

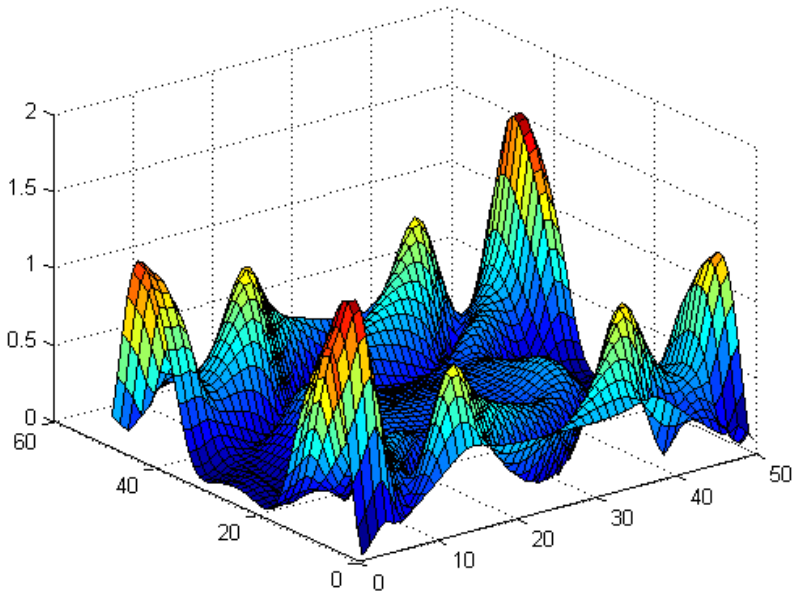


(b) Frequency response (magnitude)

Figure 4.2: Reduced order system (Beck-Doyle algorithm) – Example 1

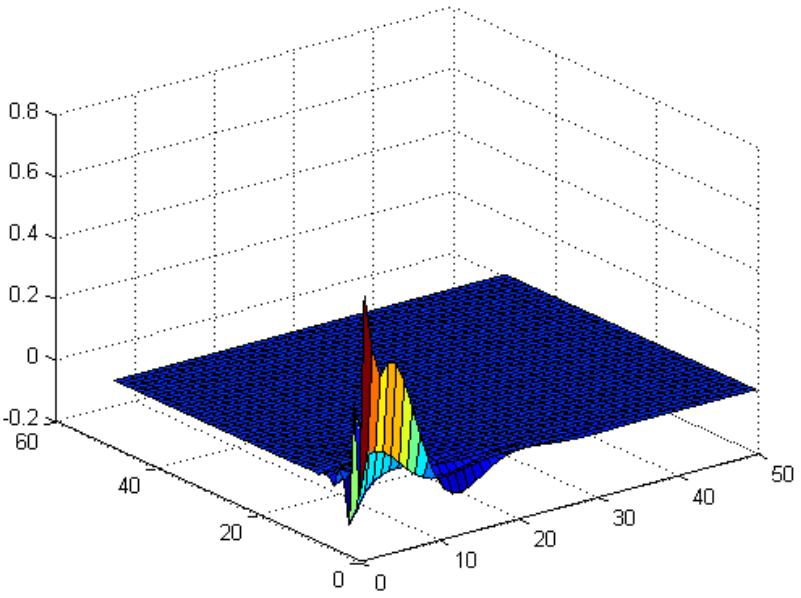


(a) Error of the impulse response

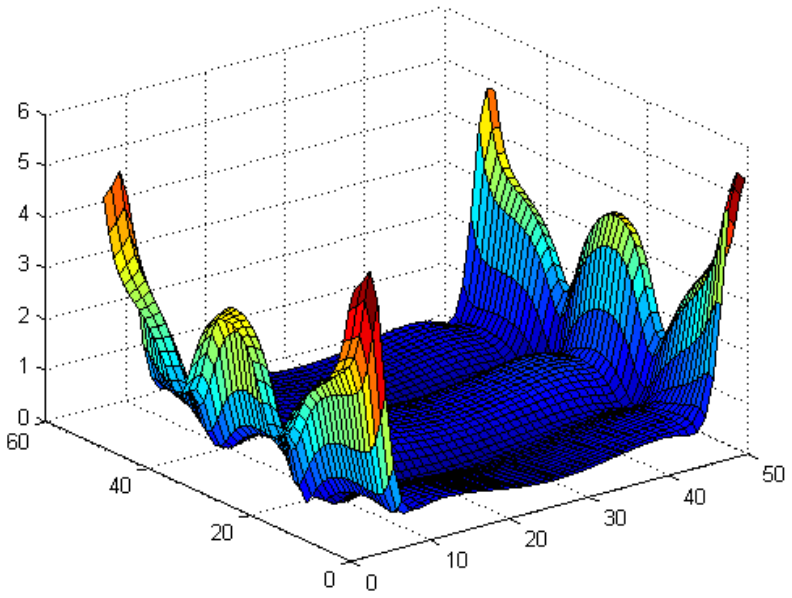


(b) Error of the frequency response (magnitude)

Figure 4.3: Error (Beck-Doyle algorithm) – Example 1

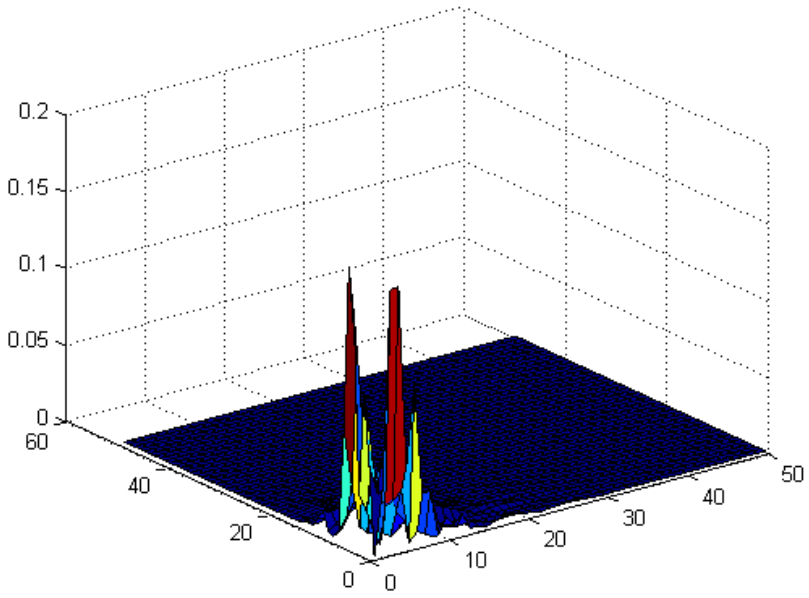


(a) Impulse response

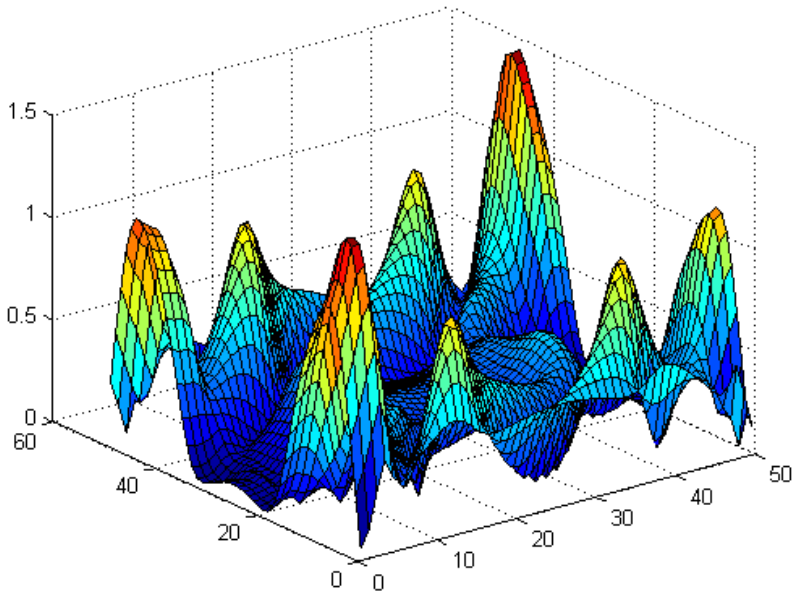


(b) Frequency response (magnitude)

Figure 4.4: Reduced order system (Alternative approach) – Example 1



(a) Error of the impulse response



(b) Error of the frequency response (magnitude) – Example 1

Figure 4.5: Error (Alternative approach) – Example 1

Example 2. *The following one-input / two-outputs stable model was presented in [22]:*

$$\begin{aligned} A_{11} &= \begin{bmatrix} 0.0300 & 0.2100 \\ 0.0200 & 0.1100 \end{bmatrix}, \; A_{12} = \begin{bmatrix} 0.0020 \\ 0.0250 \end{bmatrix}, \\ A_{21} &= \begin{bmatrix} 0 & 0.2000 \end{bmatrix}, \; A_{22} = 0.7000, \\ B_1 &= \begin{bmatrix} 1.0000 \\ 0 \end{bmatrix} \times 10^{-3}, \; B_2 = 0, \\ C_1 &= \begin{bmatrix} 0.2000 & 0.2100 \\ 1.8000 & 1.0000 \end{bmatrix}, \; C_2 = \begin{bmatrix} 0.2100 \\ 0.4000 \end{bmatrix}, \\ D &= \begin{bmatrix} 0.1000 \\ 0.8000 \end{bmatrix}. \end{aligned} \tag{4.112}$$

The impulse and frequency responses of each of the outputs are represented in figures 4.6 and 4.7. The dimension of the model will be reduced using both methods.

Beck-Doyle algorithm

The procedure is completely identical to the one followed in the previous examples. The obtained diagonal gramians are:

$$P_{bal} = Q_{bal} = S = diag(0.0020, 0.0000, \; 0.0000). \tag{4.113}$$

Since $\sigma_{1_1} = 0$ and $\sigma_{2_2} = 0$, one horizontal state and one vertical state can be discarded to obtain a zero-error reduced order model. This means that the given 2-D model is not minimal. The reduced order model is $\Sigma_r = (\hat{A}, \hat{B}, \hat{C}, D)$, where:

$$\hat{A}_{11} = 0.0432, \; \hat{B}_1 = 0.0452, \; \hat{C}_1 = \begin{bmatrix} 0.0044 \\ 0.0398 \end{bmatrix}, \tag{4.114}$$

and $\hat{A}_{12}$, $\hat{A}_{21}$, $\hat{A}_{22}$, $\hat{B}_2$ and $\hat{C}_2$ are empty matrices, since the vertical state has now null dimension. This is an exact realization of the original transfer function, $G(z_1, z_2)$. Indeed, $G(z_1, z_2) = C(Z - A)^{-1}B + D = \hat{C}(\hat{Z} - \hat{A})^{-1}\hat{B} + D$. The computation time required to perform this model order reduction was $\approx 0.52s$.

Alternative approach

The same order reduction will now be performed using the alternative approach. Once again, the steps are exactly the same that were followed in the previous example.
The horizontal and vertical diagonal gramians are:

$$P^h_{bal} = Q^h_{bal} = S^h = diag(0.0344, 0.0035), \tag{4.115}$$

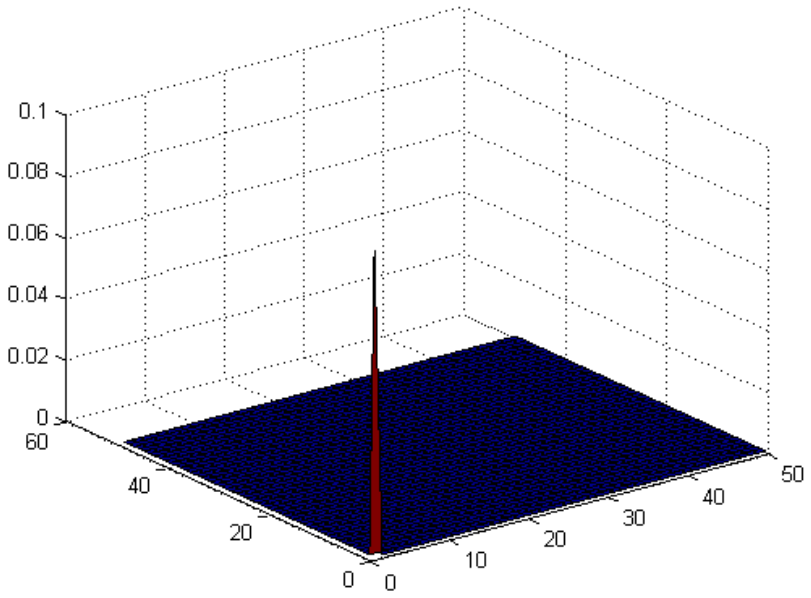
$$P^v_{bal} = Q^v_{bal} = S^v = 0.0249. \tag{4.116}$$

The states corresponding to the singular values $\sigma_2^h = 0.0035$ *and* $\sigma_1^v = 0.0249$ *are discarded, resulting in the reduced order model* $\Sigma_r = (\hat{A}, \hat{B}, \hat{C}, D)$*, where:*

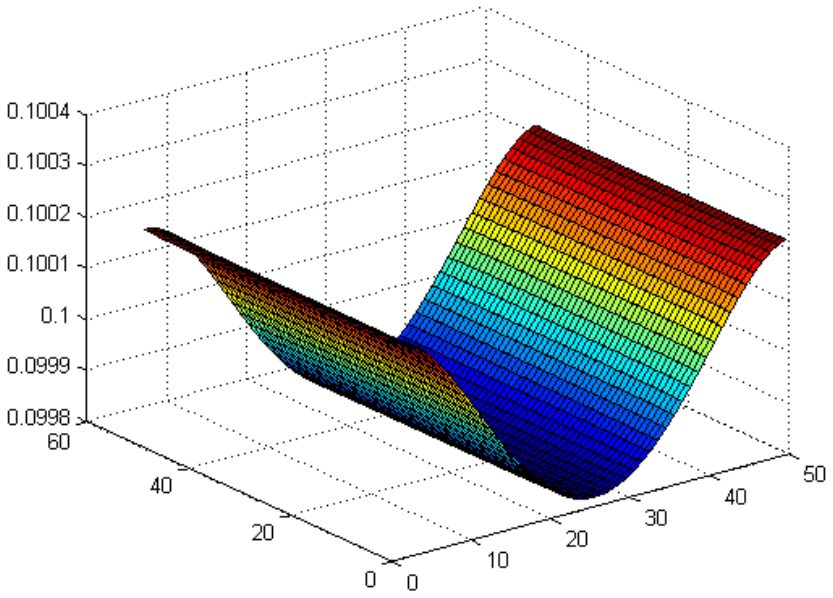
$$\hat{A}_{11} = 0.3485, \;\; \hat{B}_1 = 0.0091, \;\; \hat{C}_1 = \begin{bmatrix} 0.0316 \\ 0.1682 \end{bmatrix}, \tag{4.117}$$

and $\hat{A}_{12}$*,* $\hat{A}_{21}$*,* $\hat{A}_{22}$*,* $\hat{B}_2$ *and* $\hat{C}_2$ *are empty matrices. Note that, unlike the reduced order model obtained via Beck-Doyle algorithm, this one is not a lower order realization of the original system transfer function, since the discarded singular values are not zero. The PSNR of the respective impulse response from the input to output 1 is* $91.595dB$*, while the PSNR of the impulse response from the input to output 2 is* $97.020dB$*. The amount of time required to perform the model order reduction was* $\approx 0.066s$*.*
Some relevant plots corresponding to this model and to the respective error are presented in figures 4.8-4.11.

Clearly, for the model given in this example, it is more advantageous to use Beck-Doyle algorithm, since it gives an exact lower order realization of the original transfer function. Note, however, that although Theorem 4.3.3 provide necessary and sufficient conditions for the minimality of a 2-D model, it is not ensured that singular P or Q are found (if they exist) when they are computed according to (4.100) and (4.101). In fact, the criterion used here is minimizing the trace of the solutions, but, for the same LMI, there may exist a nonsingular solution with a smaller trace than the singular one.
Moreover, the PSNR of the impulse responses of the low order model provided by the alternative approach are rather large, what suggests that the obtained model is also a good approximation of the original one. Once again, the alternative approach proved to be faster than the Beck-Doyle algorithm.

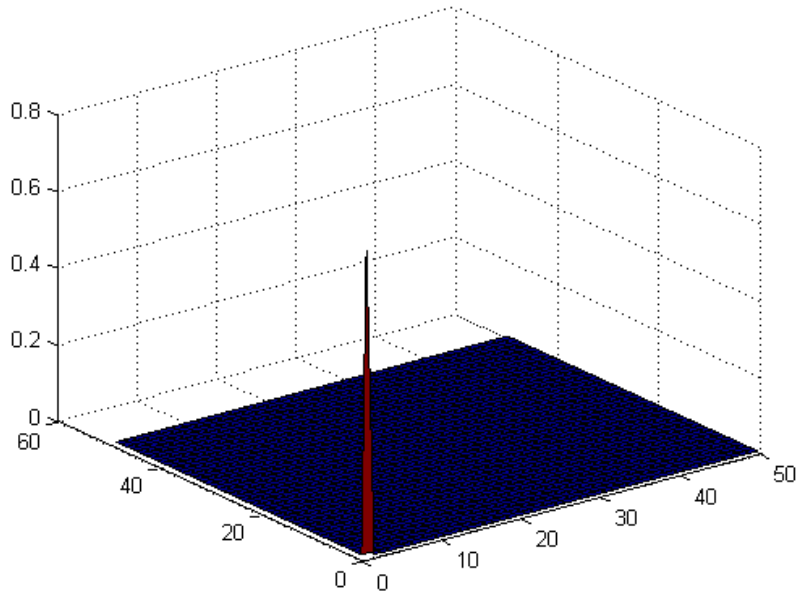


(a) Impulse response

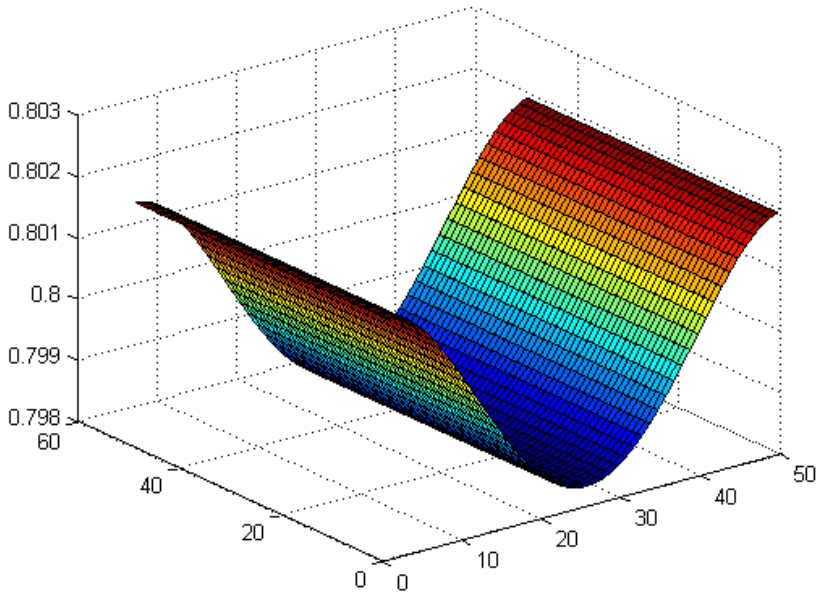


(b) Frequency response (magnitude)

