UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Data-driven control and interconnection

by

Thabiso M. Maupong

Thesis for the degree of Doctor of Philosophy

May 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES
Electronics and Computer Science

Doctor of Philosophy

DATA-DRIVEN CONTROL AND INTERCONNECTION

by Thabiso M. Maupong

In this thesis we study the design of controllers from data for a class of linear-time-invariant dynamical systems. We formulate and study conditions under which data from a system can be used to compute a controller which meets a given performance specification. Under such conditions we develop algorithms for finding a controller. Our methods do not focus only on finding control inputs trajectories but one can also find a representation of the controller under suitable conditions. We also give necessary and sufficient conditions for determining whether a system is dissipative using data.

Finally, we study the notion of difference flatness for linear systems. We define a flat system, using trajectories, as a system whose variables admits a partition such that a subset of the variables are unconstrained, and the other remaining variables are observable from the unconstrained variables. Using this definition we give a characterization of representations of flat systems. We also study the control of flat systems using the notion of control as interconnection.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Thabiso M. Maupong , declare that the thesis entitled *Data-driven control and interconnection* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: [44, 45, 47] and [46]

Signed:...

Date:...

Acknowledgements

I would like to thanks those people who have supported me during my PhD studies.

First and foremost, I would like to extend my greatest gratitude to Dr. Paolo Rapisarda for giving the opportunity to work with him. Paolo has been extremely helpful, patient and encouraging throughout my studies. Furthermore, he provided a conducive environment for learning and development of my skills in applied mathematics, writing mathematics, and presentations. I will always remember all the discussions we have had on the whiteboard.

I am also grateful to Dr. Chu Bing for the support, discussions during the control reading group, encouragement and useful comments he suggested during my 9 months examination. Professor Eric Rodgers for the suggestions during my transfer viva and for agreeing to be an examiner in my final viva. Professor Harry Trentelman for accepting to examine this thesis.

I would also like to thank Dr. Jonathan Mayo-Maldonado for not only been a good friend but for all the discussion we have had and introducing me to power electronics. For his hospitality and valuable advice during my academic visit to Tecnologico de Monterrey.

I am also grateful to all those who helped me maintain a good social/academic life during my stay in Southampton. In particular all my friend and colleagues (including all VLCCC members) in Southampton and back home in Botswana.

To my mother Kegope Maupong, daughter Rafiwa Pôpô Assan, my brother Kabelo Maupong and my extended family as a whole thanks for the unconditional love, encouragement, comfort and for the moral support.

Finally, I would like to thank the Department of Tertiary Education Financing (DTEF) in Botswana and the Botswana International University of Science and Technology (BIUST) for the scholarship.

Notation

$I_{\mathtt{n}}$	Identity matrix of dimensions $\mathtt{n} \times \mathtt{n}$.
$0_{\mathtt{n}}$	Zero matrix of dimensions $\mathtt{n} \times \mathtt{n}$.
$A^{\top}$	Transpose of the matrix A.
$\det(A)$	Determinant of a square matrix A.
$\mathrm{tr}(A)$	The trace of the matrix A.
$\mathrm{rank}(R)$	Rank of the matrix R.
$A \subseteq B$	Set A contained in, and possibly equal to, the set B.
$\mathbb{C}$	Field of complex numbers.
$\mathbb{R}$	Field of real numbers.
$\mathbb{Z}$	Ring of integers.
$\mathbb{W}^{\mathbb{T}}$	Set of all maps from $\mathbb{T}$ to $\mathbb{W}$.
$\mathbb{R}^{\mathtt{w}}$	Space of real vectors with $\mathtt{w}$ dimension.
$\mathbb{R}^{\bullet}$	Space of real vectors with unspecified but finite dimension.
$\mathbb{R}^{\mathtt{m}\times\mathtt{n}}$	Space of $\mathtt{m} \times \mathtt{n}$ dimensional real matrices.
$\mathbb{R}^{\bullet\times\mathtt{n}}$	Space of real matrices with $\mathtt{n}$ columns and a finite unspecified number of rows.
$\mathbb{R}[\xi]$	Ring of polynomials with real coefficients in the indeterminate ξ.
$\mathbb{R}[\xi^{-1}, \xi]$	Ring of Laurent polynomials with real coefficients in the indeterminate ξ.
$\mathbb{R}[\zeta, \eta]$	Ring of polynomials with real coefficients in the indeterminates ζ and η.
$\mathbb{R}(\xi)$	Field of rational functions in the indeterminate ξ.
σ	Time left shift operator.
$\deg(r)$	Degree of the polynomial r.
$\mathbb{L}^{\mathtt{w}}$	Set of all maps from $\mathbb{Z}$ to $\mathbb{R}^{\mathtt{w}}$ equipped with topology of pointwise convergence.
$\mathscr{L}^{\mathtt{w}}$	Class of linear shift invariant subspaces of $\mathbb{L}^{\mathtt{w}}$.
$\mathscr{L}^{\mathtt{w}}_{contr}$	Class of controllable linear shift-invariant subspaces of $\mathbb{L}^{\mathtt{w}}$.
$l_2^{\mathtt{w}}$	Set of square summable elements of $(\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$, i.e. $w \in l_2^{\mathtt{w}}$ implies that $\sum\limits_{t=-\infty}^{\infty} \|w(t)\|^2 < \infty$ where $\|w\|^2 := w^{\top} w$.