Figure 4.6: Original system – Example 2, output 1

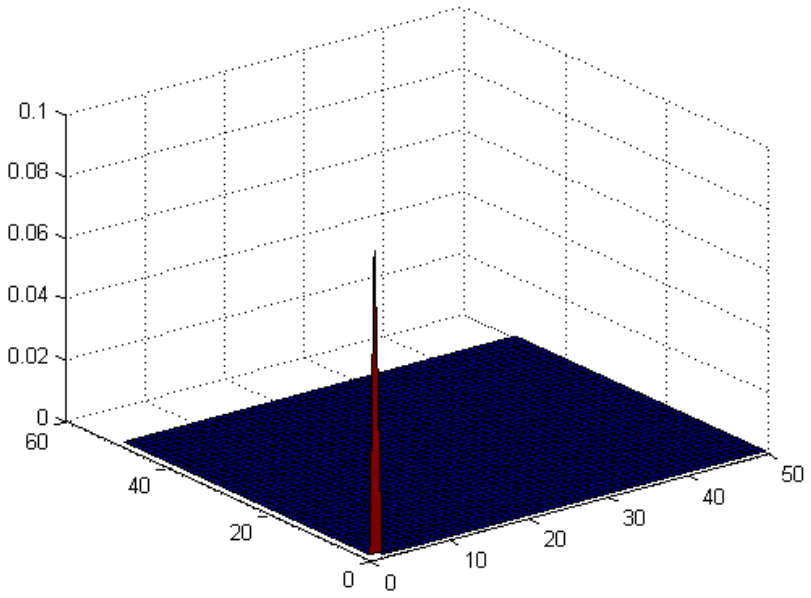


(a) Impulse response

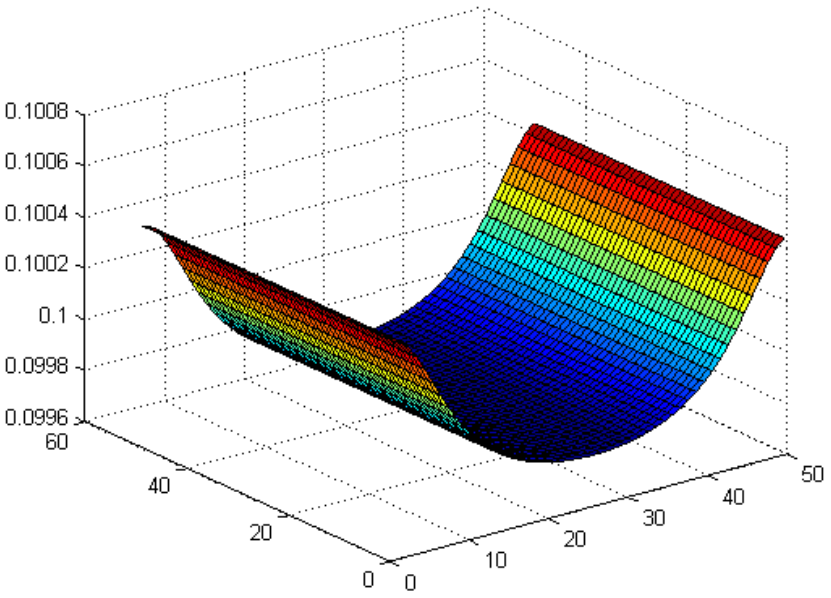


(b) Frequency response (magnitude)

Figure 4.7: Original system – Example 2, output 2

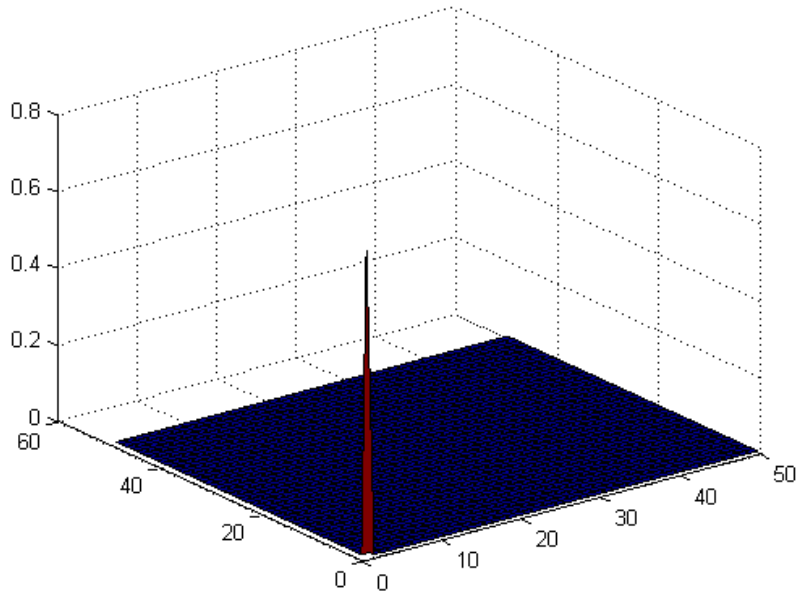


(a) Impulse response

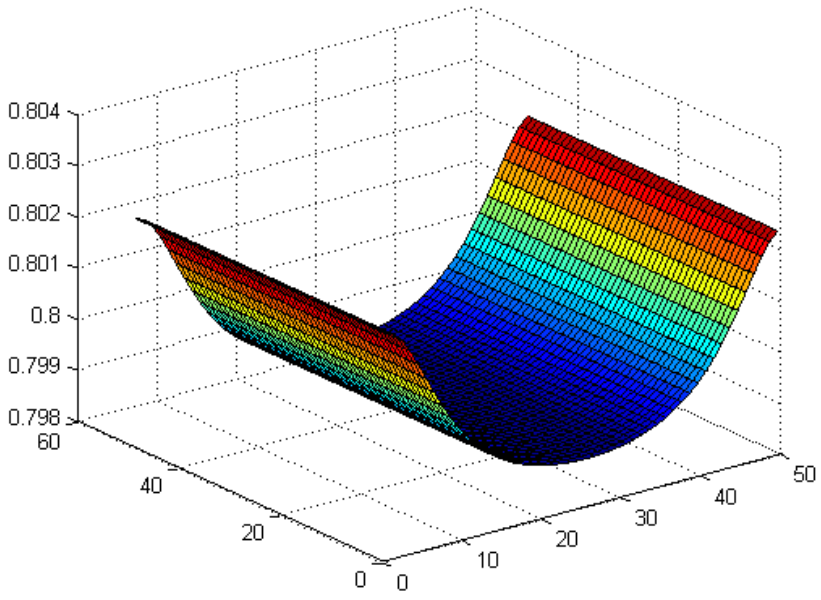


(b) Frequency response (magnitude)

Figure 4.8: Reduced order system (Alternative approach) – Example 2, output 1

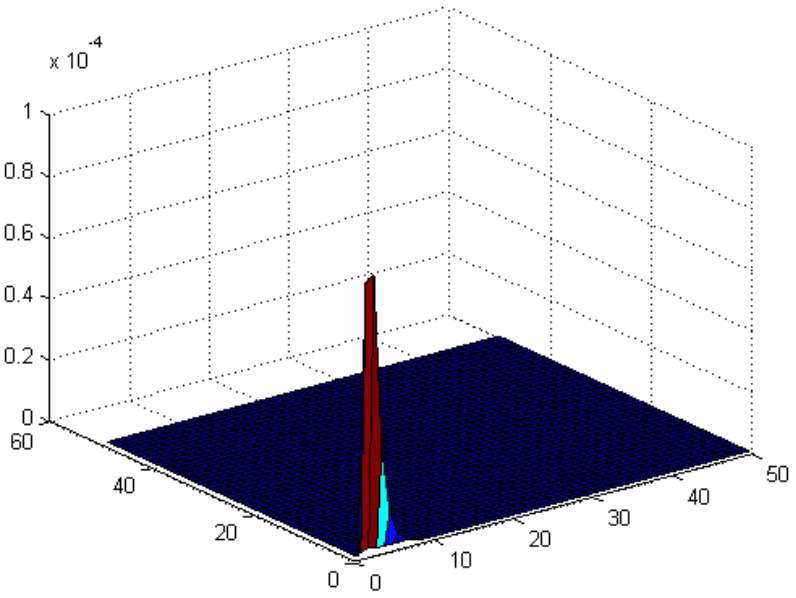


(a) Impulse response

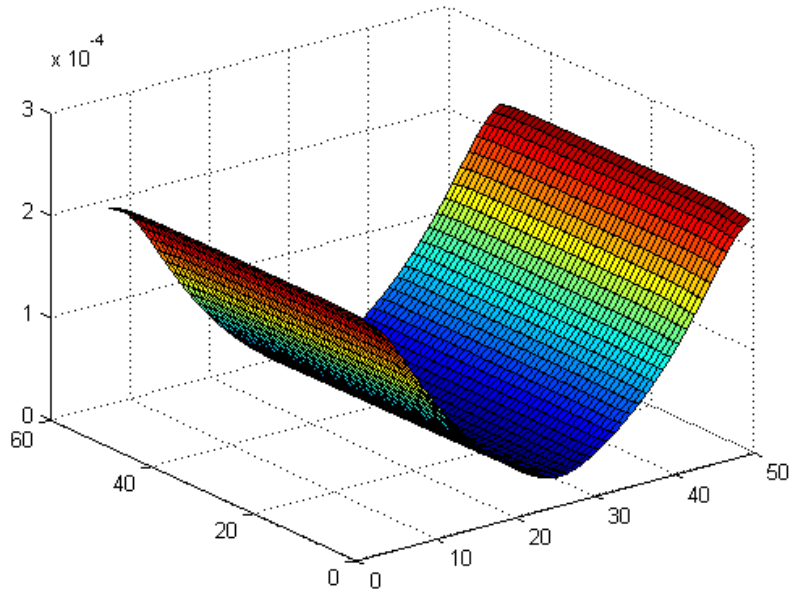


(b) Frequency response (magnitude)

Figure 4.9: Reduced order system (Alternative approach) – Example 2, output 2

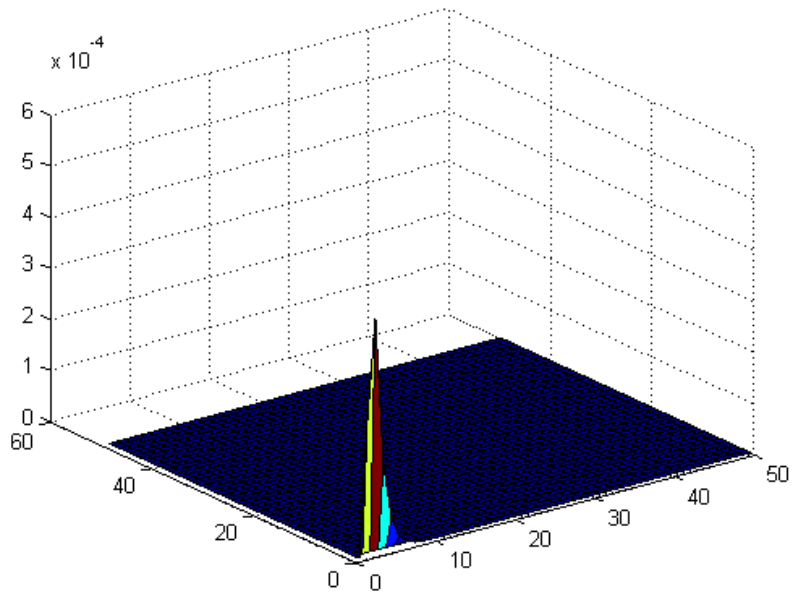


(a) Impulse response

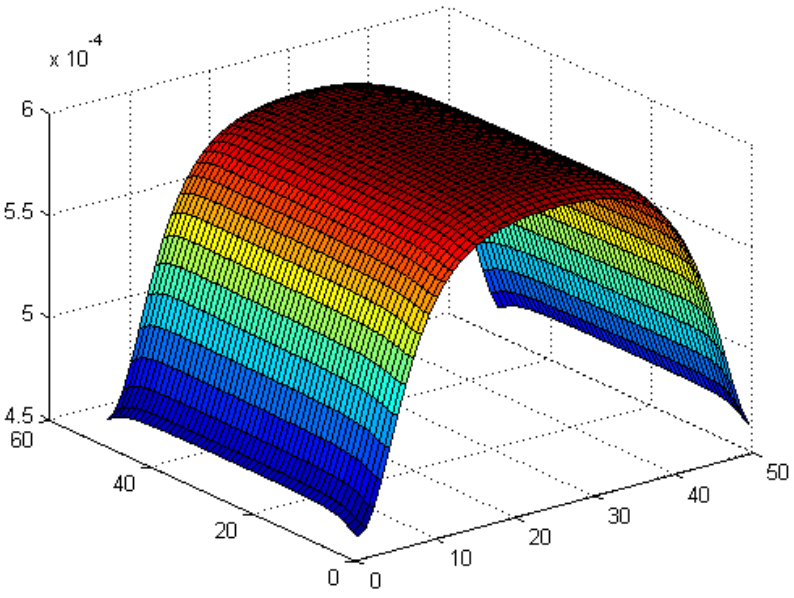


(b) Frequency response (magnitude)

Figure 4.10: Error (Alternative approach) – Example 2, output 1



(a) Impulse response



(b) Frequency response (magnitude)

Figure 4.11: Error (Alternative approach) – Example 2, output 2

Example 3. *The following model was used by K. Zhou et al. to test their model order reduction method:*

$$A_{11} = \begin{bmatrix} 0.5370 & -0.0688 & 0.9855 & 0.5039 \\ 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0.5388 & -0.0666 \\ 0 & 0 & 1.0000 & 0 \end{bmatrix},$$

$$A_{12} = \begin{bmatrix} -1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} -0.3907 & 0.2450 & -0.4836 & -0.2473 \\ 0.2512 & -0.1451 & 0.0270 & 0.0138 \\ 1.2705 & 1.1068 & 0.1981 & 0.1013 \\ 1.7964 & 0.4220 & 0.5921 & 0.3027 \\ 0 & 0 & -0.3934 & 0.2425 \\ 0 & 0 & 0.2520 & -0.1425 \\ 0 & 0 & 1.2708 & 1.1072 \\ 0 & 0 & 1.7975 & 0.4233 \end{bmatrix}, \tag{4.118}$$

$$A_{22} = \begin{bmatrix} 0.4907 & 1.0000 & 0 & 0 & 0.4907 & 0 & -0.4907 & 0 \\ -0.0274 & 0 & 0 & 0 & -0.0274 & 0 & 0.0274 & 0 \\ -0.2011 & 0 & 0 & 1.0000 & -0.2011 & 0 & 0.2011 & 0 \\ -0.6008 & 0 & 0 & 0 & -0.6008 & 0 & 0.6008 & 0 \\ 0 & 0 & 0 & 0 & 0.4912 & 1.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.0282 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.2011 & 0 & 0 & 1.0000 \\ 0 & 0 & 0 & 0 & -0.6008 & 0 & 0 & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 1.3404 \\ 0 \\ 1.3404 \\ 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} -0.6578 \\ 0.0367 \\ 0.2695 \\ 0.8054 \\ -0.6585 \\ 0.0378 \\ 0.2695 \\ 0.8054 \end{bmatrix} \times 10^{-3},$$

$$C_1 = \begin{bmatrix} 0.9837 & 0.5016 & 0.9855 & 0.5039 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \end{bmatrix},$$

$$D = 0.0013.$$

The impulse and frequency responses of this system are represented in Figure 4.12.
The dimension of the horizontal state will be reduced to 3, while the dimension of the vertical state will be reduced to 5.

Beck-Doyle algorithm

Once again, the procedure is the same. The obtained diagonal gramians are:

$$\begin{aligned} P_{bal} = Q_{bal} = S = diag(&542.5809, 407.4976, 294.1651, 249.2085, \\ &509.1803, 317.9776, 207.8968, 162.8105, 96.7848, 71.5243, 3.0771, 1.7056). \end{aligned} \tag{4.119}$$

These singular values are clearly large, thus the reduced order model is expected to be a rough approximation of the original one. More concretely, discarding one horizontal state and three vertical states yields a reduced order system whose transfer function $G_r(z_1, z_2)$ satisfies:

$$\|G(z_1, z_2) - G_r(z_1, z_2)\|_{\mathscr{H}_\infty} \leq 2(\sigma_{6_1} + \sigma_{7_1} + \sigma_{8_1} + \sigma_{4_2}) = 651.0309, \tag{4.120}$$

where $G(z_1, z_2)$ is the original system transfer function. The corresponding reduced order model is $\Sigma_r = (\hat{A}, \hat{B}, \hat{C}, D)$, where:

$$\hat{A}_{11} = \begin{bmatrix} 0.4284 & -0.0750 & -0.0022 \\ 0.1960 & 0.3663 & -0.0119 \\ 0.1192 & -0.4135 & 0.1992 \end{bmatrix},$$

$$\hat{A}_{12} = \begin{bmatrix} -0.3796 & 0.2646 & 0.2988 & 0.2475 & -0.0932 \\ -0.2843 & -0.1603 & -0.4666 & 0.2364 & -0.0745 \\ -0.0473 & -0.1145 & -0.2469 & 0.0519 & -0.0136 \end{bmatrix},$$

$$\hat{A}_{21} = \begin{bmatrix} -0.5397 & -0.0972 & -0.0748 \\ 0.1053 & -0.5664 & 0.0757 \\ 0.1713 & -0.1255 & -0.6096 \\ 0.0005 & 0.0569 & 0.0932 \\ 0.0803 & 0.1623 & -0.0536 \end{bmatrix},$$

$$\hat{A}_{22} = \begin{bmatrix} 0.3925 & -0.1703 & 0.1143 & 0.2592 & -0.1854 \\ 0.1063 & 0.3046 & -0.1947 & -0.2230 & -0.2532 \\ 0.0979 & -0.2866 & 0.0747 & 0.1342 & 0.0817 \\ 0.0718 & 0.1981 & -0.1019 & -0.0139 & -0.2164 \\ 0.1674 & 0.2276 & -0.0869 & 0.0384 & -0.3230 \end{bmatrix}, \tag{4.121}$$

$$\hat{B}_1 = \begin{bmatrix} 12.0649 \\ -10.2401 \\ -6.4287 \end{bmatrix}, \quad \hat{B}_2 = \begin{bmatrix} -0.0042 \\ 0.0084 \\ 0.0090 \\ -0.0036 \\ -0.0024 \end{bmatrix},$$

$$\hat{C}_1 = \begin{bmatrix} 1.9632 & 2.9827 & -1.5114 \end{bmatrix},$$

$$\hat{C}_2 = \begin{bmatrix} -8.0253 & -10.2356 & 3.4016 & -5.3367 & -4.9216 \end{bmatrix}.$$

The obtained PSNR is $41.0756dB$*, which may be considered a surprisingly large value, given that the error of the reduced model transfer function is also large. The computation time was* $\approx 0.55s$.