$\mathfrak{B}$	a behaviour.
$\mathtt{w}(\mathfrak{B})$	The number of components of an element w in $\mathfrak{B}$.
$\mathtt{m}(\mathfrak{B})$	The input cardinality of $\mathfrak{B}$.
$\mathtt{p}(\mathfrak{B})$	The output cardinality of $\mathfrak{B}$.
$\mathtt{n}(\mathfrak{B})$	The McMillan degree of $\mathfrak{B}$.
$\mathtt{L}(\mathfrak{B})$	The Lag of $\mathfrak{B}$.
$\mathtt{l}(\mathfrak{B})$	The shortest lag of $\mathfrak{B}$.
$\Pi_{w'}(\mathfrak{B})$	The projection of $\mathfrak{B}$ on the subset w' of its variables
im	Image of a linear map.
$\mathrm{col}(A, B)$	The matrix obtained by stacking the matrix A over B, if A and B have the same number of columns.
$\mathrm{row}(A, B)$	The matrix obtained by placing the matrix A next to B, if A and B have the same number of rows.
ker	Kernel of a linear map.
col span(A)	Column span of A.
leftkernel(A)	Left kernel of matrix A.
row span(A)	Row span of A.

Chapter 1

Introduction

In order to effectively design a controller for a dynamical to-be-controlled system one needs a good understanding of how the to-be-controlled system variables evolve over time, i.e. the system dynamics. In the so called *model-based control* this understanding is given explicitly in the form of *mathematical models*. Examples of such mathematical models include first-order *state space models*, *transfer functions* and high-order *differential/difference* equations. The models are normally derived from first principles, e.g. *Newton's laws of motion* for mechanical systems, *Ohm's* and *Kirchhoff's circuit* laws for electrical circuits. After finding equations from first principles, in the case of state space models, one has to do some mathematical manipulations to, first, introduce an artificial variable called the *state* and, second, to describe the to-be-controlled systems using first-order state space equations. This is necessary because state variables do not appear naturally when modelling, moreover, it is not always the case that one gets first-order equations when modelling dynamical systems. Mathematical manipulations are also needed to obtain rational functions from first principles equations in the case of transfer functions.

Remark 1.1. It is worth noting that despite its considerable popularity and application in control, describing to-be-controlled systems using state space models and transfer functions has got limitations and shortcomings. This fact is well articulated with explanatary examples on pp. 1-3 of [53] for the general case and pp. 1-9 of [48] for *switched systems*. One of the shortcomings of state space models is that, when modelling one naturally gets high-order differential or difference equations while state space models consists of first-order equations. Both state space models and transfer functions are also limited in the sense that there is the need to identify system variables as either inputs or outputs but in reality there are situations in which systems variables cannot be easily classified as inputs or outputs. For example, take Ohm's law which states that the voltage (V) across a resistor (R) is directly propotional to the currunt (I) with the resistance as the proportionality factor. However, the law does not specify which of the current and voltage is the input and which the output. Another problem comes from the fact that state space models and transfer functions do not take into consideration that an effective way to device mathematical models is to break down a system into subsystems, then obtain subsystems models and interconnect them. This is normally referred to as *modelling by tearing, zooming and linking*, see [9, 48, 85], where

- Tearing: Viewing the to-be-controlled system as an interconnection of subsystems.
- Zooming: Modelling of the subsystems.
- Linking: Modelling of the interconnections among the subsystems.

□

Model-based control has had a considerable amount of success over the years, this is evident from the development and use of several model-based techniques such as *robust control, linear quadratic control (LQG), model-predictive control* and *optimal control.* Moreover, the study of dynamical systems properties such as *dissipativity* is also based on models. Such has been the success of model-based control that in the case where only data from the to-be-controlled system is available, then the control problem is split into two stages as follows.

1. Model identification;
2. Controller design.

We shall call this the *two stage approach to controller design.* Step number (1) above is referred to as *system identification* in the literature, see [37, 41, 67, 76]. This involves using data from the to-be-controlled system in conjunction with various mathematical techniques to infer a mathematical model that best describes the to-be-controlled system. Once a model is obtained then a model-based control technique is used to find a suitable controller for the to-be-controlled system.

In recent years, a considerable amount of research effort has gone into the development of control design strategies which avoid the two stage approach to controller design above. Most of these methods aim to use data from the to-be-controlled system to directly design a controller, hence, avoiding the identification stage. In the literature, this is referred to as *data-driven control.* Data-driven control, covered in more details in the next section, is not only suitable for avoiding the two stage approach to controller design, but also offers an alternative to model-based control in various situations including, but not limited to, the following.

- *To-be-controlled system models are not available.* Such models may not be available for various reasons for example: the system could be difficult or expensive to model;

some aspect of the system may be difficult to quantify using models; there is a time constraint in the design of the controller, i.e. there is not enough time to come up with a suitable model and verify such a model.

- *Model shortcomings.* Models can have some shortcomings when applied in practice. These shortcoming could be as a result of excessive assumptions made during the modelling stage or due over simplification of the model in order to make analysis easy to follow, see p. 15 of [11].

Data-driven control approaches have also gained more popularity in application. This is evident from the recent call for a special issue on data-driven control by *IEEE Transaction on Industrial Electronics* and recent publications in areas such as: real-time, fault-tolerant controller design for electrical circuits [24]; on-line adaptive control [62, 92]; fault tolerant control design [64]; and filtering [38].

1.1 Data-driven control

Data-driven control, also known as *model-free* control or *data-based* control, is not completely new, see for example papers from the 1990's [7, 60], in fact it can be argued that one of the oldest data-driven approaches to control is the Ziegler-Nichols method for tuning PID parameters. The Ziegler-Nichols PID tuning approach, is a graphical method where the PID paramters are determined by first connecting the system with a propotional controller then increasing the propotional gain until the system output reaches consistent oscilliations. Then the oscillition period and propotional gain are used to set the value of the gains of the PID. In the literature, there are several definitions of data-driven control, for example see p. 4 of [22], we state one below.

> *Data-driven control includes all control theories and methods in which a controller is designed by directly using on-line or off-line data of the to-be-controlled plant or knowledge from the data processing but not any explicit information from mathematical models of the to-be-controlled plant, and whose stability, convergence and robustness can be guaranteed by rigorous mathematical analysis under certain reasonable assumptions.*

It follows from the above definition that in data-driven control, the to-be-controlled system models are not used in the design of the controller. Moreover, data is not used to infer any models that describe the system. Rather, data-driven control is a procedure that derives a control law or a controller for a to-be-controlled system directly from data and the given performance specification.

1.1.1 A review of data-driven control literature

We briefly review some of the data-driven control methods for a class of *linear time-invariant* systems, see [22] for an exhaustive review. First, in Table 1.1 we give a summary of some of the characteristics of data-driven control methods.

Features	**Description**
Off-line	Data from the to-be-controlled system is stored and the design of a controller is off-line.
On-line	A window of real-time data from the system is used on-line. Compared to off-line, on-line approaches contain the latest information about the system.
Iterative	A technique is referred to as iterative if the design of a controller involves an algorithm which is recursive.
Finite-horizon	A technique is considered finite-horizon if the designed controller action is defined over a finite time, i.e. if the performance criterion is specified over a finite time t.

Table 1.1: Outline of data-driven control features

We now discuss a summary of some data-driven control algorithms and their application, available in the literature. The first methods are *subspace* based, that is, the solutions use numerical tools like: QR-decompositions, singular value decomposition (SVD), linear operations such as projection and a solution of systems of equations.

- *Output matching control* [39, 40]. The authors present an algorithm for finding an input trajectory such that the system output trajectory matches a given reference output trajectory for some initial conditions. The solution is based on the *behavioural system theory* [81] and on the assumption that the data is from a controllable system and the input trajectory is *persistently exciting* (see p. 412 of [37]) of some order.
- *Linear quadratic tracking* [39, 40, 54]. An algorithm for finding a trajectory that minimises a quadratic cost function with respect to a given reference trajectory is presented. The algorithm also takes into consideration the initial conditions and is based on the same assumptions as the output matching control solution above.
- *Tracking control based on input-output* [23]. The authors illustrates, with the help of an example, an algorithm which uses input-output data from the to-be-controlled system to find a control input such that the system tracks a given reference trajectory. In their algorithm the input-output data is used to represent the dynamics of the system. The data is used to construct a matrix which the authors call the *array matrix*. The dimensions of the matrix depend on the *order* and *relative degree* of the system. The solution is based on the assumptions that the order and relative degree of the system are known a priori, that the given

data is noise free and comes from a *minimum phase* system. In common with the previous two methods the input trajectories are required to be persistently exciting of some order.

- *Markov data-based LQG control* [65]. Data from the to-be-controlled system is used to find *Markov parameters* [8]. Then using the Markov parameters, the authors show how to find an optimal control input sequence that minimises a given quadratic function.

- *Model-free H_∞ control* [95, 93]. The authors present an algorithm for solving the H_∞ control design problem using input-output data. In the first steps of the algorithm, the data is used to find a *subspace predictor* of the to-be-controlled system. The predictor is then used to find an optimal control input which minimizes the H_∞ norm. The solution is based on the assumption that the data has persistently exciting inputs.