Alternative approach

The gramians of the 1-D balanced realizations are:

$$P^h_{bal} = Q^h_{bal} = S^h = diag(34.2824, 10.9502, 2.2389, 1.2802) \tag{4.122}$$

$$P^v_{bal} = Q^v_{bal} = S^v = diag(6.6260, 2.3896, 1.1610, 0.7580, 0.2196, 0.0954, 0.4036, 0.0107) \tag{4.123}$$

which are clearly lower than the singular values obtained with Beck-Doyle algorithm. The corresponding reduced order model is $\Sigma_r = (\hat{A}, \hat{B}, \hat{C}, D)$*, where:*

$$\hat{A}_{11} = \begin{bmatrix} 0.8026 & -0.2912 & -0.0017 \\ 0.3046 & 0.4392 & -0.3252 \\ -0.0139 & 0.3709 & 0.0315 \end{bmatrix},$$

$$\hat{A}_{12} = \begin{bmatrix} 0.3767 & 0.3902 & -0.2859 & -0.0001 & 0.3432 \\ 0.2275 & 0.1174 & -0.0353 & 0.6639 & -0.1337 \\ -0.2319 & -0.1674 & 0.0914 & -0.4088 & -0.0013 \end{bmatrix},$$

$$\hat{A}_{21} = \begin{bmatrix} 1.8617 & 0.2220 & -0.1314 \\ 0.4380 & -1.1033 & -0.3210 \\ 0.4468 & 0.6423 & 0.0286 \\ -0.2782 & 0.2440 & -0.7015 \\ -0.1550 & 0.0130 & 0.2199 \end{bmatrix},$$

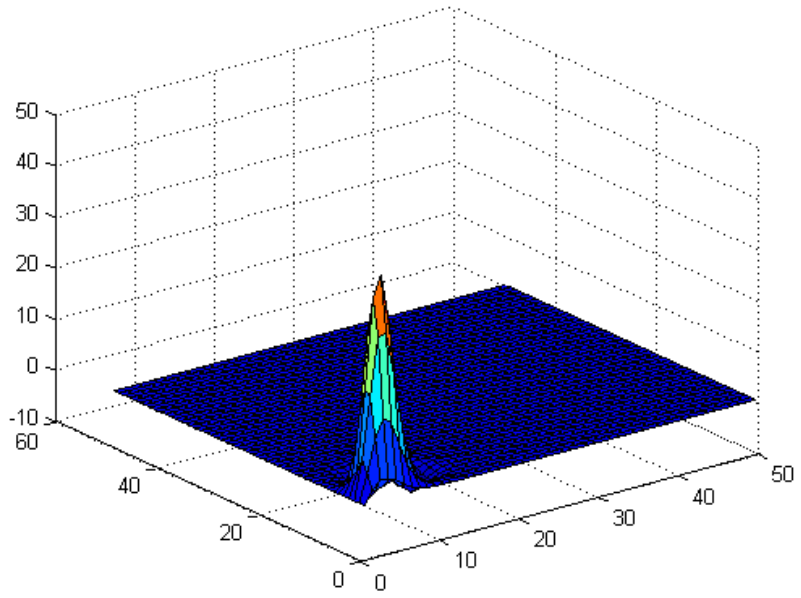
$$\hat{A}_{22} = \begin{bmatrix} 0.6385 & 0.3987 & 0.1004 & -0.0361 & -0.0152 \\ -0.2029 & 0.3197 & -0.5431 & -0.1389 & -0.1225 \\ -0.2233 & 0.0259 & -0.3708 & 0.2562 & 0.1732 \\ 0.0506 & 0.1815 & -0.1422 & -0.0437 & -0.0630 \\ 0.0439 & 0.1617 & -0.1699 & 0.2131 & -0.1311 \end{bmatrix}, \tag{4.124}$$

$$\hat{B}_1 = \begin{bmatrix} -2.6085 \\ 1.8216 \\ -0.4862 \end{bmatrix}, \quad \hat{B}_2 = \begin{bmatrix} -0.0019 \\ -0.0021 \\ 0.0007 \\ 0.0019 \\ 0.0002 \end{bmatrix},$$

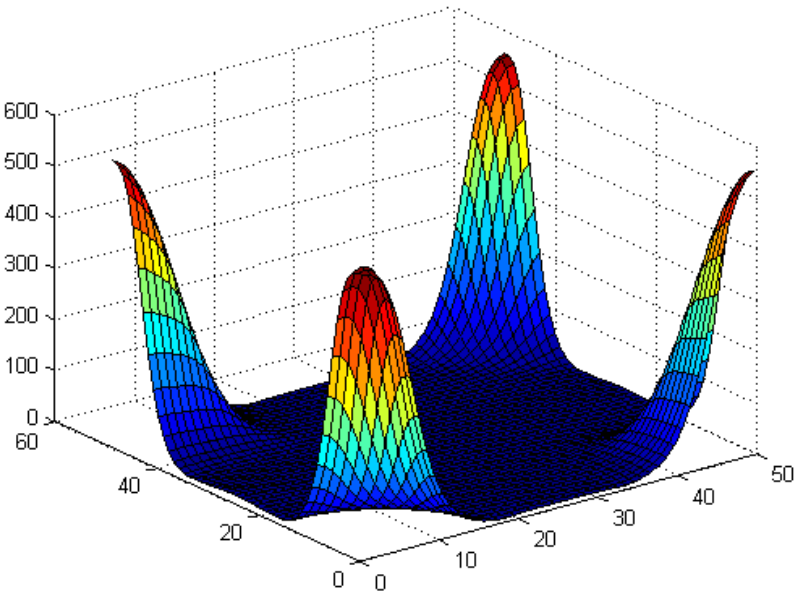
$$\hat{C}_1 = \begin{bmatrix} -1.5720 & -0.8011 & -0.1150 \end{bmatrix},$$

$$\hat{C}_2 = \begin{bmatrix} -1.8759 & 0.9308 & 0.3413 & 0.0503 & 0.0244 \end{bmatrix}.$$

Note that, for this reduced order model, one has $\rho(\hat{A}_{11}) = 0.5243 < 1$ *and* $\rho(\hat{A}_{22}) = 0.4080 < 1$, *thus these matrices are stable in the 1-D sense. This fact was ensured a priori by Theorem 2.2.1. However, it is easy to check that the matrix* $\hat{A}$ *does not accomplish the sufficient BIBO stability conditions stated in Theorem 3.5.4. If one computes the spectral radius of* $\hat{A}$, *one obtains* $\rho(\hat{A}) = 1.2084 > 1$, *thus the matrix* $(\hat{Z} - \hat{A})$ *is singular at some* $(z_1^*, z_2^*) \in \mathbb{C}^2$ *such that* $z_1 = z_2 = z$ *and* $|z| = 1.2084$. *This* (z_1^*, z_2^*) *is in* $\bar{U}$, *so* $\hat{A}$ *does not satisfy that BIBO stability test. Hence, one may suspect that this reduced order 2-D model is unstable. Plots 4.15(a) and 4.15(b) clearly show that this suspicion is true. In fact, this is the main drawback of the alternative approach: given a BIBO stable 2-D model, there is no guarantee that the reduced order model is also BIBO stable.*
The computation time of the alternative approach for the model order reduction was $\approx 0.050s$.

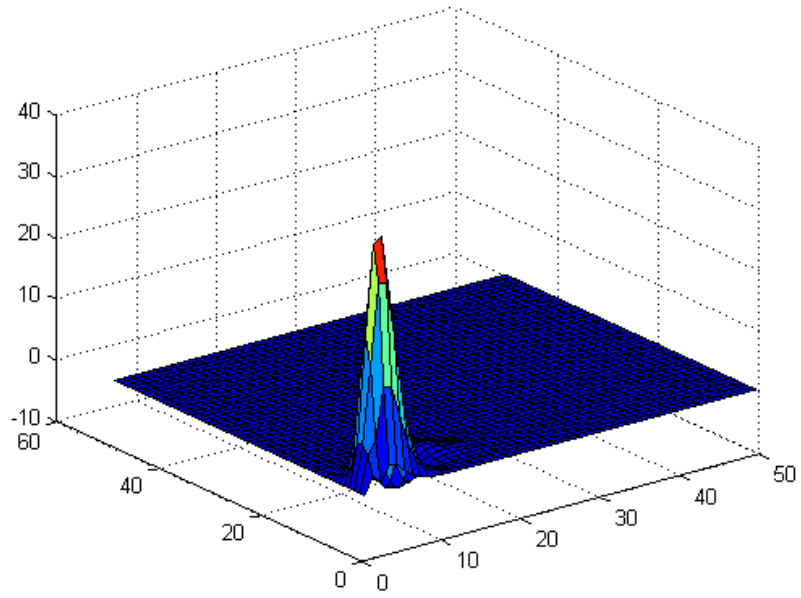


(a) Impulse response

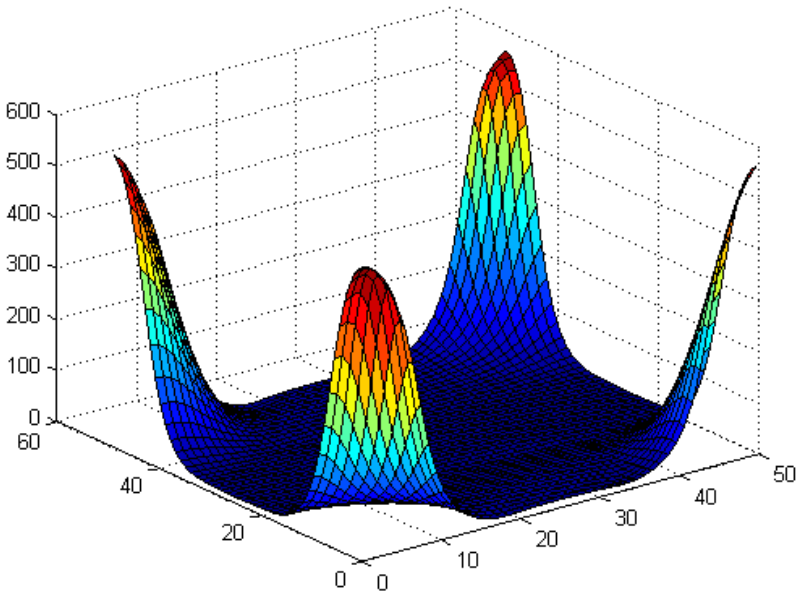


(b) Frequency response (magnitude)

Figure 4.12: Original system – Example 3

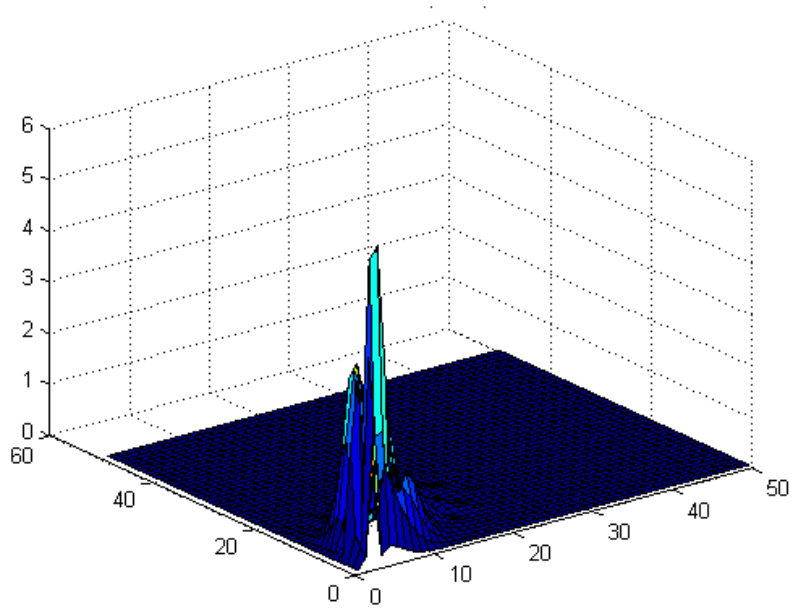


(a) Impulse response

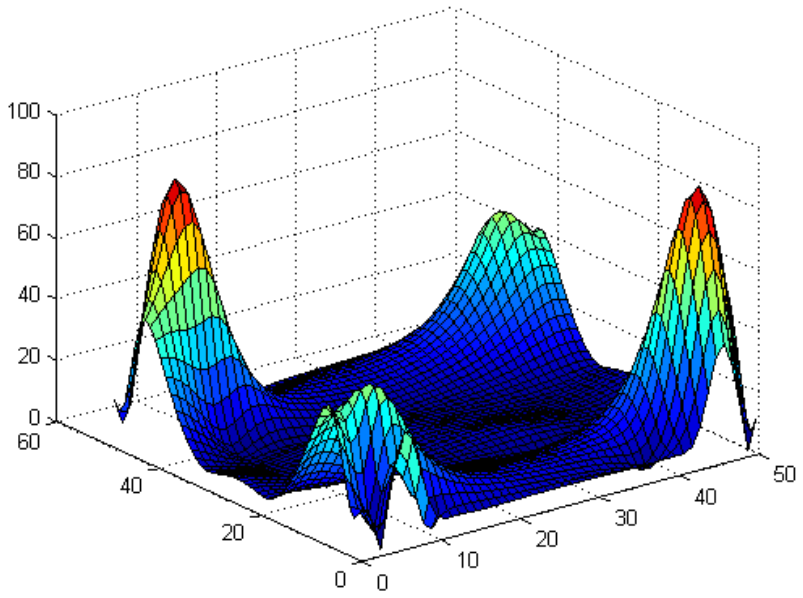


(b) Frequency response (magnitude)

Figure 4.13: Reduced order system (Beck-Doyle algorithm) – Example 3

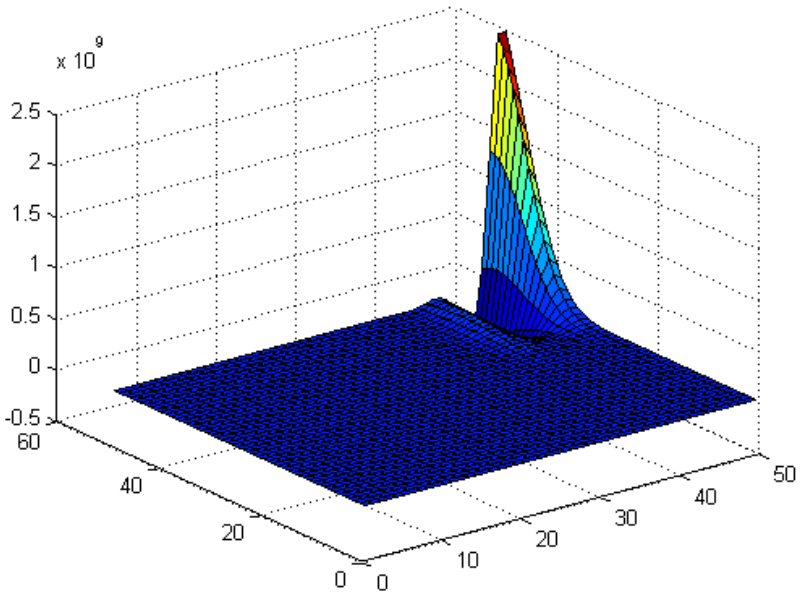


(a) Error of the impulse response

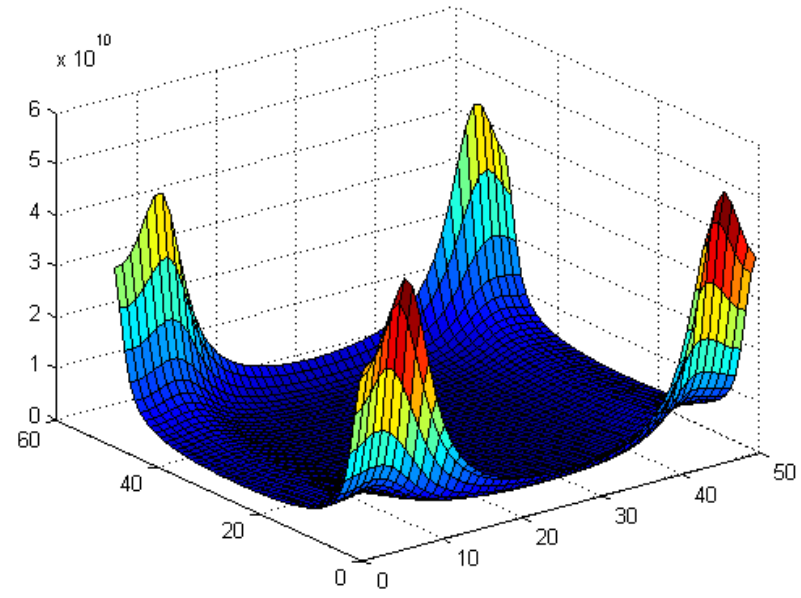


(b) Error of the frequency response (magnitude)

Figure 4.14: Error (Beck-Doyle algorithm) – Example 3

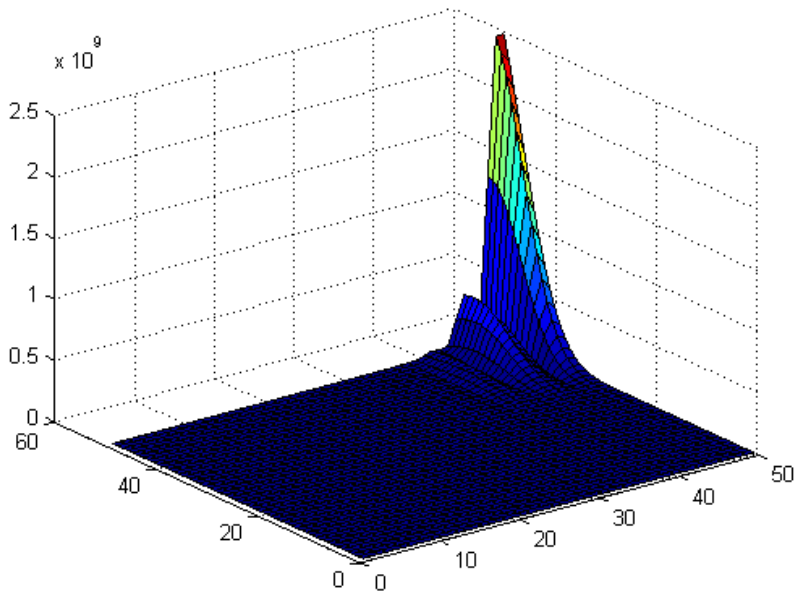


(a) Impulse response

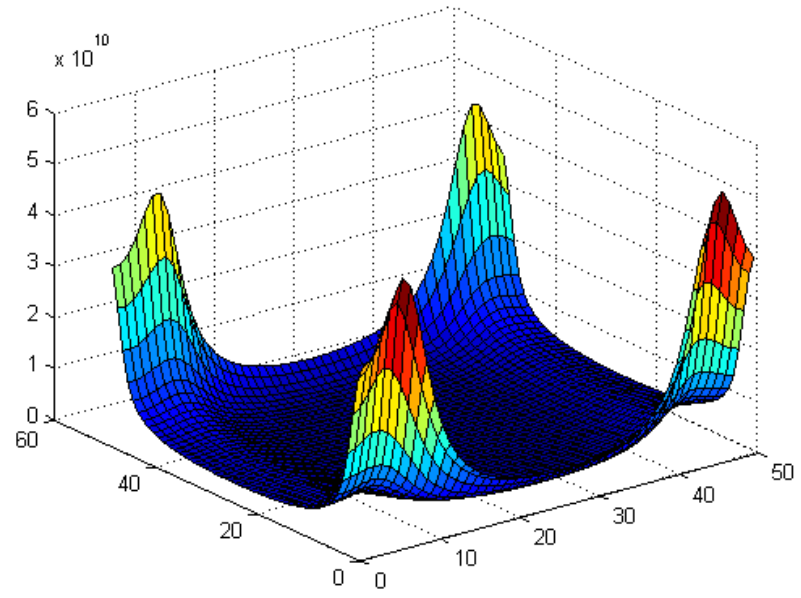


(b) Frequency response (magnitude)

Figure 4.15: Reduced order system (Alternative approach) – Example 3



(a) Error of the impulse response



(b) Error of the frequency response (magnitude) – Example 3

Figure 4.16: Error (Alternative approach) – Example 3

The following table summarizes the comparison between the two presented algorithms for the three examples just considered.

	Sum of the discarded singular vals.		PSNR [dB]		Comput. time [s]	
	B-D	Alt.	B-D	Alt.	B-D	Alt.
Example 1	2.5345	0.7213	36.6647	36.9496	0.47	0.048
Example 2 (out. 1)	0	0.0035	∞	91.5953	0.52	0.066
Example 2 (out. 2)			∞	97.0196		
Example 3	325.5155	1.3756	41.0756	-132.1468	0.55	0.050

Table 4.1: Quantitative comparison between the two model order reduction approaches for the given examples.

Chapter 5

Two case studies

In this chapter, the approaches for model order reduction of 2-D systems presented in section 4.3 are applied to two models with practical relevance.
The first one describes the distribution of the pollution along a river and its evolution along time. Hence, it is intrinsically a 2-D model and the application of both model order reduction algorithms is straightforward.
The second one is the state-space model of an active suspension system of a bus. There, the only considered independent variable is time, hence it is a one-dimensional model. However, one of the parameters on which the system depends has not a fixed value, thus that system is *uncertain*. That example shows how the 2-D model order reduction algorithms described previously may be extended to model order reduction of 1-D uncertain systems.

5.1 River pollution model

A big variety of river water quality problems are caused by pollutants that are discharged into the river as a consequence of human activities. Organic compounds included in these discharges are thereby oxidized by bacteria, algae and fish and converted into inorganic substances and heat.
A simple approach is reducing the variety of polluting compounds to one class of substances and evaluating their concentration by the amount of oxygen needed for their complete biochemical degradation. Such quantity is designated by *biological oxygen demand* (BOD). Another variable that plays an important role in this process is the *dissolved oxygen* (DO) *concentration*. Besides providing a good criterion for water quality, this quantity occurs naturally in the selfpurification process of the river.
Based on these considerations, E. Fornasini proposed in [30] a two-dimensional model that describes the variation of the BOD concentration and DO deficit (with respect to the saturation level) along time and space. The model assumes that the variations of BOD and DO along the cross sections of the river are much less relevant than the variations along the longitudinal one, thus a single space variable is enough to describe the dynamics of the system. Moreover, in the model that is presented here, the stream velocity of the river is assumed to be constant all over the river

stretch.
Let $\beta(t,l)$ and $\delta(t,l)$ denote, respectively, BOD concentration and DO deficit (with respect to the saturation level) that exist in the elementary river reach centered in l at time t. BOD and DO values at $(t+\Delta t, l+\Delta l)$ are computed taking into account [30]:

1. The self purification process, due to the degradation of the discharged pollutants, which results in a decrease of $a_1\beta(t,l)\Delta t$ in the BOD concentration and in a decrease of the DO deficit by the same amount.

2. The reaeration process, which takes place at the water/atmosphere interface, and that is assumed to result in a decrease in the DO deficit of $a_2\delta(t,l)\Delta t$.

3. The diffusion, which is a natural process in which the molecules of a substance move along their concentration gradient (i.e., from a more concentrated area to a less concentrated area). This is modelled assuming that the BOD concentration and DO deficit of the elementary water volume centred on l at time t undergoes variations in Δt that are proportional to $\beta(t,l-\Delta l)-\beta(t,l)$ and $\beta(t,l+\Delta l)-\beta(t,l)$, for BOD, and to $\delta(t,l-\Delta l)-\delta(t,l)$ and $\delta(t,l+\Delta l)-\delta(t,l)$, for DO deficit.

4. BOD sources (effluents, local runoff, etc.) and reoxygenation agents (plants, algae, etc.), respectively $u_\beta(t,l)$ and $u_\delta(t,l)$.

The 2-D model presented here is discrete in both time and space. The time is divided into intervals of Δt and the river is divided into elementary reaches of length Δl, which are related by:

$$\Delta l = v\Delta t, \tag{5.1}$$

where v is the stream velocity. To simplify the notation, $\beta(i\Delta t, j\Delta l)$, $\delta(i\Delta t, j\Delta l)$, $u_\beta(i\Delta t, j\Delta l)$ and $u_\delta(i\Delta t, j\Delta l)$, $(i,j)\in\mathbb{Z}_0^{+^2}$, are, from now on, abusively denoted by $\beta(i,j)$, $\delta(i,j)$, $u_\beta(i,j)$ and $u_\delta(i,j)$, respectively.
The dynamics of the system is described by the following equations:

$$\beta(i+1,j+1) = D_\beta\Delta t\beta(i,j-1)+(1-\bar{a}_1\Delta t)\beta(i,j)+D_\beta\Delta t\beta(i,j+1)+M(1-a_1\Delta t)u_\beta(i,j) \tag{5.2}$$

$$\delta(i+1,j+1) = a_1\Delta t\beta(i,j)+D_\delta(i,j-1)+(1-\bar{a}_2\Delta t)\delta(i,j)+D_\delta\Delta t\delta(i,j+1)-N(1-a_2\Delta t)u_\delta(i,j) \tag{5.3}$$

where $\bar{a}_1 := a_1+2D_\beta$ and $\bar{a}_2 := a_2+2D_\delta$. Defining

$$x_h(i,j) := \begin{bmatrix}\beta(i,j+1)\\ \delta(i,j+1)\end{bmatrix}, \tag{5.4}$$

$$x_v(i,j) := \begin{bmatrix}\beta(i,j-1)\\ \beta(i,j)\\ \delta(i,j-1)\\ \delta(i,j)\end{bmatrix}, \tag{5.5}$$

and choosing $\beta(i,j)$ and $\delta(i,j)$ as the outputs of the system, these equations can be translated into the following Roesser model:

$$\begin{bmatrix} x_h(i+1,j) \\ x_v(i,j+1) \end{bmatrix} = \begin{bmatrix} D_\beta\Delta t & 0 & D_\beta\Delta t & 1-\bar{a}_1\Delta t & 0 & 0 \\ 0 & D_\delta\Delta t & 0 & a_1\Delta t & D_\delta\Delta t & 1-\bar{a}_2\Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_h(i,j) \\ x_v(i,j) \end{bmatrix} + \\ + \begin{bmatrix} M(1-a_1\Delta t) & 0 \\ 0 & -N(1-a_2\Delta t) \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} u(i,j) \tag{5.6}$$

$$y(i,j) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} x(i,j), \tag{5.7}$$

where $u(i,j) := \begin{bmatrix} u_\beta(i,j) & u_\delta(i,j) \end{bmatrix}^T$ and $y(i,j) := \begin{bmatrix} \beta(i,j) & \delta(i,j) \end{bmatrix}^T$.
Here, the following reasonable values are used:

$$v = 0.40\text{m/s},\ a_1 = 25\%/\text{day},\ a_2 = 45\%/\text{day},\ D_\beta = 30\%/\text{day},\ D_\delta = 35\%/\text{day}. \tag{5.8}$$

The time step is $\Delta t = 1$s, thus, according to (5.1), the river is divided into reaches of $\Delta l = 0.40$m. The order of this model will be reduced by using both the Beck-Doyle algorithm and the alternative approach proposed in this thesis. In both, the total order of the model is going to be reduced to 5.

Beck-Doyle algorithm

One has clearly two possible ways to reduce the order of the model from 6 to 5: discarding one dimension either in the horizontal or in the vertical state. Given that the order of the horizontal state is 2 and the order of the vertical state is 4, one can suspect that it is more reasonable to discard one vertical dimension. Computing the diagonal gramians of the balanced realization yields:

$$P_{bal} = Q_{bal} = S = diag(2.2581, 0.9975,\ 2.2581, 1.4677, 0.9975, 0.2953) \times 10^4. \tag{5.9}$$

The optimization criterion that was used for computing the gramians is the same that was used in 4.3.3. The smallest "vertical" singular value is $\sigma_{4_1} = 2.953 \times 10^3$, while the smallest "horizontal" singular value is $\sigma_{2_2} = 9.975 \times 10^3$. Therefore, the error bound for the reduced order system is tighter if a vertical dimension is discarded. In fact, computation of both possibilities confirms that the obtained error is smaller if that option is chosen. For a matter of brevity, only the model

corresponding to that possibility is shown here:

$$\hat{A}_{11} = \begin{bmatrix} 0.3587 & 0.0209 \\ 0.0255 & 0.3936 \end{bmatrix} \times 10^{-4},\ \hat{A}_{12} = \begin{bmatrix} 0.9999 & 0.0060 & -0.0000 \\ -0.0000 & 0.0025 & 0.9999 \end{bmatrix},$$
$$\hat{A}_{21} = \begin{bmatrix} 0.9999 & -0.0000 \\ 0.0095 & 0.0019 \\ -0.0000 & 1.0000 \end{bmatrix},\ \hat{A}_{22} = \begin{bmatrix} 0.0001 & 0.0000 & -0.0000 \\ -0.0075 & -0.0001 & 0.0031 \\ 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \tag{5.10}$$
$$\hat{B}_1 = \begin{bmatrix} 0.8995 & -0.3659 \\ -0.4943 & -0.8120 \end{bmatrix},\ \hat{B}_2 = 0_{3\times 2},$$
$$\hat{C}_1 = 0_{2\times 2},\ \hat{C}_2 = \begin{bmatrix} 1.1292 & -118.7876 & 0.2038 \\ 0.5424 & 0.0076 & 0.9871 \end{bmatrix}.$$

Alternative approach

Here, also, one can discard either one horizontal or one vertical dimension. Computing both the horizontal and the vertical gramians of both (1-D) balanced realizations yields:

$$P^h_{bal} = Q^h_{bal} = S^h = diag(1.4141, 1.4141), \tag{5.11}$$
$$P^v_{bal} = Q^v_{bal} = S^v = diag(1.4141, 1.4141, 1, 2.8646 \times 10^{-5}). \tag{5.12}$$

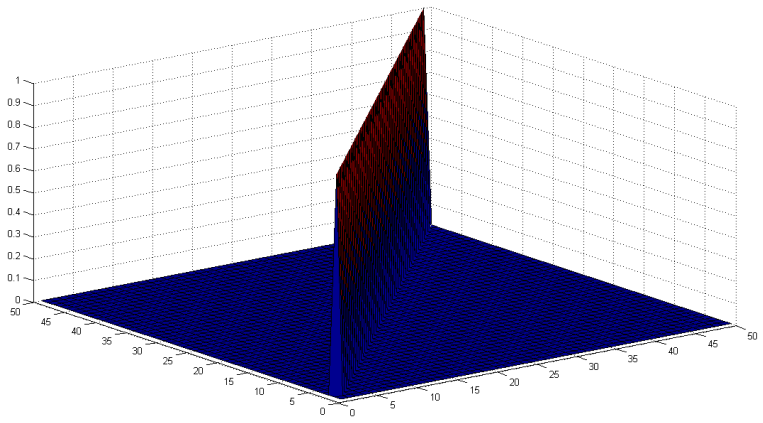
Therefore, once again, discarding one vertical state yields a reduced order model with a much better error bound than discarding a horizontal one and, by computation of both alternatives, this option proved to be actually the most adequate. This reduced order model is defined by the matrices:

$$\hat{A}_{11} = \begin{bmatrix} 0.3472 & 0 \\ 0 & 0.4051 \end{bmatrix} \times 10^{-4},\ \hat{A}_{12} = \begin{bmatrix} -0.7906 & 0.2862 & 0.0000 \\ -0.2862 & -0.7906 & -0.0000 \end{bmatrix},$$
$$\hat{A}_{21} = \begin{bmatrix} -1.1182 & -0.4048 \\ 0.4048 & -1.1181 \\ -0.0000 & 0.0000 \end{bmatrix},\ \hat{A}_{22} = \begin{bmatrix} 0.0000 & -0.0000 & 0.0000 \\ -0.0000 & 0.0000 & -0.0000 \\ -0.7907 & 0.2863 & -0.0000 \end{bmatrix}, \tag{5.13}$$
$$\hat{B}_1 = \begin{bmatrix} 1.0000 & 0 \\ 0 & -0.9999 \end{bmatrix},\ \hat{B}_2 = 0_{3\times 2},$$
$$\hat{C}_1 = 0_{2\times 2},\ \hat{C}_2 = \begin{bmatrix} -0.0000 & 0.0000 & 1.0000 \\ -0.2863 & -0.7907 & 0.0000 \end{bmatrix}.$$

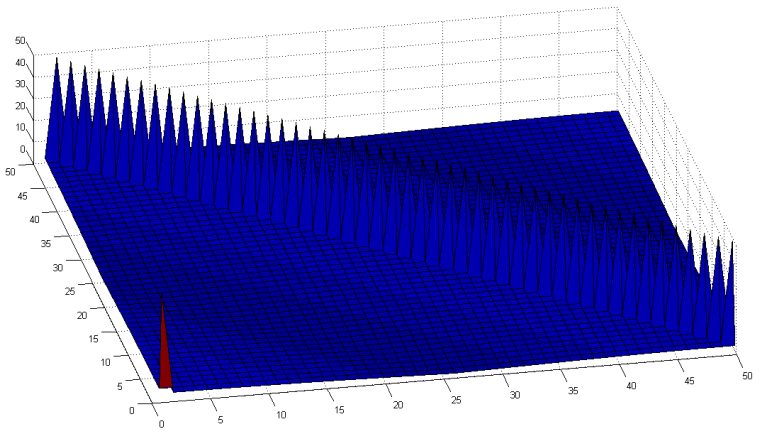
	Discarded singular value		PSNR [dB]		Comput. time [s]	
	B-D	Alt.	B-D	Alt.	B-D	Alt.
Output 1 ($\beta(i,j)$)	2.9530×10^{3}	2.8646×10^{-5}	44.5673	179.1492	0.46	0.043
Output 2 ($\delta(i,j)$)			95.1741	82.6402		

Table 5.1: Quantitative comparison between the approaches for the order reduction of a river pollution model.

In Table 5.1, the performance of both methods in this example is compared. The discarded singular value is much larger in the Beck-Doyle algorithm than in the alternative approach, therefore the error bound is much tighter for the latter. Furthermore, the PSNR for the impulse response of the output $\beta(i,j)$ is much larger in the alternative approach, although for the output $\delta(i,j)$ it is slightly smaller. In terms of computational time, the alternative approach was able to reduce the order of the model in more than 10 times quicker than the Beck-Doyle algorithm, revealing again to be a much faster algorithm. Hence, for this example, it is clearly more advantageous to use the alternative approach. Figures 5.1 and 5.2 show the plots of the impulse and frequency responses of the original model. Figures 5.3-5.6 show the errors of the reduced order models obtained with both approaches.

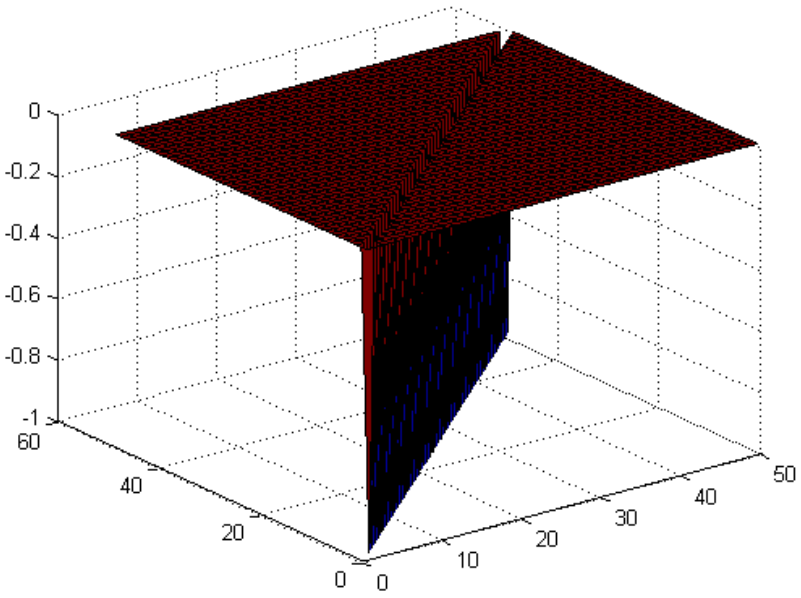


(a) Impulse response

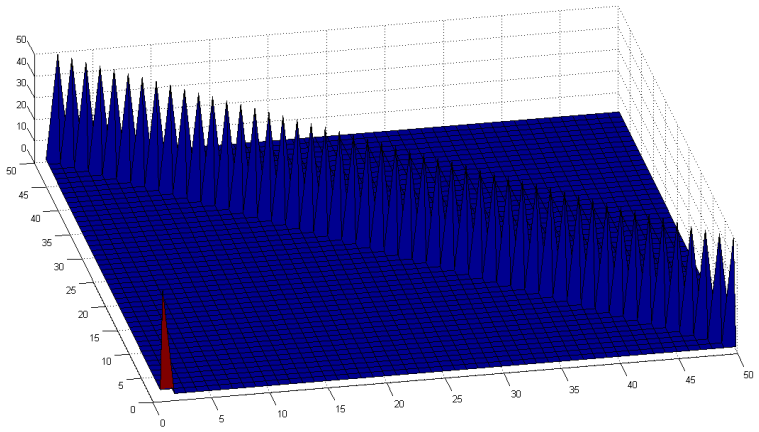


(b) Frequency response (magnitude)

Figure 5.1: Original river model – output $\beta(i, j)$

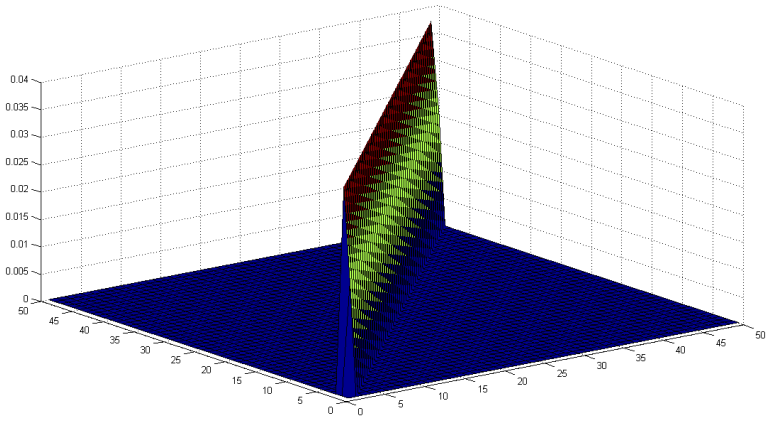


(a) Impulse response

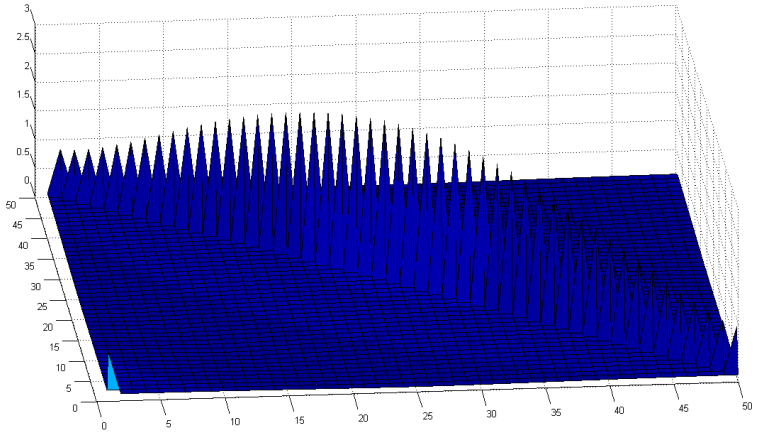


(b) Frequency response (magnitude)