In the following we review a data-driven adaptive control method.

- *Unfalsified control* [60, 61, 62, 94]. Unfalsified control is an approach whereby input/output data is used to falsify a control law from a set of available admissible control laws for a given system. A controller is said to be unfalsified by data if when such controller and the system are connected in a closed loop, then the system achieves the required specification; otherwise, it is falsified by data. The process of falsifying control laws based on data is carried out until an unfalsified control law is found.

We summaries the techniques covered in Table 1.2 on p. 6.

Algorithm and Technique	**Outcomes**	**Requirements**	**Features**
Output matching control algorithm and Linear quadratic tracking control algorithm	control inputs	1. Open loop experimental data with persistently exciting inputs 2. Initial condition as prefix trajectory 3. Reference trajectory	- Off-line - Non-iterative - Finite-horizon
A model-less algorithm for tracking control based on input-output data	control inputs	1. Open loop experimental data with persistently exciting inputs 2. Order of the system 3. Relative degree of the system 4. Reference output	- Off-line - Non-iterative - Finite-horizon
Unfalsified control	controller representation	1.Either open or closed loop experimental data 2. Set of admissible control laws 3. A set of performance specifications	- Applicable both off and on-line
Markov data-based LQG control	optimal control input	1.Markov parameters 2. Performance criterion as quadratic cost function	- Off-line - Iterative - Finite-horizon
Model-free H_∞ control	future optimal control input	1. Experimental data with persistently exciting inputs	- Off-line - Iterative - Finite-horizon

Table 1.2: Summary and comparison of data-driven control techniques

1.2 Goals and contributions of this thesis

We outline the contributions of this thesis.

Note that the goal of the data-driven approaches discussed in the previous section is to find a control sequence of some sort, except for unfalsified control where the aim is to use the system data to identify a suitable controller at a particular moment. Moreover, in all of the the approaches outlined, there is the need to partition the to-be-controlled system data into input and output trajectories. Hence, these approaches are likely to have some of the limitations outlined in Remark 1.1, particularly when there is no clear separation of the to-be-controlled system variables into inputs and outputs, in which case the approaches will not be applicable.

In this thesis, we aim to achieve the following.

1. Study and develop mathematically sound conditions such that given data from the to-be-controlled system and some performance specification we are able to compute a controller, i.e. to find some representation of the controller,
2. develop step-by-step algorithms to compute a controller on the basis of the obtained conditions.

In order to achieve (1)-(2) above we use the behavioural approach to dynamical systems as in [39, 40]. However, rather than partitioning the to-be-controlled system data into input and output trajectories we use the *interconnection paradigm* (see [50]), therefore, our approach is based on control as an interconnection between two systems, i.e. the interconnection of the controller and the to-be-controlled system. We develop data-driven solutions using the two types of interconnections: the *partial interconnection*, when there is separation of system variables into *control* and *to-be-controlled* variables (see [2]), and *full interconnection*, when the control variables and to-be-controlled variables coincide (see [50]). Our solutions are characterised as follows.

i. Since we use the notion of control as interconnection there is no need to partition system variables into inputs and outputs. Therefore, our results could be applicable when there is no clear separation of system variables into input and outputs. Consequently, our solution avoids one of the shortcomings of state space models and transfer functions outlined in Remark 1.1.

ii. One can find high-order equations that best describes the controller.

iii. The algorithms developed are off-line and non-iterative.

Based on the notion of control as interconnection some of the contributions of this thesis include the following.

3. Throughout the thesis we present some results on *parametrization* of controllers. Parametrization of *stabilizing controllers* in feedback control has been studied in [96] and in the behavioural framework in [33] for full interconnection and [51] for partial interconnection. In this thesis, among other parametrization, we present a parametrization of all controllable controllers and also parametrization of controlled behaviours *implemented* via full interconnections.

4. We also study *flat systems* (see [14]) in a behavioural system theory setting. We present a new definition of *flatness*, then we give a characterization of flat systems based on the definition. We also study control of flat systems using the interconnection paradigm.

1.3 Outline of the thesis

The outline of this thesis is as follows:

- Chapter 2. We cover some of the concepts of the behavioural system theory that are necessary to understand the results presented in this thesis. This includes a formal definition of linear time-invariant discrete dynamical systems, notions of controllability, observability, identifiability and persistency of excitation.

- Chapter 3. In this chapter, we review the literature on control as interconnection. We study the two types of interconnections, i.e. full and partial interconnection. We also cover the notion of implementability, that is, necessary and sufficient conditions for the existence of a controller that can be interconnected with a system to achieve some desired controlled behaviour.

The following chapters contain the main contributions of this thesis.

- Chapter 4. In this chapter, we present the first data-driven control problem. The problem is a hybrid, since we use both data and system representations to find a solution to a prescribed path problem. We show how to compute trajectories of the desired controlled systems using representations and to-be-controlled system trajectories. We prove sufficient conditions such that the computed trajectory can be used to find a controller such that the variables of the to-be-controlled system follow a prescribed trajectory for a given time interval.