Figure 5.2: Original river model – output $\delta(i, j)$

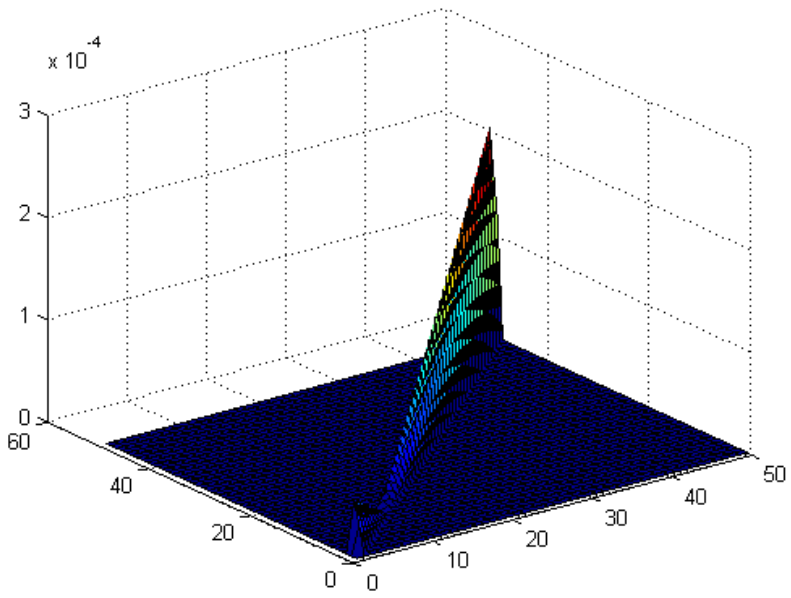


(a) Impulse response

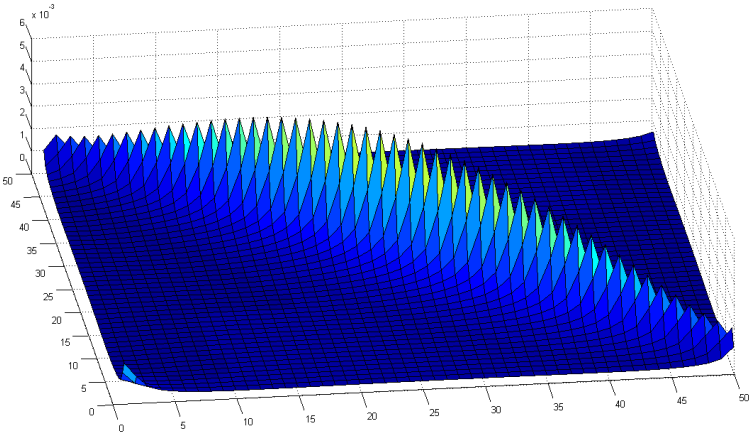


(b) Frequency response (magnitude)

Figure 5.3: Error of the reduced order river model (Beck-Doyle algorithm) – output $\beta(i, j)$

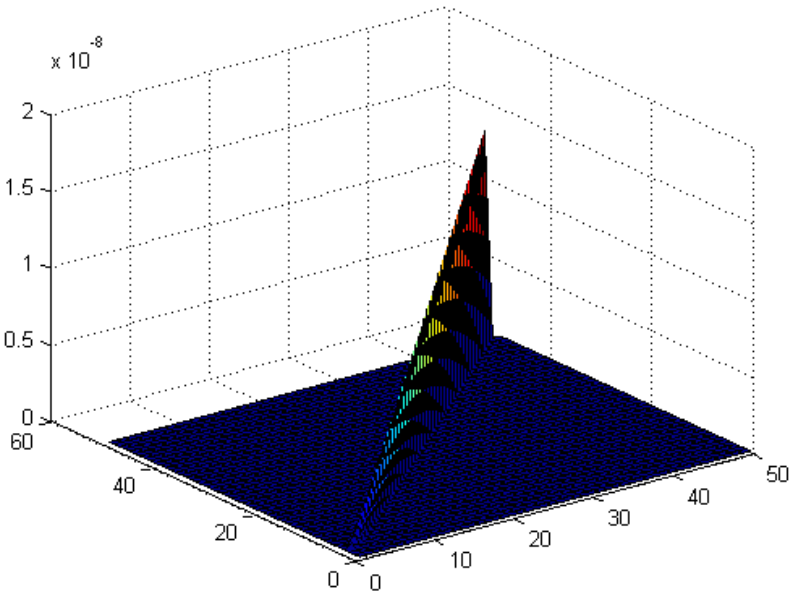


(a) Impulse response

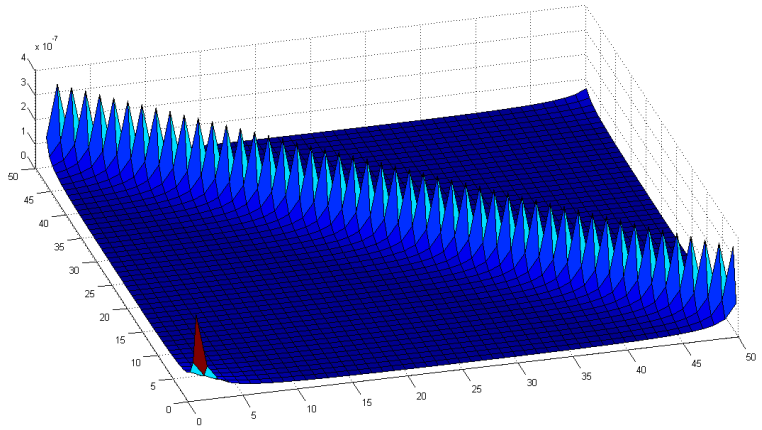


(b) Frequency response (magnitude)

Figure 5.4: Error of the reduced order river model (Beck-Doyle algorithm) – output $\delta(i, j)$

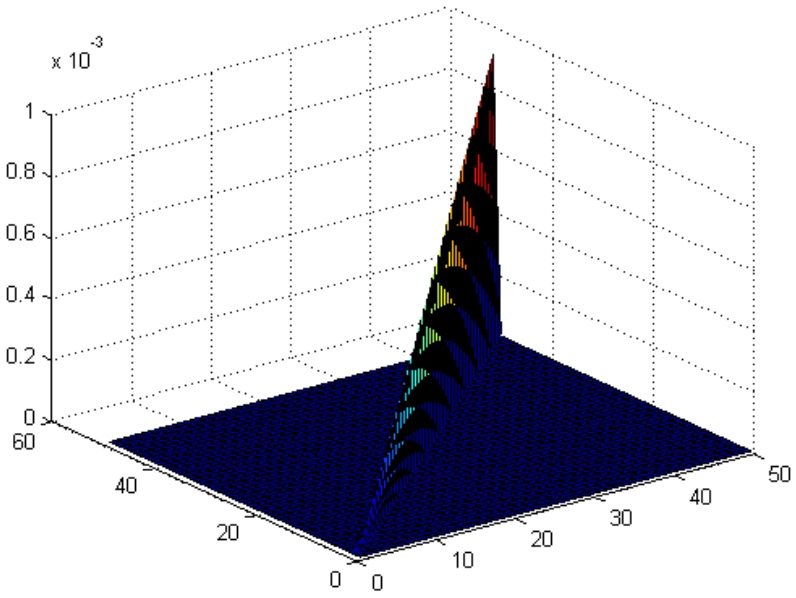


(a) Impulse response

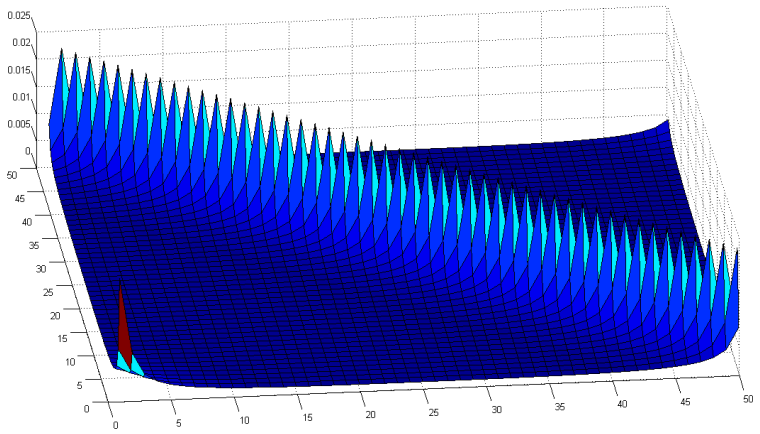


(b) Frequency response (magnitude)

Figure 5.5: Error of the reduced order river model (Alternative approach) – output $\beta(i, j)$



(a) Impulse response



(b) Frequency response (magnitude)

Figure 5.6: Error of the reduced order river model (Alternative approach) – output $\delta(i, j)$

5.2 Active suspension system of a bus

Designing an active suspension system for a bus is an interesting control problem. The focus here will be on reducing the order of the corresponding model by applying 2-D order reduction algorithms. To simplify the problem, only one of the four wheels is considered and modelled as the spring-damper system represented in Figure 5.7 and presented in [31].

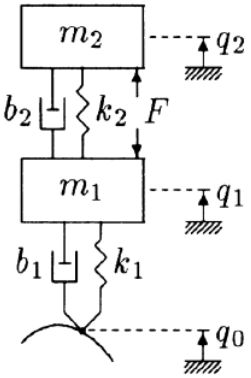


Figure 5.7: Spring-damper system.

Here, m_1 is the suspension mass, m_2 is the body mass, k_1 is the spring constant of the wheel and tire, k_2 is the spring constant of the suspension system, b_1 is the damping constant of the wheel and tire and b_2 is the damping constant of the suspension system. The force F is the control input that is applied to the system, q_0 represents the position of the ground, q_1 the position of the wheel and q_2 the position of the bus body.
This system is described by the second order differential equations:

$$m_2\ddot{q}_2 + b_2(\dot{q}_2 - \dot{q}_1) + k_2(q_2 - q_1) - F = 0, \tag{5.14}$$

$$m_1\ddot{q}_1 + b_1(\dot{q}_1 - \dot{q}_2) + k_2(q_1 - q_2) + k_1(q_1 - q_0) + b_1(\dot{q}_1 - \dot{q}_0) - F = 0. \tag{5.15}$$

The inputs of the model considered here are the position of the ground, q_0, and the applied force, F. The outputs are the body acceleration, $\ddot{q}_2$, the stretch of the suspension, $q_2 - q_1$, and the deformation of the tire, $q_1 - q_0$. A possible state-space realization for this system is:

$$\dot{x}(t) = \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k_1+k_2}{m_1} & \frac{k_2}{m_1} & -\frac{b_1+b_2}{m_1} & \frac{b_2}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{b_2}{m_2} & -\frac{b_2}{m_2} \end{bmatrix}}_{A_c} x(t) + \underbrace{\begin{bmatrix} \frac{b_1}{m_1} & 0 \\ 0 & 0 \\ \frac{-b_1^2-b_1b_2}{m_1^2} + \frac{k_1}{m_1} & -\frac{1}{m_1} \\ \frac{b_1b_2}{m_1m_2} & \frac{1}{m_2} \end{bmatrix}}_{B_c} u(t) \tag{5.16}$$

$$y(t) = \underbrace{\begin{bmatrix} \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{b2}{m_2} & \frac{-b_2}{m_2} \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{C_c} x(t) + \underbrace{\begin{bmatrix} \frac{b_1b_2}{m_1m_2} & \frac{1}{m_2} \\ 0 & 0 \\ -1 & 0 \end{bmatrix}}_{D_c} u(t), \tag{5.17}$$

where $u(t) := \begin{bmatrix} q_0(t) & F(t) \end{bmatrix}^T$ and $y(t) := \begin{bmatrix} \ddot{q}_2(t) & q_2(t)-q_1(t) & q_1(t)-q_0(t) \end{bmatrix}^T$. This is a one-dimensional continuous-time state-space model, denoted by $\Sigma_c = (A_c, B_c, C_c, D_c)$. Here, the sub-index c is used to recall that these matrices belong to a continuous-time system.
The nominal values for the parameters described above are:

$$m_1 = m_2 = 1500\text{Kg},\ \ k_1 = k_2 = 1500\text{N/m},\ \ b_1 = 1\times 10^{-3}\text{Ns/m and } b_2 = 1\times 10^{4}\text{Ns/m}. \quad (5.18)$$

These values correspond to the situation when the car is completely unloaded. When the car is loaded, the mass m_2 can increase up to 1×10^4Kg, but all the other parameters remain constant. Thus, this system has an *uncertainty* in the mass m_2. Taking into account that m_2 belongs to the interval $[1500,\ 1\times 10^4]$, whose middle point is 5750, this uncertainty can be parametrized as follows:

$$m_2(\delta) = 5750 + 4250\delta,\ -1 \le \delta \le 1. \quad (5.19)$$

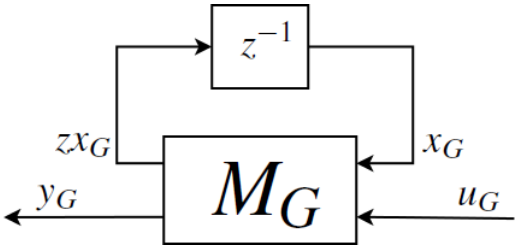
Taking system (5.16)-(5.17), replacing m_2 by the parametrization above and each of the other parameters by their respective values, yields a realization $\Sigma_c = (A_c(\delta), B_c(\delta), C_c(\delta), D_c(\delta))$, where:

$$\begin{aligned}
A_c(\delta) &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -666.7 & 333.3 & -33.33 & 33.33 \\ \frac{5\times 10^5}{5750+4250\delta} & \frac{-5\times 10^5}{5750+4250\delta} & \frac{-5\times 10^4}{5750+4250\delta} & -\frac{-5\times 10^4}{5750+4250\delta} \end{bmatrix}, \\
B_c(\delta) &= \begin{bmatrix} 6.667\times 10^{-7} & 0 \\ 0 & 0 \\ 333.3 & -6.667\times 10^{-4} \\ \frac{50}{8.625\times 10^6+6.375\times 10^6\delta} & \frac{1}{5750+4250\delta} \end{bmatrix}, \\
C_c(\delta) &= \begin{bmatrix} \frac{5\times 10^5}{5750+4250\delta} & -\frac{5\times 10^5}{5750+4250\delta} & \frac{5\times 10^4}{5750+4250\delta} & -\frac{5\times 10^4}{5750+4250\delta} \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \\
D_c(\delta) &= \begin{bmatrix} \frac{50}{8.625\times 10^6+6.375\times 10^6\delta} & \frac{1}{5750+4250\delta} \\ 0 & 0 \\ -1 & 0 \end{bmatrix}.
\end{aligned} \quad (5.20)$$

The given uncertain model Σ_c is said to be *robustly* stable (i.e., stable against all possible values of the uncertainty) if it is stable for all $\delta \in [-1,1]$. It is possible to verify that this is the case, so this model is robustly stable.
Since this model depends on the parameter δ, performing its order reduction is not a trivial task. Note that if one had a model order reduction method for first level realizations of 2-D systems, one could possibly apply it here, since these matrices can be seen as rational matrices in the variable δ. However, such method does not exist, therefore a different approach must be used.
The order of this model will be reduced by using the algorithms for model order reduction de-

scribed in section 4.3.

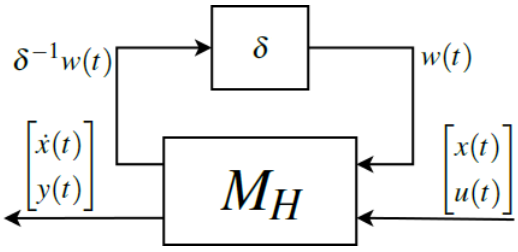
Before that, though, the relation between the current uncertain system and a 2-D system must be clarified. It is easy to verify that a realization $\Sigma_G = (A_G, B_G, C_G, D_G)$ of a transfer matrix $G(z)$ that maps u_G into y_G can be seen as the following block interconnection:



where $M_G = \begin{bmatrix} A_G & B_G \\ C_G & D_G \end{bmatrix}$ and x_G is the state vector. Hence, finding a realization for

$$H_c(\bar{\delta}) := \begin{bmatrix} A_c(\bar{\delta}) & B_c(\bar{\delta}) \\ C_c(\bar{\delta}) & D_c(\bar{\delta}) \end{bmatrix}, \tag{5.21}$$

where $\bar{\delta} := \delta^{-1}$, allows representing the uncertain system as the following block interconnection:



Here, M_H is defined analogously to M_G, but considering a realization for $H_c(\bar{\delta})$. The computation of these blocks is commonly designated as *pulling-out* the uncertainty. Such realization can be translated into a model which is structurally similar to a two-dimensional one:

$$\begin{bmatrix} \dot{x}(t) \\ \bar{\delta} w(t) \end{bmatrix} = \begin{bmatrix} A_{11c} & A_{12c} \\ A_{21c} & A_{22c} \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix} + \begin{bmatrix} B_{1c} \\ B_{2c} \end{bmatrix} u(t) \tag{5.22}$$

$$y(t) = \begin{bmatrix} C_{1c} & C_{2c} \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix} + D_c u(t). \tag{5.23}$$

In order to obtain a description of the form (5.22)-(5.23), the model Σ is now rewritten as a function

of $\bar{\delta}$:

$$
\begin{aligned}
A_c(\bar{\delta}) &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -666.7 & 333.3 & -33.33 & 33.33 \\ \frac{5\times10^5\bar{\delta}}{5750\bar{\delta}+4250} & \frac{-5\times10^5\bar{\delta}}{5750\bar{\delta}+4250} & \frac{-5\times10^4\bar{\delta}}{5750\bar{\delta}+4250} & -\frac{-5\times10^4\bar{\delta}}{5750\bar{\delta}+4250} \end{bmatrix}, \\
B_c(\bar{\delta}) &= \begin{bmatrix} 6.667\times10^{-7} & 0 \\ 0 & 0 \\ 333.3 & -6.667\times10^{-4} \\ \frac{50\bar{\delta}}{8.625\times10^6\bar{\delta}+6.375\times10^6} & \frac{\bar{\delta}}{5750\bar{\delta}+4250} \end{bmatrix}, \\
C_c(\bar{\delta}) &= \begin{bmatrix} \frac{5\times10^5\bar{\delta}}{5750\bar{\delta}+4250} & -\frac{5\times10^5\bar{\delta}}{5750\bar{\delta}+4250} & \frac{5\times10^4\bar{\delta}}{5750\bar{\delta}+4250} & -\frac{5\times10^4\bar{\delta}}{5750\bar{\delta}+4250} \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \\
D_c(\bar{\delta}) &= \begin{bmatrix} \frac{50\bar{\delta}}{8.625\times10^6\bar{\delta}+6.375\times10^6} & \frac{\bar{\delta}}{5750\bar{\delta}+4250} \\ 0 & 0 \\ -1 & 0 \end{bmatrix}.
\end{aligned}
\tag{5.24}
$$

A minimal realization of $H_c(\bar{\delta})$ yields:

$$
\begin{aligned}
A_{11c} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -666.6667 & 333.3333 & -33.3333 & 33.3333 \\ 86.9565 & -86.9565 & 8.6957 & -8.6957 \end{bmatrix}, \; A_{12c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 11.3137 \end{bmatrix}, \\
A_{21c} &= \begin{bmatrix} -5.6809 & 5.6809 & -0.5681 & 0.5681 \end{bmatrix}, \; A_{22c} = -0.7391, \\
B_{1c} &= \begin{bmatrix} 0.0000 & 0 \\ 0 & 0.0000 \\ 333.3333 & -0.0007 \\ 0.0000 & 0.0002 \end{bmatrix}, \; B_{2c} = \begin{bmatrix} -0.0038 & -0.1136 \end{bmatrix} \times 10^{-4}, \\
C_{1c} &= \begin{bmatrix} 86.9565 & -86.9565 & 8.6957 & -8.6957 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; C_{2c} = \begin{bmatrix} 11.3137 \\ 0 \\ 0 \end{bmatrix}, \\
D_c &= \begin{bmatrix} 0.0000 & 0.0002 \\ 0 & 0 \\ -1 & 0 \end{bmatrix}.
\end{aligned}
\tag{5.25}
$$

Since the model order reduction procedures to be applied here were developed for discrete 2-D models, the model (5.22)-(5.23) must be discretized. For this purpose, it is useful to rewrite this

model as:

$$\dot{x}(t) = A_{11c}x(t) + \begin{bmatrix} A_{12c} & B_{1c} \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix} \tag{5.26}$$

$$\begin{bmatrix} \bar{\delta}w(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A_{21c} \\ C_{1c} \end{bmatrix} x(t) + \begin{bmatrix} A_{22c} & B_{2c} \\ C_{2c} & D_c \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix} \tag{5.27}$$

This model has the structure of a 1-D model with input $\begin{bmatrix} w(t) & u(t)^T \end{bmatrix}^T$, output $\begin{bmatrix} \bar{\delta}w(t) & y(t)^T \end{bmatrix}^T$ and state vector $x(t)$. Therefore, it can be discretized by using any traditional discretization procedure. Here, zero-order hold is used. The matrices of the corresponding discretized model $\Sigma_d^{2D} = (A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, D)$ are related with the ones of the continuous time model by the following equalities (see, for instance, [32]):

$$A_{11} = e^{A_{11c}T_s} := \sum_{k=0}^{\infty} \frac{A_{11c}T_s^k}{k!}, \tag{5.28}$$

$$\begin{bmatrix} A_{12} & B_1 \end{bmatrix} = \int_0^{T_s} e^{A_{11c}\tau} \begin{bmatrix} A_{12c} & B_{1c} \end{bmatrix} d\tau, \tag{5.29}$$

$$\begin{bmatrix} A_{21} \\ C_1 \end{bmatrix} = \begin{bmatrix} A_{21c} \\ C_{1c} \end{bmatrix}, \tag{5.30}$$

$$\begin{bmatrix} A_{22} & B_2 \\ C_2 & D \end{bmatrix} = \begin{bmatrix} A_{22c} & B_{2c} \\ C_{2c} & D_c \end{bmatrix}, \tag{5.31}$$

where T_s denotes the chosen sampling period. Here, $T_s = 0.01$s and the matrices of the discretized model are:

$$\begin{aligned}
A_{11} &= \begin{bmatrix} 0.9706 & 0.0145 & 0.0085 & 0.0015 \\ 0.0033 & 0.9962 & 0.0004 & 0.0096 \\ -5.5045 & 2.6872 & 0.7019 & 0.2832 \\ 0.5752 & -0.7054 & 0.0739 & 0.9257 \end{bmatrix}, \; A_{12} = \begin{bmatrix} 0.0001 \\ 0.0006 \\ 0.0169 \\ 0.1087 \end{bmatrix}, \\
A_{21} &= \begin{bmatrix} -5.6809 & 5.6809 & -0.5681 & 0.5681 \end{bmatrix}, \; A_{22} = -0.7391, \\
B_1 &= \begin{bmatrix} 0.0149 & -0.0000 \\ 0.0004 & 0.0000 \\ 2.8173 & -0.0000 \\ 0.1302 & 0.0000 \end{bmatrix}, \; B_2 = \begin{bmatrix} -0.0038 & -0.1136 \end{bmatrix} \times 10^{-4}, \\
C_1 &= \begin{bmatrix} 86.9565 & -86.9565 & 8.6957 & -8.6957 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} C_2 = \begin{bmatrix} 11.3137 \\ 0 \\ 0 \end{bmatrix}, \\
D &= \begin{bmatrix} 0.0000 & 0.0002 \\ 0 & 0 \\ -1 & 0 \end{bmatrix}.
\end{aligned} \tag{5.32}$$

Now, the order of the model may be reduced. Before that, though, it is important to relate the internal stability of this 2-D structured model, Σ_d^{2D}, with the robust stability of the uncertain model Σ_c.

Defining

$$A_{2D} := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{5.33}$$

and

$$Z := \begin{bmatrix} zI_2 & 0 \\ 0 & \bar{\delta} \end{bmatrix}, \tag{5.34}$$

one has that Σ_d^{2D} is internally stable if and only if $(Z - A_{2D})$ is nonsingular for all $(z, \bar{\delta}) \in \mathbb{C}^2 : |z| \geq 1, |\bar{\delta}| \geq 1$. It was stated previously that the uncertain model Σ_c is robustly stable if it is stable for all $\delta \in [-1, 1]$. Equivalently, since $\bar{\delta} = \delta^{-1}$, Σ_c is robustly stable if it is stable for all $\bar{\delta} \in \mathbb{R}$ such that $|\bar{\delta}| \geq 1$. In fact, the uncertainty considered in this example is intrinsically real (i.e., it makes no sense to admit complex values for the value of the mass m_2), thus one should not expect that the robust stability of the uncertain model Σ_c implies the internal stability of the 2-D model Σ_d^{2D}. Therefore, there is no guarantee that the Beck-Doyle algorithm is applicable in this case. In fact, when one tries to compute solutions for (4.80) and (4.81), one concludes that these LMI are infeasible. Hence, Σ_d^{2D} is internally unstable and the Beck-Doyle algorithm is not applicable.

Fortunately, one has that $\rho(A_{11}) = 0.9917 < 1$ and $\rho(A_{22}) = 0.7391 < 1$, so the alternative approach proposed in this thesis may be used. Clearly, since the dimension of the state vector associated with the uncertainty (i.e., the dimension of A_{22}) is one, it is pointless to reduce the order of this state. The dimension of the state vector of the nominal system (i.e., the dimension of A_{11}) is 4 and it will be reduced to 3. The reduction is made as in the previous examples and yields the reduced order model $\hat{\Sigma}_d^{2D} = (\hat{A}_{11}, \hat{A}_{12}, \hat{A}_{21}, A_{22}, \hat{B}_1, B_2, \hat{C}_1, C_2, D)$, where:

$$\begin{gathered} \hat{A}_{11} = \begin{bmatrix} 0.9925 & -0.0791 & 0.0452 \\ 0.0789 & 0.9677 & 0.0857 \\ -0.0439 & 0.0833 & 0.7120 \end{bmatrix}, \ \hat{A}_{12} = \begin{bmatrix} 0.2653 \\ -0.2598 \\ -0.2255 \end{bmatrix}, \\ \hat{A}_{21} = \begin{bmatrix} 0.0710 & 0.1498 & -0.2993 \end{bmatrix}, \ \hat{B}_1 = \begin{bmatrix} 1.0528 & 0.0000 \\ -2.2802 & 0.0000 \\ 4.5918 & -0.0000 \end{bmatrix}, \\ \hat{C}_1 = \begin{bmatrix} -1.0869 & -2.2930 & 4.5821 \\ 0.0000 & 0.0192 & 0.0010 \\ 0.0270 & 0.0114 & 0.0055 \end{bmatrix}. \end{gathered} \tag{5.35}$$

The discarded singular value is $\sigma_4^h = 11.8107$. The reduced order discrete-time uncertain state-space model is obtained by computing a first level realization corresponding to the model $\hat{\Sigma}_d^{2D}$ and replacing $\bar{\delta}$ by δ^{-1}. Such model is $\hat{\Sigma}_d = (\hat{A}(\delta), \hat{B}(\delta), \hat{C}(\delta), \hat{D}(\delta))$, where these matrices are

obtained through:

$$\hat{A}(\delta) = \hat{A}_{12}(\bar{\delta}I - A_{22})^{-1}\hat{A}_{21} + \hat{A}_{11}\,\big|_{\bar{\delta}=\delta^{-1}}, \tag{5.36}$$

$$\hat{B}(\delta) = \hat{A}_{12}(\bar{\delta}I - A_{22})^{-1}B_2 + \hat{B}_1\,\big|_{\bar{\delta}=\delta^{-1}}, \tag{5.37}$$

$$\hat{C}(\delta) = C_2(\bar{\delta}I - A_{22})^{-1}\hat{A}_{21} + \hat{C}_1\,\big|_{\bar{\delta}=\delta^{-1}}, \tag{5.38}$$

$$\hat{D}(\delta) = C_2(\bar{\delta}I - A_{22})^{-1}B_2 + D\,\big|_{\bar{\delta}=\delta^{-1}}. \tag{5.39}$$

Performing these calculations, one obtains finally:

$$\begin{aligned}
\hat{A}(\delta) &= \begin{bmatrix} \frac{0.7525\delta+0.9925}{0.7391\delta+1} & \frac{-0.0187\delta-0.07907}{0.7391\delta+1} & \frac{-0.04602\delta+0.04518}{0.7391\delta+1} \\ \frac{0.03985\delta+0.07887}{0.7391\delta+1} & \frac{0.6764\delta+0.9677}{0.7391\delta+1} & \frac{0.1411\delta+0.08568}{0.7391\delta+1} \\ \frac{-0.04843\delta-0.04386}{0.7391\delta+1} & \frac{0.02779\delta+0.08331}{0.7391\delta+1} & \frac{0.5938\delta+0.712}{0.7391\delta+1} \end{bmatrix}, \\
\hat{B}(\delta) &= \begin{bmatrix} \frac{0.7781\delta+1.053}{0.7391\delta+1} & \frac{-1.556\times10^{-6}\delta+1.972\times10^{-6}}{0.7391\delta+1} \\ \frac{-1.685\delta-2.28}{0.7391\delta+1} & \frac{3.371\times10^{6}\delta+5.666\times10^{-7}}{0.7391\delta+1} \\ \frac{3.394\delta+4.592}{0.7391\delta+1} & \frac{-6.788\times10^{-6}\delta-1.265\times10^{-5}}{0.7391\delta+1} \end{bmatrix}, \\
\hat{C}(\delta) &= \begin{bmatrix} \frac{4.441\times10^{-16}\delta-1.087}{0.7391\delta+1} & \frac{-2.293}{0.7391\delta+1} & \frac{-1.332\times10^{-15}\delta+4.582}{0.7391\delta+1} \\ 1.779\times10^{-5} & 0.01916 & 0.0009805 \\ 0.02695 & 0.01144 & 0.005477 \end{bmatrix}, \\
\hat{D}(\delta) &= \begin{bmatrix} \frac{-8.47\times10^{-22}\delta+5.797\times10^{-6}}{0.7391\delta+1} & \frac{-2.711\times10^{-20}\delta+0.0001739}{0.7391\delta+1} \\ 0 & 0 \\ -1 & 0 \end{bmatrix}.
\end{aligned} \tag{5.40}$$

In figures 5.8 and 5.9, the responses of the non-reduced discretized uncertain model to independent impulses of 1m in q_2 and 15kN in F are displayed for various admissible values of δ. Note that, due to the type of discretization that was used, a discrete-time impulse of amplitude K corresponds to a continuous-time rectangular pulse $s(t)$ of amplitude K and width T_s, i.e.:

$$s(t) := \begin{cases} K, & 0 \le t \le T_s \\ 0, & t > T_s \end{cases}. \tag{5.41}$$

The error of the reduced order system for the same inputs is plotted in figures 5.10 and 5.11. Observing these error plots, one concludes that the error is considerably larger when δ is close to -1, i.e., when the mass m_2 is close to its minimum possible value.
The relative error from input k to output l is defined as:

$$\varepsilon_{k,l} := \frac{\sum_{n=0}^{N_s} |h_{k,l}(n) - \hat{h}_{k,l}(n)|}{\sum_{n=0}^{N_s} |h_{k,l}(n)|}, \tag{5.42}$$

where $h_{k,l}$ and $\hat{h}_{k,l}$ are the impulse responses from input k to output l of the original and reduced order systems, respectively, and N_s is the total number of samples. In Table 5.2 the values of the *relative error* from each input to each output are presented, for $\delta = -1, 0$ and 1.

	Input 1 (q_2)			Input 2 (F)		
	$\delta=-1$	$\delta=0$	$\delta=1$	$\delta=-1$	$\delta=0$	$\delta=1$
Output 1 ($\ddot{q}_2$)	0.6693	0.2231	0.4337	0.7559	0.4321	1.0998
Output 2 (q_2-q_1)	0.9320	0.5343	0.6856	1.2059	1.1415	1.3501
Output 3 (q_1-q_0)	0.6700	0.1983	0.3826	1.1260	0.6873	1.3008

Table 5.2: Relative error of the impulse responses of the reduced order system

The values in this table confirm what was stated before: the error is actually large for $\delta=-1$. For the input F, the error is even larger for $\delta=1$. Furthermore, the table shows that the error for the input q_2 is smaller than the error for the input F.

As it is clear through the analysis of Table 5.2, the approximation obtained with this reduced order system yields a rather large error for the input F and also for the input q_2 when $\delta=-1$. It might be interesting, therefore, to compare these results with those obtained by performing the 1-D balanced truncation reduction for the original system when one takes $\delta=-1,0$ and 1. Note that, when one replaces δ by a constant value, the obtained system is not an uncertain system anymore, thus the "classic" 1-D model order reduction algorithm may be applied. However, the obtained system is an approximation of the original one only for that pre-specified value of δ. For that reason, it will be called a *local* reduced order model.

In figures 5.12-5.17, the responses of the original and reduced order uncertain and local models are compared for $\delta=-1,0$ and 1. In Table 5.3 the relative errors of the local reduced order systems are presented, for the same values of δ.

	Input 1 (q_2)			Input 2 (F)		
	$\delta=-1$	$\delta=0$	$\delta=1$	$\delta=-1$	$\delta=0$	$\delta=1$
Output 1 ($\ddot{q}_2$)	0.0260	0.2288	0.3265	0.1824	0.4390	0.5022
Output 2 (q_2-q_1)	0.1540	0.5572	0.6342	0.7361	1.1662	1.1733
Output 3 (q_1-q_0)	0.0208	0.2051	0.2907	0.2895	0.6980	0.7492

Table 5.3: Relative errors of the impulse responses of the local reduced order systems

As this table shows, for $\delta=-1$ and 1, the errors of the local reduced order systems are rather smaller than the errors of the uncertain reduced order system. This could be expected *a priori*. In fact, in the computation of the uncertain reduced order model, there is an implicit constraint: the entries of the uncertain model must be proper rational functions in the variable $\bar{\delta}=\delta^{-1}$ in which the degree of the denominator may be at most one. Hence, these entries should not be expected to interpolate exactly the entries of each local reduced order model that may be computed for each

value of δ in the interval $[-1,\ 1]$. Therefore, the obtained error is expected to be larger for the uncertain reduced order model. However, for $\delta = 0$, the relative errors of both the uncertain reduced order model and the local reduced order model are very close. Figures 5.13 and 5.16 confirm that the local reduced order model and the uncertain reduced order model impulse responses almost coincide for $\delta = 0$.

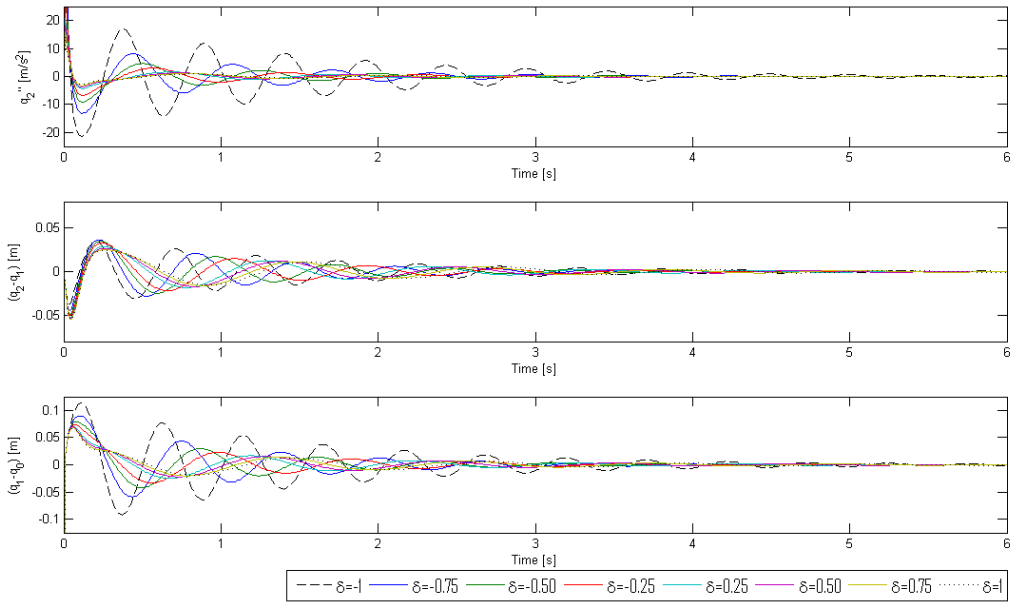


Figure 5.8: Response of the original system to an impulse of 1m in q_2, for the indicated values of δ.

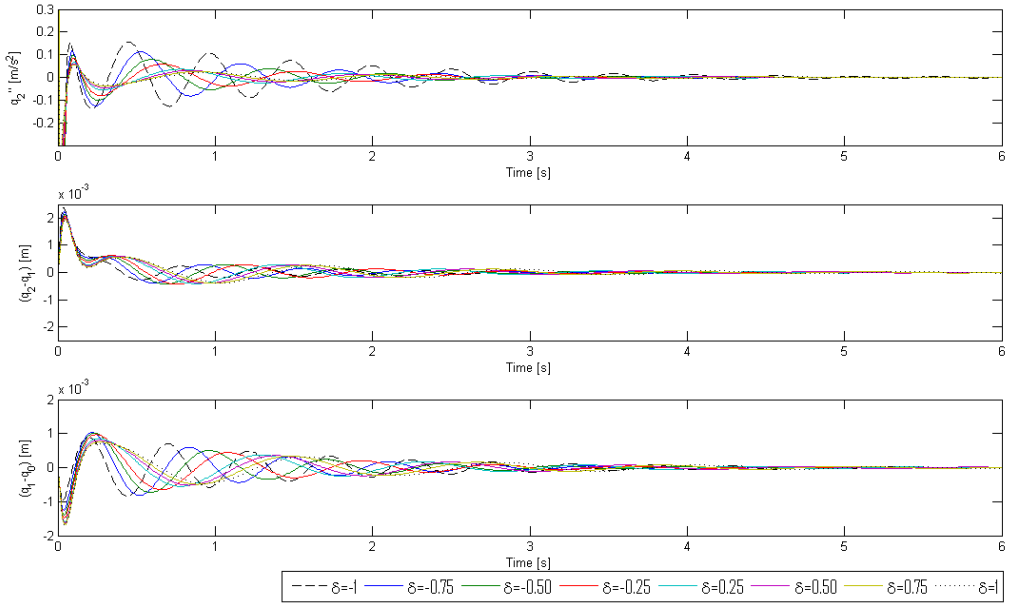


Figure 5.9: Response of the original system to an impulse of 15kN in F, for the indicated values of δ.