- Chapter 5. We provide solutions to two completely data-driven control problems. In the first problem, given trajectories from of the to-be-controlled system and desired controlled system we show how to find a controller that when partially interconnected with the to-be-controlled through the control variables implements the given desired controlled behaviour. To find a solution we state necessary and sufficient conditions for finding a control variable trajectory from to-be-controlled variable trajectory. The second is the full interconnection case. Again we consider given trajectories from of the to-be-controlled system and desired controlled system and show how to find a controller that implements the desired controlled system via full interconnection. Both solutions are summarised by a step-by-step algorithm.

- Chapter 6. We show how to determine if a system is half-line dissipative using data. To do this we first cover some aspects of *quadratic difference forms* (QdFs) that are necessary to study theory of dissipativity from a behavioural framework point of view.

- Chapter 7. We study flat systems, focusing mainly on linear flat systems. We propose a trajector-based definition of flatness, and provide a characterization of flat systems. We also characterize and parametrize controllers interconnected with flat systems. We end the chapter by showing the parametrizations of the image representations of controlled systems.

- The last chapter of this thesis is Chapter 8 where we provide some general conclusions and possible future research.

Appendix A contains definitions of some of the important terms used throughout this thesis. It also contains some results on polynomial matrices.

Chapter 2

Behavioural system theory

We cover some concepts of the behavioural approach necessary for this thesis, focusing mainly on linear discrete-time dynamical systems. The material in this chapter is taken from various sources all of which are referenced within, and some of the terms used are explained in Appendix A.

2.1 Linear difference systems

Definition 2.1. A *dynamical system* Σ is defined by $\Sigma := (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, where $\mathbb{T}$ is called time axis, $\mathbb{W}$ the signal space and $\mathfrak{B}$ a subset of $\mathbb{W}^{\mathbb{T}}$ called the *behaviour* of the system.

The set $\mathbb{T}$ takes its values from $\mathbb{R}$ for continuous-time systems, and $\mathbb{Z}$ for discrete time systems.

Definition 2.2. $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is *linear* if

1. $\mathbb{W}$ is a vector space over $\mathbb{R}$, and
2. $\mathfrak{B}$ is a linear subspace of $\mathbb{W}^{\mathbb{T}}$.

The second statement in Definition 2.2 above is equivalent with Σ obeying the principle of superposition, i.e.

$$[w_1, w_2 \in \mathfrak{B} : \alpha, \beta \in \mathbb{R}] \Rightarrow [\alpha w_1 + \beta w_2 \in \mathfrak{B}] .$$

Let the *backward shift* operator

$$\sigma : \mathbb{W}^{\mathbb{T}} \to \mathbb{W}^{\mathbb{T}}$$

be defined by

$$(\sigma f)(t) := f(t+1).$$

Definition 2.3. $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is *time-invariant* if $\sigma^t\mathfrak{B} \subseteq \mathfrak{B}$ for all $t \in \mathbb{T}$.

The property of time-invariance implies that if a trajectory w belongs to a behaviour then the shifted trajectory also belongs to that behaviour as well, i.e.

$$[w(\cdot) \in \mathfrak{B}] \Longrightarrow [w(\cdot + t) \in \mathfrak{B} \; \forall \, t \in \mathbb{T}] .$$

Let $t_0, t_1 \in \mathbb{Z}$. We define the restriction of $\mathfrak{B}$ on the interval $[t_0, t_1]$ by

$$\mathfrak{B}_{|[t_0,t_1]} := \left\{ w : [t_0, t_1] \to \mathbb{W} \mid \text{there exists } w' \in \mathfrak{B} \text{ such that } w(t) = w'(t) \text{ for } t_0 \leqslant t \leqslant t_1 \right\} .$$

Definition 2.4. A dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is called *complete* if

$$[w \in \mathfrak{B}] \Longleftrightarrow \left[w_{|[t_0,t_1]} \in \mathfrak{B}_{|[t_0,t_1]} \text{ for all } -\infty < t_0 \leqslant t_1 < \infty \right]$$

Let $w \in \mathfrak{B}$, completeness implies that one can verify that $w \in \mathfrak{B}$ by verifying that $w_{|[t_0,t_1]} \in \mathfrak{B}_{|[t_0,t_1]}$ for all $t_0, t_1 \in \mathbb{Z}$. Σ is called $L-$*complete* if there exists $L \in \mathbb{Z}_+$ such that

$$[w \in \mathfrak{B}] \Longleftrightarrow \left[w_{|[t,t+L]} \in \mathfrak{B}_{|[t,t+L]} \text{ for all } t \in \mathbb{T} \right] . \tag{2.1}$$

The integer L appearing in (2.1) is called the *lag* and we shall discuss it in more details in section 2.6 along with other *integer invariants* of interest.