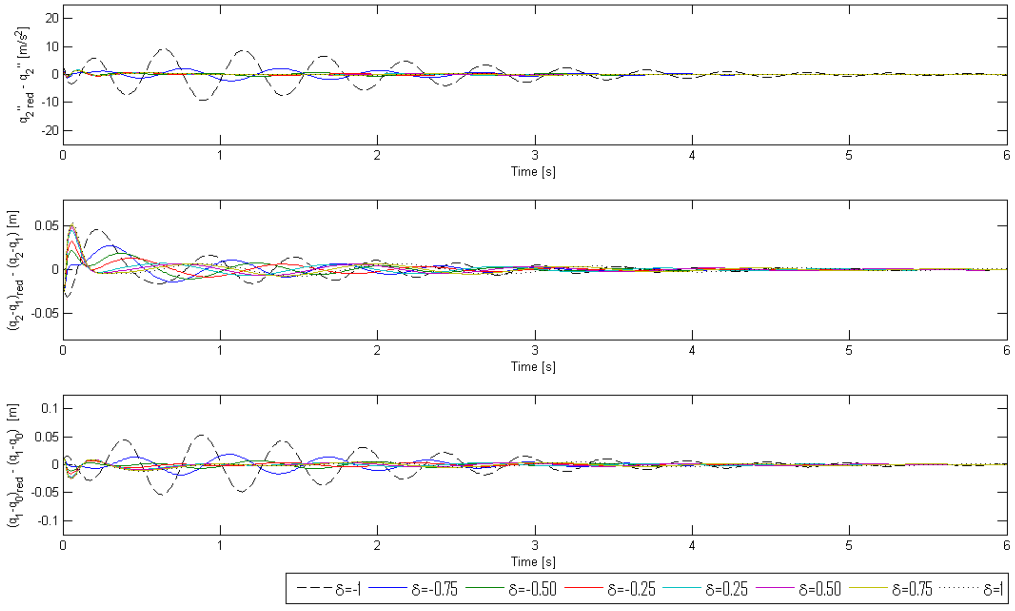


Figure 5.10: Error of the response of the uncertain reduced order system to an impulse of 1m in q_2, for the indicated values of δ.

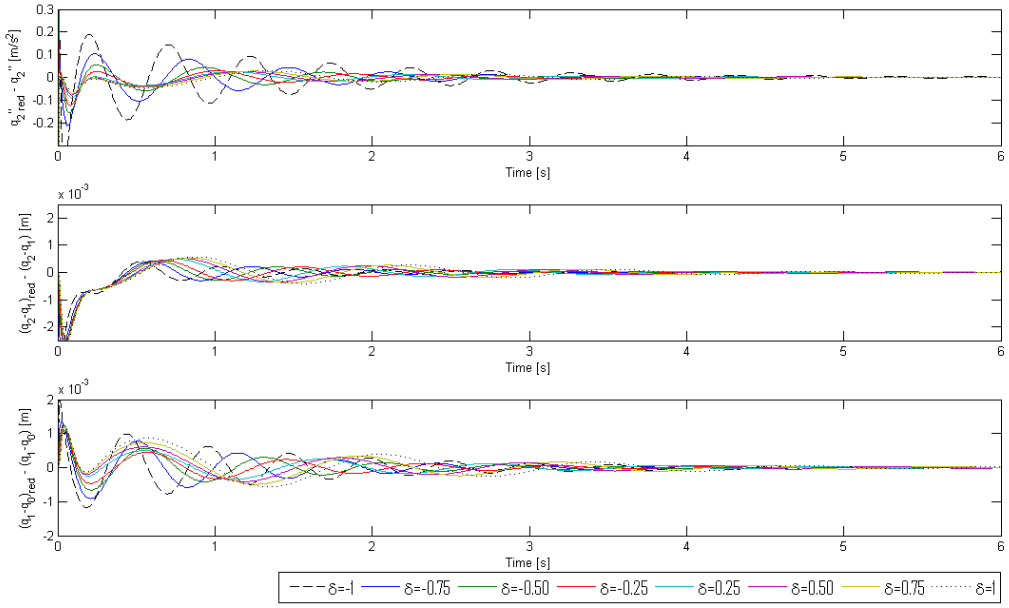


Figure 5.11: Error of the response of the uncertain reduced order system to an impulse of 15kN in F, for the indicated values of δ.

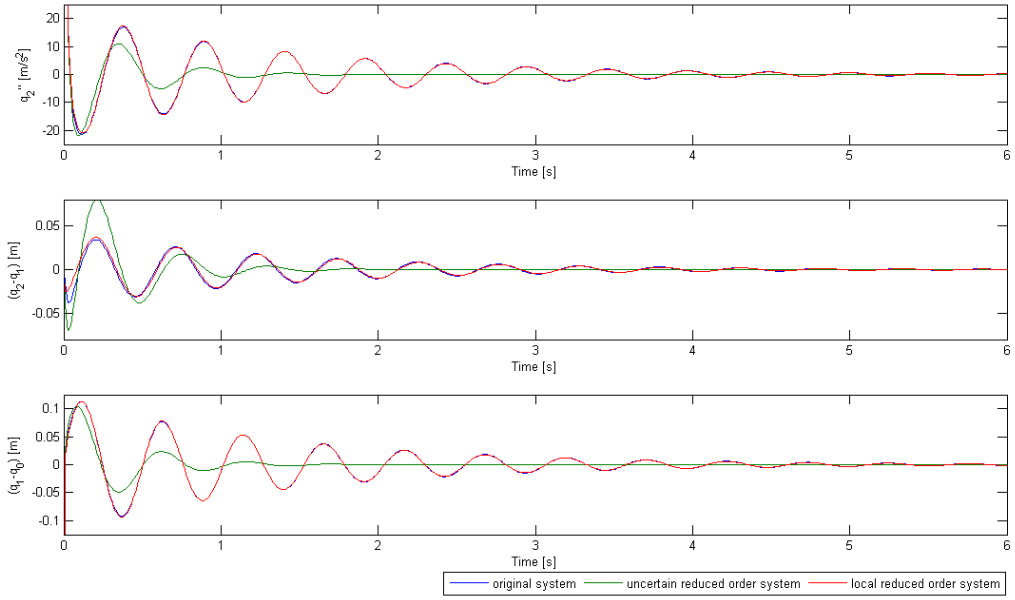


Figure 5.12: Responses of the original and reduced order systems (uncertain and local) to an impulse of 1m in q_2, for $\delta = -1$.

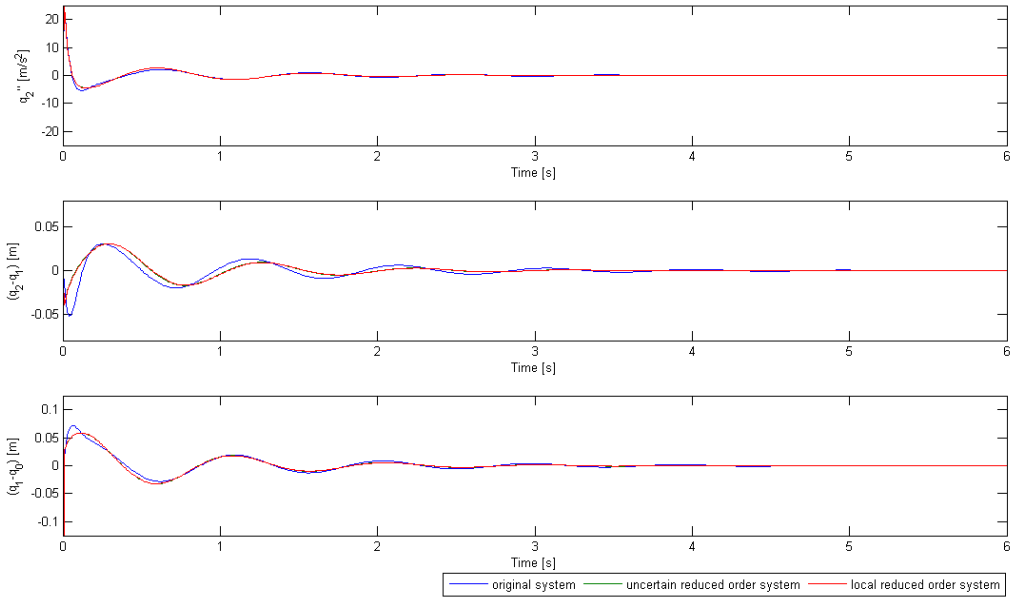


Figure 5.13: Responses of the original and reduced order systems (uncertain and local) to an impulse of 1m in q_2, for $\delta = 0$.

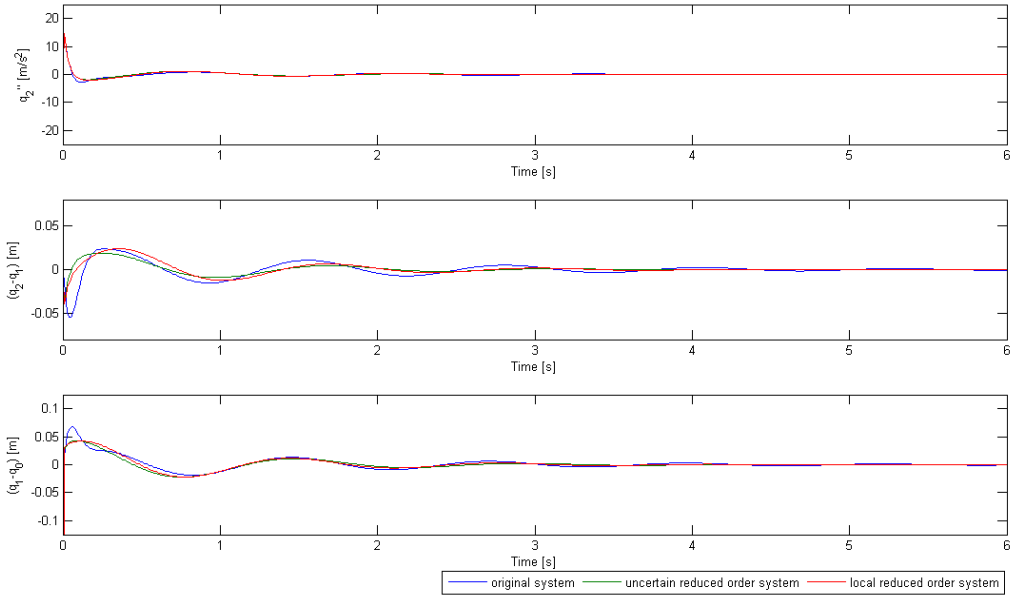


Figure 5.14: Responses of the original and reduced order systems (uncertain and local) to an impulse of 1m in q_2, for $\delta = 1$.

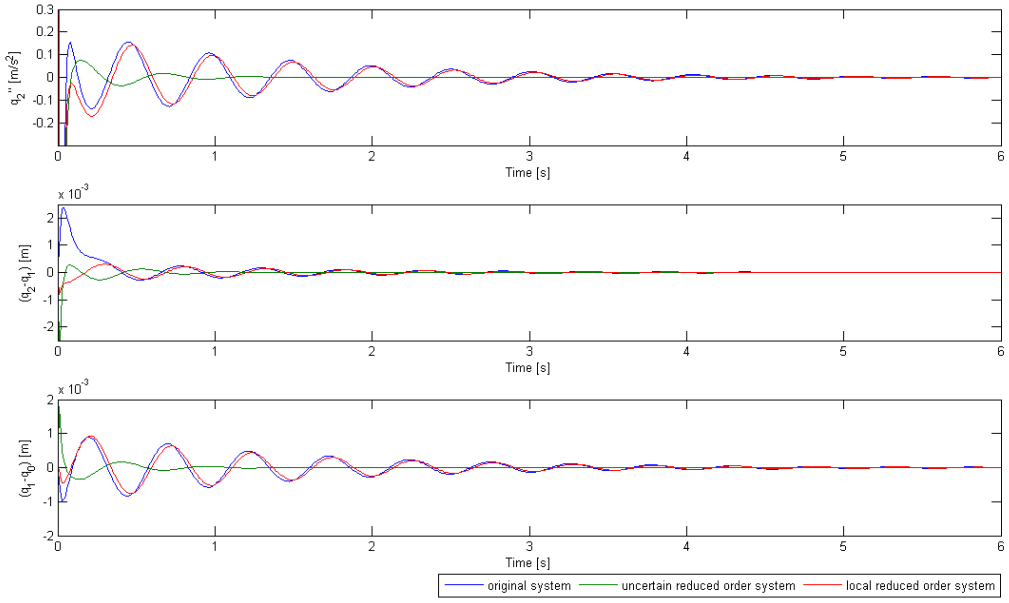


Figure 5.15: Responses of the original and reduced order systems (uncertain and local) to an impulse of 15kN in F, for $\delta = -1$.

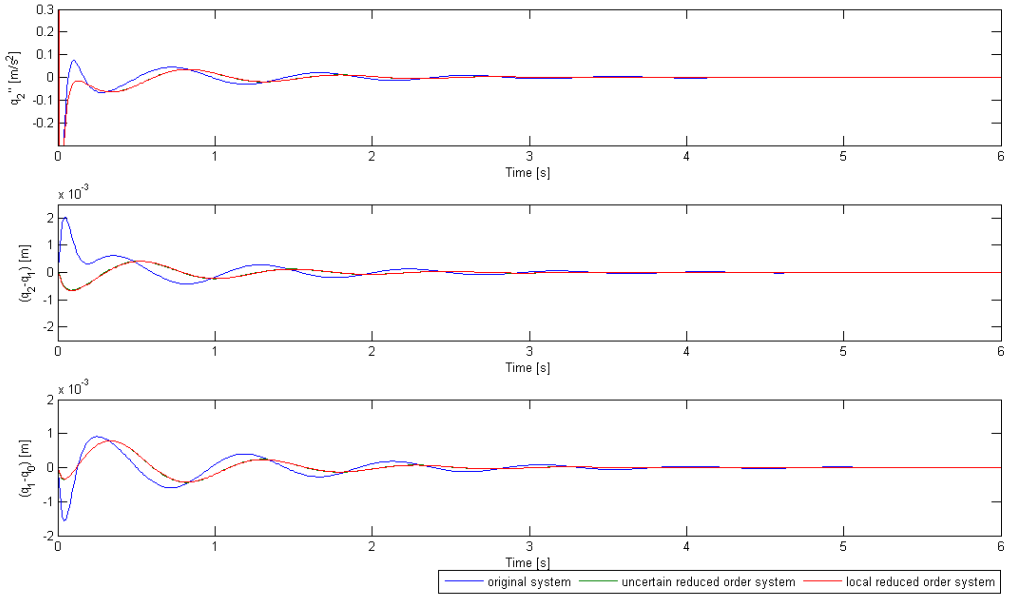


Figure 5.16: Responses of the original and reduced order systems (uncertain and local) to an impulse of 15kN in F, for $\delta = 0$.

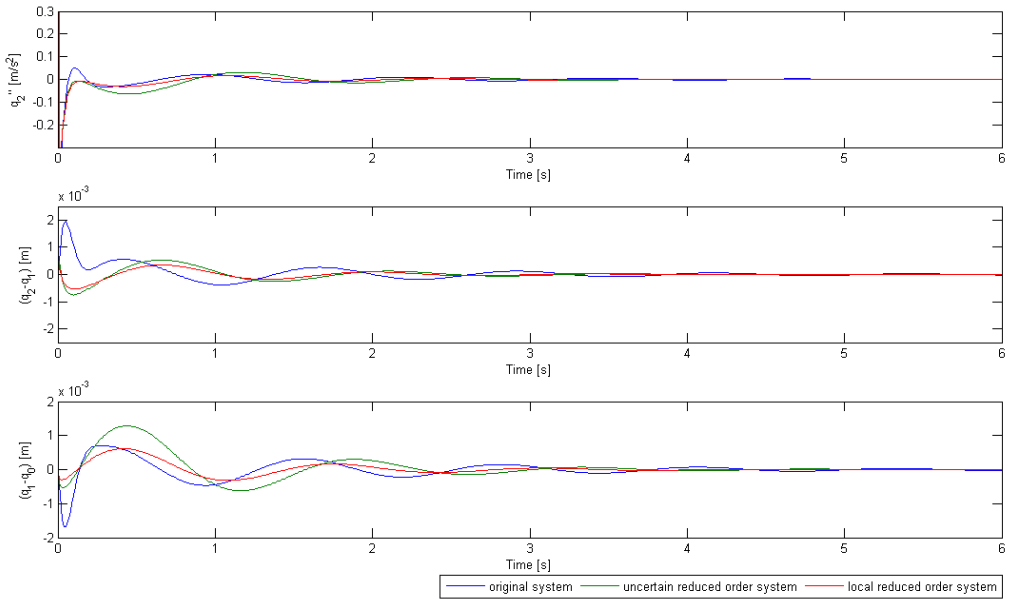


Figure 5.17: Responses of the original and reduced order systems (uncertain and local) to an impulse of 15kN in F, for $\delta = 1$.

Chapter 6

Conclusions and future work

In this thesis, some results on 2-D state-space realization theory were presented, with a special focus on the subject of model order reduction. In this topic, two possible approaches were explored: order reduction of second level realizations and order reduction of first level realizations.

The former approach has been followed previously by many authors (see section 4.1) since, in fact, it rises much less difficulties than the latter. Here, two algorithms for order reduction of second level realizations were presented. The first one, proposed by C.Beck and J.Doyle, yields some remarkable results about the minimality of 2-D models and always ensures the stability of the obtained reduced order model. The second one, proposed in this thesis, sometimes leads to an unstable reduced order model. However, it is much faster than the first one: in the examples given here, computations show that a reduced order model can be computed up to ten times quicker with this alternative approach, comparatively to the Beck-Doyle algorithm. Furthermore, the alternative approach can be applied to some unstable 2-D systems, which is not the case for the Beck-Doyle's. This feature can be particularly useful in the context of model order reduction of uncertain systems, which is, in some sense, equivalent to the model order reduction of n-D systems (recall section 5.2). In terms of the produced error, it is not possible to say which of the algorithms is better: in some cases the Beck-Doyle algorithm produces a better approximation than the alternative approach (recall examples 2 and 3 in section 4.3.3) and in others the inverse happens (recall Example 1 in section 4.3.3 and section 5.1).

The order reduction of first level realizations is still an incomplete task. In this thesis, some efforts towards balancing a first level realization were presented. It was shown that, for a certain class of models, it is possible to transform a first level realization into another whose (truncated) controllability or observability gramian is the identity matrix, by applying an identified similarity transformation. As pointed out before, this is the first step towards finding a balanced realization. To complete this task, more information about the structure of inner matrices and/or about the class of matrices that have a PSVD seems to be needed. As a conclusion, it is important to point out that, besides the theoretical interest of this approach, being able to perform direct order reduction of first level realizations could be specially useful for model order reduction of uncertain systems, since it would not require pulling-out the uncertainty from the nominal model.