In this thesis, we consider linear time-invariant complete dynamical systems with the vector of *manifest variables* w, the signal space $\mathbb{R}^{\mathtt{w}}$ and the time axis $\mathbb{Z}$. Such systems are called *linear difference systems.* The behaviour $\mathfrak{B}$ of a linear difference systems consists of all trajectories $w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}}$ satisfying a system of linear difference equations with constant coefficients

$$R_L w(t + L) + R_{L-1} w(t + L - 1) + R_{L-2} w(t + L - 2) + \cdots + R_0 w(t) = 0 \tag{2.2}$$

where $R_L, R_{L-1}, R_{L-2}, \ldots, R_0 \in \mathbb{R}^{\mathtt{g} \times \mathtt{w}}$. Define from (2.2) the *polynomial matrix* $R \in \mathbb{R}^{\mathtt{g} \times \mathtt{w}}[\xi]$ by

$$R(\xi) := R_L \xi^L + R_{L-1} \xi^{L-1} + R_{L-2} \xi^{L-2} + \cdots + R_0 .$$

Then the system of linear difference equations in (2.2) can be written compactly as

$$R(\sigma) w = 0. \tag{2.3}$$

Equation (2.3) describes a linear difference system whose behaviour is defined by

$$\mathfrak{B} := \{ w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid R(\sigma) w = 0 \} ,$$

and (2.3) is called a *kernel representation*, thus we write $\mathfrak{B} = \ker(R(\sigma))$. The operator

$$R(\sigma) : (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}} \longrightarrow (\mathbb{R}^{\mathtt{g}})^{\mathbb{Z}}$$

is defined by

$$(R(\sigma)f)(t) := R_L f(t+L) + R_{L-1} f(t+L-1) + R_{L-2} f(t+L-2) + \cdots + R_0 f(t),$$

and is called *polynomial operator in the shift*. The following result holds true, see Definition A.2 in Appendix A for definition of a *unimodular matrix*.

Proposition 2.5. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ and $R(\sigma) : (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}} \longrightarrow (\mathbb{R}^{\mathtt{g}})^{\mathbb{Z}}$ be a polynomial operator in the shift. Then $R(\sigma)$ is:*

i. injective if and only if $R(\lambda)$ is full column rank for all $\lambda \in \mathbb{C}$,

ii. surjective if and only if $\operatorname{rank}(R) = \mathtt{g}$, and

iii. bijective if and only if $\mathtt{w} = \mathtt{g}$ and R is a unimodular matrix.

Proof. See section 4.1.3 p. 231 of [81]. □

Let $\mathscr{L}^{\mathtt{w}}$ denote the collection of all linear, closed and shift invariant subspaces of $(\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ equipped with the topology of pointwise convergence. In the following result we give a characterization of linear discrete time-invariant complete systems.

Theorem 2.6. *Let $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$, then the following statements are equivalent:*

i. Σ is time-invariant and complete,

ii. there exists $R \in \mathbb{R}^{\bullet\times\mathtt{w}}[\xi]$ such that $\mathfrak{B} = \ker(R(\sigma))$, and

iii. $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$.

Proof. Follows from Proposition 4.1A p. 232 of [81]. □

It has been shown that $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ such that $\mathfrak{B} = \ker(R(\sigma))$ is not unique, this point shall be covered in more details in section 2.5. We define the notion of a minimal kernel representation as follows.

Definition 2.7. Let $\mathfrak{B} = \ker(R(\sigma))$ where $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$. R induces a *minimal representation* of $\mathfrak{B}$ if

$$\left[R' \in \mathbb{R}^{\mathtt{g}'\times\mathtt{w}}[\xi]\right] \text{ and } \left[\mathfrak{B} = \ker(R'(\sigma))\right] \Longrightarrow \left[\mathtt{g}' \geq \mathtt{g}\right].$$

The following result holds true about minimal kernel representations.

Theorem 2.8. *Let* $\mathfrak{B} = \ker(R(\sigma))$ *with* $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$. *Then* R *induces a minimal representation if and only if* R *is of full row rank.*

Proof. See Proposition III.3 p. 263 of [82]. □

2.2 Linear difference systems with latent variables

When modelling dynamical systems, along with the variables of interest, manifest variables w, sometimes there is need to introduce some auxiliary variables called *latent* (or *auxiliary*) *variables* and denoted by ℓ. We define a dynamical system with latent variables as follows.

Definition 2.9. A *dynamical system with latent variables* is defined by a quadruple $\Sigma_L := (\mathbb{T}, \mathbb{W}, \mathbb{A}, \mathfrak{B}_{full})$ with $\mathbb{T}$ the time axis, $\mathbb{W}$ the manifest variable signal space, $\mathbb{A}$ the latent variable signal space and $\mathfrak{B}_{full} \subseteq (\mathbb{W} \times \mathbb{A})^{\mathbb{T}}$.

$\mathfrak{B}_{full}$ is called the *full behaviour*, and consists of all trajectories (w, ℓ) with w a manifest variable trajectory and ℓ a latent variable trajectory. Let $\mathfrak{B}_{full} \in \mathscr{L}^{\mathtt{w}+1}$, then $\mathfrak{B}_{full}$ consist of all trajectories (w, ℓ) satisfying a system of linear difference equations with constant coefficients

$$R_0 w(t) + R_1 w(t+1) + \cdots + R_L w(t+L) = M_0 \ell(t) + M_1 \ell(t+1) + \cdots + M_{L'} \ell(t+L') \quad (2.4)$$

where $R_i \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}$ for $i = 0, 1 \ldots L$ and $M_j \in \mathbb{R}^{\mathtt{g}\times 1}$for $j = 0, 1 \ldots L'$. Equation (2.4) is called a *hybrid representation* and is compactly written as

$$R(\sigma) w = M(\sigma) \ell,$$

where $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ and $M \in \mathbb{R}^{\mathtt{g}\times 1}[\xi]$ are defined in the obvious way. $\mathfrak{B}_{full}$ induces the *manifest behaviour* $\mathfrak{B}$, as we define now.

Definition 2.10. Let $\Sigma_{\mathrm{L}} = (\mathbb{T}, \mathbb{W}, \mathbb{A}, \mathfrak{B}_{full})$ be a dynamical system with latent variables. Then the *manifest dynamical system* induced by Σ_{L} is defined by $\Sigma := (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ with the *manifest behaviour* $\mathfrak{B}$ defined by

$$\mathfrak{B} := \{w : \mathbb{T} \to \mathbb{W} \mid \exists\ \ell \in \mathbb{A}^{\mathbb{T}} \text{such that}\,(w, \ell) \in \mathfrak{B}_{full}\}.$$

The process of obtaining the manifest behaviour from the full behaviour is called *elimination of latent variables* and is achieved by projecting $\mathfrak{B}_{full}$ onto the manifest variables.

Let the operator

$$\Pi_w : (\mathbb{W} \times \mathbb{A})^{\mathbb{T}} \to \mathbb{W}^{\mathbb{T}},$$

defined by

$$\Pi_w(w, \ell) := w$$

be the *projection onto the w variables* of $\mathfrak{B}_{full}$. Then manifest behaviour is defined by

$$\mathfrak{B} := \Pi_w(\mathfrak{B}_{full}),$$

where $\Pi_w(\mathfrak{B}_{full})$ is the projection of all $(w, \ell) \in \mathfrak{B}_{full}$ onto the w variables of $\mathfrak{B}_{full}$.

Theorem 2.11. *Let $\Sigma_L = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathbb{R}^{\mathtt{l}}, \mathfrak{B}_{full})$ be linear time-invariant and complete, then the manifest dynamical system $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$ is linear time-invariant and complete.*

Proof. See Theorem IV.3 p. 265 of [82]. □

In terms of representations the process of elimination of latent variable can be carried out by unimodular matrix multiplication as follows.

Theorem 2.12. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ and $M \in \mathbb{R}^{\mathtt{g}\times\mathtt{l}}[\xi]$ be such that*

$$\mathfrak{B}_{full} = \left\{(w, \ell) : \mathbb{Z} \to (\mathbb{R}^{\mathtt{w}} \times \mathbb{R}^{\mathtt{l}}) \mid R(\sigma)w = M(\sigma)\ell\right\}$$

and let $U \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ be a unimodular matrix such that

$$UM = \begin{bmatrix} 0 \\ M_1 \end{bmatrix}$$

with $M_1 \in \mathbb{R}^{\mathtt{g}'\times\mathtt{l}}[\xi]$ having full row rank. Define

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} := UR$$

where $R_2 \in \mathbb{R}^{\mathtt{g}'\times\mathtt{w}}[\xi]$. Then the manifest behaviour $\mathfrak{B}$ is described by

$$\mathfrak{B} = \{w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid R_1(\sigma)w = 0\}\,.$$

Proof. See Theorem 6.2.6 pp. 206-207 of [50]. □

2.3 Observability

We introduce the notion of observability.

Definition 2.13. Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$, with partition of the variables of $\mathfrak{B}$ defined by $w =: (w_1, w_2)$, where $w_k \in (\mathbb{R}^{\mathtt{w}_k})^{\mathbb{Z}}$ for $k = 1, 2$ such that $\mathtt{w} = \mathtt{w}_1 + \mathtt{w}_2$. Then w_2 is *observable* from w_1 if

$$\left[(w_1, w_2), (w_1, w_2') \in \mathfrak{B}\right] \Longrightarrow \left[w_2 = w_2'\right].$$

w_1 is called the *observed* variable and w_2 the *deduced* variable. Definition 2.13 is equivalent to the existence of a polynomial operator in the shift

$$F(\sigma) : (\mathbb{R}^{\mathtt{w}_1})^{\mathbb{Z}} \to (\mathbb{R}^{\mathtt{w}_2})^{\mathbb{Z}},$$

such that

$$w_2 = F(\sigma) w_1,$$

see p.268 of [82].