Appendix A

Useful algebraic structures

Ring

A *ring* is a set R equipped with two binary operators $+$ (addition) and $\cdot$ (multiplication) satisfying the following properties, for all $a, b, c \in R$:

1. $(a+b)+c = a+(b+c)$ ($+$ is associative)
2. There exists a $0 \in R$ such that $0+a = a+0 = a$ (existence of additive identity)
3. $a+b = b+a$ ($+$ is commutative)
4. For each a there exists a $-a \in R$ such that $-a+a = a+(-a) = 0$ (existence of additive inverse)
5. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ ($\cdot$ is associative)
6. $a \cdot (b+c) = a \cdot b + a \cdot c$ ($\cdot$ is left distributive over $+$)
7. $(b+c) \cdot a = b \cdot a + c \cdot a$ ($\cdot$ is right distributive over $+$)

Ideal

An *ideal* is a subset I of elements in a ring R satisfying the following properties, for all $i_1, i_2 \in I$ and $r \in R$:

1. $i_1 + i_2 \in I$ (closure under addition)
2. additive associativity
3. existence of additive identity
4. existence of additive inverse
5. $i_1 \cdot r \in I$ (closure under right multiplication by an element of the ring)

6. $r \cdot i_1 \in I$ (closure under left multiplication by an element of the ring)

Integral domain

An *integral domain* is a ring that is commutative under multiplication (commutative ring), has a multiplicative identity and has no divisors of 0 (i.e., for all $a, b \in R,\ a \cdot b = 0 \Longrightarrow a = 0 \vee b = 0$).

Principal ideal domain

A *principal ideal domain* is an integral domain in which every ideal is generated by one element.

Field

A *field* is a commutative ring S that, in addition, has the following properties:

1. There exists an element $1 \in S$ such that for all $a \in S,\ 1 \cdot a = a \cdot 1 = a$ (existence of multiplicative identity)
2. For each $a \neq 0 \in S$, there exists an element $a^{-1} \in S$ such that $a^{-1} \cdot a = a \cdot a^{-1} = 1$ (existence of multiplicative inverse)

Appendix B

Relative degree of rational functions – some algebraic properties

Consider rational functions $g_1(z) = p_1(z)/q_1(z)$ and $g_2(z) = p_2(z)/q_2(z)$, where $p_i(z)$ and $q_i(z)$ are polynomials with degree m_i and n_i, respectively. Let $d(g_i(z)) := n_i - m_i$. Then,

1. $d(g_1(z)g_2(z)) = d(g_1(z)) + d(g_2(z))$.

   Proof.
   $$d(g_1(z)g_2(z)) = (n_1+n_2) - (m_1+m_2) = (n_1-m_1) + (n_2-m_2). \tag{B.1}$$
   □

2. $d(g_1(z)+g_2(z)) = \min(d(g_1(z)), d(g_2(z)))$, if no leading monomial cancellations occur.

   Proof. If a monomial of $p_1(z)q_2(z)$ does not cancel the leading monomial of $p_2(z)q_1(z)$, and vice-versa, the degree of the numerator of $p_1(z)q_2(z)+p_2(z)q_1(z)$ is $\max(m_1+n_2, m_2+n_1)$. Thus,
   $$\begin{aligned} d(g_1(z)+g_2(z)) &= (n_1+n_2) - \max(m_1+n_2, m_2+n_1) = \\ &= (n_1+n_2) + \min(-(m_1+n_2), -(m_2+n_1)) = \\ &= \min((n_1+n_2)-(m_1+n_2), (n_1+n_2)-(m_2+n_1)) = \\ &= \min(n_1-m_1, n_2-m_2). \end{aligned} \tag{B.2}$$
   □

3. $d(g_1(z)^k) = kd(g_1(z))$, $k \in \mathbb{Z}$.

   Proof.
   $$d(g_1(z)^k) = kn_1 - km_1 = k(n_1-m_1). \tag{B.3}$$
   □

4. If, moreover, $p_1(0) \neq 0$ and $q_1(0) \neq 0$, then $d(g_1(z^{-1})) = 0$.

Proof. Let

$$g_1(z) = \frac{\alpha_{m_1} z^{m_1} + \alpha_{m_1-1} z^{m_1-1} + \cdots + \alpha_0}{z^{n_1} + \beta_{n_1-1} z^{n_1-1} + \cdots + \beta_0}, \tag{B.4}$$

where $\alpha_0 \neq 0$ and $\beta_0 \neq 0$, because $p_1(0) \neq 0$ and $q_1(0) \neq 0$.
Then,

$$\begin{aligned} g_1(z^{-1}) &= \frac{\alpha_{m_1} z^{-m_1} + \alpha_{m_1-1} z^{-(m_1-1)} + \cdots + \alpha_0}{z^{-n_1} + \beta_{n_1-1} z^{-(n_1-1)} + \cdots + \beta_0} = \\ &= \frac{\alpha_0 z^{n_1} + \cdots + \alpha_{m_1-1} z^{n_1-(m_1-1)} + \alpha_{m_1} z^{n_1-m_1}}{\beta_0 z^{n_1} + \cdots + \beta_{n_1-1} z + 1}. \end{aligned} \tag{B.5}$$

Thus, $d(g_1(z^{-1})) = n_1 - n_1 = 0$. □

Appendix C

MATLAB implementation of the algorithms

C.1 Beck-Doyle algorithm

```
function [ A11r,A12r,A21r,A22r,B1r,B2r,C1r,C2r,D,err ] = ...
          BeckDoyleModRed( A11,A12,A21,A22,B1,B2,C1,C2,D,hdisc,vdisc )
% Model order reduction of the 2-D model Sigma=(A11,A12,A21,A22,B1,B2,
% C1,C2,D) using the Beck-Doyle algorithm, described in section 4.3.1

% hdisc and vdisc contain the number of horizontal and vertical states
% to be discarded

% err gives the sum of the discarded singular values

A2d=[A11, A12; A21, A22];
B2d=[B1; B2];
C2d=[C1, C2];

nh=length(A11);
nv=length(A22);

%LMI for the controllability gramian
setlmis([]);
p = lmivar(1,[nh 1;nv 1]);
lmiterm([-1 1 1 p],1,1);
lmiterm([2 1 1 p],A2d,A2d');            % LMI : A*P*A'
lmiterm([2 1 1 p],1,-1);                % LMI : -P
lmiterm([2 1 1 0],B2d*B2d');            % LMI : B*B'
ctrgram=getlmis;
c=mat2dec(ctrgram,eye(nh+nv));
options=[1e-5,0,0,0,0];
```

```
[~,popt]=mincx(ctrgram,c,options);
P=dec2mat(ctrgram,popt,p);
P1=P(1:nh,1:nh);
P2=P(nh+1:nh+nv,nh+1:nh+nv);

%LMI for the observability gramian
setlmis([]);
q=lmivar(1,[nh 1;nv 1]);
lmiterm([-1 1 1 q],1,1);
lmiterm([2 1 1 q],A2d',A2d);            % LMI : A'*P*A
lmiterm([2 1 1 q],1,-1);                % LMI : -Q
lmiterm([2 1 1 0],C2d'*C2d);            % LMI : C'*C
obsgram=getlmis;
c=mat2dec(obsgram,eye(nh+nv));
options=[1e-5,0,0,0,0];
[~,qopt]=mincx(obsgram,c,options);
Q=dec2mat(obsgram,qopt,q);
Q1=Q(1:nh,1:nh);
Q2=Q(nh+1:nh+nv,nh+1:nh+nv);

%balancing the realization
[U,S,~]=svd(P1);
T1=U*sqrt(S);
[U,S,~]=svd(P2);
T2=U*sqrt(S);
T=[T1 zeros(nh,nv);zeros(nv,nh) T2];
Pp=T\P/T';
Qq=T'*Q*T;

Aa=T\A2d*T;
Bb=T\B2d;
Cc=C2d*T;

Qq1=Qq(1:nh,1:nh);
Qq2=Qq(nh+1:nh+nv,nh+1:nh+nv);

[U,~,~]=svd(Qq1);
T1=U;
[U,~,~]=svd(Qq2);
T2=U;
T=[T1 zeros(nh,nv); zeros(nv,nh) T2];
Ppp=T\Pp/T';
Qqq=T'*Qq*T;

Aaa=T\Aa*T;
Bbb=T\Bb;
Ccc=Cc*T;

Qqq1=Qqq(1:nh,1:nh);
```

```
Qqq2=Qqq(nh+1:nh+nv,nh+1:nh+nv);

[~,S,~]=svd(Qqq1);
T1=inv(S.^(1/4));
[~,S,~]=svd(Qqq2);
T2=inv(S.^(1/4));
T=[T1 zeros(nh,nv); zeros(nv,nh) T2];

Pbal=T\Ppp/T'
Qbal=T'*Qqq*T

A2dbal=T\Aaa*T;
B2dbal=T\Bbb;
C2dbal=Ccc*T;

%the 2-D realization (A2dbal,B2dbal,C2dbal,D) is a balanced one

%now, perform the reduction
A11r=A2dbal(1:nh-hdisc,1:nh-hdisc);
A12r=A2dbal(1:nh-hdisc,nh+1:nh+nv-vdisc);
A21r=A2dbal(nh+1:nh+nv-vdisc,1:nh-hdisc);
A22r=A2dbal(nh+1:nh+nv-vdisc,nh+1:nh+nv-vdisc);
B1r=B2dbal(1:nh-hdisc,:);
B2r=B2dbal(nh+1:nh+nv-vdisc,:);
C1r=C2dbal(:,1:nh-hdisc);
C2r=C2dbal(:,nh+1:nh+nv-vdisc);

err=0;
for i=nh-hdisc+1:nh
    err=err+Pbal(i,i);
end
for i=nh+nv-vdisc+1:nh+nv
    err=err+Pbal(i,i);
end

end
```

C.2 Alternative approach

```
function [ A11r,A12r,A21r,A22r,B1r,B2r,C1r,C2r,D,err ] = ...
          altModRed( A11,A12,A21,A22,B1, B2,C1,C2,D,hdisc,vdisc )
% Model order reduction of the 2-D model Sigma=(A11,A12,A21,A22,B1,B2,
% C1,C2,D) using the alternative approach described in section 4.3.2

% hdisc and vdisc contain the number of horizontal and vertical states
% to be discarded
```

```
nh=length(A11);
nv=length(A22);

[ny,nu]=size(D);

A11r=A11;
A12r=A12;
A21r=A21;
A22r=A22;
B1r=B1;
B2r=B2;
C1r=C1;
C2r=C2;
err=0;

if hdisc~=0

% representation of the system in the form (4.86)−(4.87)
Ah=A11;
Bh=[A12, B1];
Ch=[A21; C1];
Dh=[A22, B2; C2, D];

%calculating the gramians
Ph=dlyap(Ah,Bh*Bh');
Qh=dlyap(Ah',Ch'*Ch);

%balancing the realization
[U,S,~]=svd(Ph);
T=U*sqrt(S);
Pph=T\Ph/T';
Qqh=T'*Qh*T;

Aah=T\Ah*T;
Bbh=T\Bh;
Cch=Ch*T;

[U,~,~]=svd(Qqh);
T=U;
Ppph=T\Pph/T';
Qqqh=T'*Qqh*T;

Aaah=T\Aah*T;
Bbbh=T\Bbh;
Ccch=Cch*T;

[~,S,~]=svd(Qqqh);
T=inv(S.^(1/4));
Phbal=T\Ppph/T';
```

```
Qhbal=T'*Qqqh*T;

Abal=T\Aaah*T;
Bbal=T\Bbbh;
Cbal=Ccch*T;

%performing the order reduction
Ar=Abal(1:nh-hdisc,1:nh-hdisc);
Br=Bbal(1:nh-hdisc,:);
Cr=Cbal(:,1:nh-hdisc);
Dr=Dh;

for i=nh-hdisc+1:nh
    err=err+Phbal(i,i);
end

nh=nh-hdisc;

A11r=Ar;
A12r=Br(:,1:nv);
A21r=Cr(1:nv,:);
A22r=Dr(1:nv,1:nv);
B1r=Br(:,nv+1:nv+1+nu-1);
B2r=Dr(1:nv,nv+1:nv+1+nu-1);
C1r=Cr(nv+1:nv+1+ny-1,:);
C2r=Dr(nv+1:nv+1+ny-1,1:nv);

% the horizontal state of the system has now reduced order nh-hdisc

end

if vdisc~=0

% representation of the system in the form (4.88)-(4.89)
Av=A22r;
Bv=[A21r, B2r];
Cv=[A12r; C2r];
Dv=[A11r, B1r; C1r, D];

%calculating the gramians
Pv=dlyap(Av,Bv*Bv');
Qv=dlyap(Av',Cv'*Cv);

%balancing the realization
[U,S,~]=svd(Pv);
T=U*sqrt(S);
Ppv=T\Pv/T';
Qqv=T'*Qv*T;
```

```
Aav=T\Av*T;
Bbv=T\Bv;
Ccv=Cv*T;

[U,~,~]=svd(Qqv);
T=U;
Pppv=T\Ppv/T';
Qqqv=T'*Qqv*T;

Aaav=T\Aav*T;
Bbbv=T\Bbv;
Cccv=Ccv*T;

[~,S,~]=svd(Qqqv);
T=inv(S.^(1/4));
Pvbal=T\Pppv/T';
Qvbal=T'*Qqqv*T;

Abal=T\Aaav*T;
Bbal=T\Bbbv;
Cbal=Cccv*T;

Ar=Abal(1:nv-vdisc,1:nv-vdisc);
Br=Bbal(1:nv-vdisc,:);
Cr=Cbal(:,1:nv-vdisc);
Dr=Dv;

for i=nv-vdisc+1:nv
    err=err+Pvbal(i,i);
end

nv=nv-vdisc;

A11r=Dr(1:nh,1:nh);
A12r=Cr(1:nh,:);
A21r=Br(:,1:nh);
A22r=Ar;
B1r=Dr(1:nh,nh+1:nh+1+nu-1);
B2r=Br(:,nh+1:nh+1+nu-1);
C1r=Dr(nh+1:nh+1+ny-1,1:nh);
C2r=Cr(nh+1:nh+1+ny-1,:);

% the vertical state of the system has now reduced order nv-vdisc
end
% the model order reduction is now complete

end
```

References

[1] R. Eising. *2-D Systems, an algebraic approach*. Mathematisch Centrum, Amsterdam, The Netherlands, 1979.

[2] R.P. Roesser. A discrete state-space model for linear image processing. *IEEE Transactions on Automatic Control*, AC-20(1):1 – 10, 1975/02.

[3] E. Fornasini and G. Marchesini. Doubly-indexed dynamical systems: state-space models and structural properties. *Mathematical Systems Theory*, 12(1):59 – 72, 1978.

[4] E.D. Sontag. On first-order equations for multidimensional filters. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-26(5):480 – 2, 1978/10.

[5] Hong Luo, Wu-Sheng Lu, and A. Antoniou. A weighted balanced approximation for 2-D discrete systems and its application to model reduction. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 42(8):419 – 29, 1995/08.

[6] M. Diab, V. Sreeram, and W.Q. Liu. Model reduction of 2-D separable-denominator transfer functions via quasi-kalman decomposition. volume 3, pages 1680 – 1683, Hong Kong, 1997.

[7] P.N. Paraskevopoulos, P.E. Panagopoulos, G.K. Vaitsis, S.J. Varoufakis, and G.E. Antoniou. Model reduction of 2-D systems via orthogonal series. *Multidimensional Systems and Signal Processing*, 2(1):69 – 83, 1991/03.

[8] C.L. Beck, John Doyle, and Keith Glover. Model reduction of multidimensional and uncertain systems. *IEEE Transactions on Automatic Control*, 41(10):1466–1477, 1996.

[9] Chi-Tsong Chen. *Linear System Theory and Design*. Oxford University Press, Third edition, 1999.

[10] Tohru Katayama. *Subspace Methods for System Identification*. Springer, First edition, 2005.

[11] L. Hogben. *Handbook of Linear Algebra*. Discrete Mathematics and Its Applications. Taylor & Francis, 2007.

[12] H. Schumacher and J.C. Willems. *Lecture Notes of the course "Mathematical Systems Theory and Linear Systems"*. Dutch Institute for Systems and Control – DISC, The Netherlands, 1987.

[13] A.V. Oppenheim, A.S. Willsky, and I.T. Young. *Signals and systems*. Prentice-Hall signal processing series. Prentice-Hall, 1983.

[14] M. Vidyasagar and N. K. Bose. Input-output stability of linear systems defined over measure spaces. In *Proceedings of 1975 Midwest Symposium on Circuits and Systems*, pages 394–397, 1975.

[15] D. Goodman. Some stability properties of two-dimensional linear shift-invariant digital filters. *IEEE Transactions on Circuits and Systems*, CAS-24(4):201 – 8, 1977/04.

[16] Tong-Yi Guo, Chyi Hwang, Leang-San Shieh, and Chen-Hung Chen. Reduced-order models of 2-D linear discrete separable-denominator system using bilinear routh approximations. *IEEE Proceedings, Part G: Circuits, Devices and Systems*, 139(1):45 – 56, 1992.

[17] Qing Wang, James Lam, Huijun Gao, and Qingyang Wang. Energy-to-peak model reduction for 2-D discrete systems in Fornasini-Marchesini form. *European Journal of Control*, 12(4):420 – 430, 2006.

[18] P. Misra and T. Manickam. Balanced realization of separable-denominator multidimensional systems. *Linear Algebra and Its Applications*, 188-189:521 – 47, 1993/07.

[19] K. Zhou, J.L. Aravena, Guoxiang Gu, and Dapeng Xiong. 2-D model reduction by quasi-balanced truncation and singular perturbation. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 41(9):593 – 602, 1994/09.

[20] W.-S. Lu, E. B. Lee, and Q.-T. Zhang. Balanced approximation of two-dimensional and delay-differential systems. *International Journal of Control*, 46(6):2199–2218, 1987.

[21] Chengshan Xiao, V. Sreeram, W.Q. Liu, and A.N. Venetsanopoulos. Identification and model reduction of 2-D systems via the extended impulse response Gramians. *Automatica*, 34(1):93 – 101, 1998/01.

[22] Cuihong Wang and Lin Jia. H_∞ model reduction for positive 2-D discrete systems in Roesser model. pages 1733 – 8, Piscataway, NJ, USA, 2012.

[23] E.D. Sontag. The lattice of minimal realizations of response maps over rings. *Mathematical Systems Theory*, 11(2):169 – 75, 1977.

[24] Jana Nemcova and Jan H. van Schuppen. Realization theory for rational systems: Minimal rational realizations. *Acta Applicandae Mathematicae*, pages 1 – 22, 2009.

[25] A.V. Oppenheim, R.W. Schafer, and J.R. Buck. *Discrete-time signal processing*. Prentice-Hall signal processing series. Prentice Hall, 1999.

[26] Mi-Ching Tsai, Da-Wei Guf, Ian Postlethwaite, and Brian D. O. Anderson. Inner functions and a pseudo-singular-value decomposition in super-optimal h-infinity control. *International Journal of Control*, 51(5):1119–1131, 1990.

[27] C.L. Beck and John Doyle. A necessary and sufficient minimality condition for uncertain systems. *IEEE Transactions on Automatic Control*, 44(10):1802–1813, 1999.

[28] P. Bendotti and C.L. Beck. On the role of LFT model reduction methods in robust controller synthesis for a pressurized water reactor. *IEEE Transactions on Control Systems Technology*, 7(2):248–257, 1999.

[29] P. Gahinet, A. Nemirovski, A.J. Laub, and M. Chilali. *LMI Control Toolbox For Use With MATLAB*. The Mathworks, 1995.

[30] Ettore Fornasini. A 2-D systems approach to river pollution modelling. *Multidimensional Systems and Signal Processing*, 2(3):233–265, 1991.

[31] S. Weiland. *Lecture Notes of the course "Robust Control"*. Eindhoven University of Technology – The Netherlands, 2011.

[32] G.F.A. Franklin, J.D.A. Powell, and M.L.A. Workman. *Digital control of dynamic systems*. Addison-Wesley world student series. Addison Wesley Longman, 1998.