Observability is characterized in terms of representations as follows.

Theorem 2.14. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ be described by $R_1(\sigma)w_1 = R_2(\sigma)w_2$ with $R_1 \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}_1}[\xi]$ and $R_2 \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}_2}[\xi]$, then w_2 is observable from w_1 if and only if* $\operatorname{rank}(R_2(\lambda)) = \mathtt{w}_2$ *for all $\lambda \in \mathbb{C}$.*

Proof. See Theorem VI.2 p. 268 of [82]. □

2.4 Controllability and autonomous systems

Definition 2.15. $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ is *controllable* if for any two trajectories $w_1, w_2 \in \mathfrak{B}$ there exists a $t_1 \in \mathbb{Z}$, $t_1 \geq 0$ and $w \in \mathfrak{B}$ such that

$$w(t) = \begin{cases} w_1(t) & t \leq 0 \\ w_2(t - t_1) & t \geq t_1. \end{cases}$$

We shall denote by $\mathscr{L}^{\mathtt{w}}_{contr}$ the collection of all controllable elements of $\mathscr{L}^{\mathtt{w}}$. The following result holds true.

Theorem 2.16. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$. Then the following statements are equivalent:*

1. $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$,

2. *there exists an integer* $\mathtt{l}$ *and a matrix* $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ *such that*

$$\mathfrak{B} = \left\{ w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid \exists\ \ell \in (\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}} \text{ such that } w = M(\sigma)\ell \right\},$$

3. *if* $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ *is such that* $\mathfrak{B} = \ker(R(\sigma))$, *then* $\operatorname{rank}(R(\lambda))$ *is the same for all* $\lambda \in \mathbb{C}$.

Proof. (1) $\Longleftrightarrow$ (2) See Theorem 6.6.1 p. 229 of [50].
(1) $\Longleftrightarrow$ (3) See Theorem V.2 p. 266 of [82]. □

From statement (2) in Theorem 2.16, $\mathfrak{B} = \operatorname{im}(M(\sigma))$ is called an *image representation.* Note that in this case, the latent variable ℓ remains unconstrained, i.e. it can be any value in $(\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}}$.

Definition 2.17. Let $\mathfrak{B} = \operatorname{im}(M(\sigma))$ where $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$, then M induces an *observable image representation* if ℓ is observable from w.

It follows directly from Theorem 2.14 that $w = M(\sigma)\ell$ is an observable image representation if and only if $\operatorname{rank}(M(\lambda)) = \mathtt{l}$ for all $\lambda \in \mathbb{C}$.

In the following result we give necessary and sufficient conditions for a given $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ to induce a kernel representation of $\mathfrak{B} = \operatorname{im}(M(\sigma))$, see Definition A.23 in Appendix A for the definition of *left syzygy.*

Proposition 2.18. *Let* $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ *be such that* $\mathfrak{B} = im(M(\sigma))$, *and let* $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$. *Write* $R = \operatorname{col}(r_1, \ldots, r_{\mathtt{g}})$, *then the following statements are equivalent:*

1. $\mathfrak{B} = \ker(R(\sigma))$, *and*

2. $r_1, \ldots, r_{\mathtt{g}}$ *are generators of the left syzygy of* M.

Moreover, $r_1, \ldots, r_{\mathtt{g}}$ *are bases generators if and only if* R *induces a minimal kernel representation of* $\mathfrak{B}$.

Proof. (1) $\Longrightarrow$ (2) Since

$$\mathfrak{B} = \operatorname{im}(M(\sigma)) = \ker(R(\sigma))$$

then

$$w = M(\sigma)\ell,$$

for some latent variable, and

$$R(\sigma)w = 0.$$

Hence,

$$RM(\sigma)\ell = 0.$$

Now since M induces an image representation and ℓ is not restricted, i.e. it can take any value in $(\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}}$, then $RM(\sigma)\ell = 0$ implies that $RM = 0$ for all non-zero ℓ. Therefore, it follows from the definition of left syzygy that the rows of R are generators of the left syzygy of M.
(2) $\Longrightarrow$ (1) Assume $r_1, \ldots, r_{\mathtt{g}}$ span the left syzygy of M. Since $\mathfrak{B} = \text{im}(M(\sigma))$ then $w = M(\sigma)\ell$ for some latent variable. Define $R := \text{col}(r_1, \ldots, r_{\mathtt{g}})$ and compute

$$R(\sigma)w = RM(\sigma)\ell.$$

Since the rows of R span the left syzygy of M then $RM = 0$ which implies that $R(\sigma)w = RM(\sigma)\ell = 0$, hence, $R(\sigma)w = 0$.
To complete the proof (Only if) Recall that if r_i for $i = 1, \ldots, \mathtt{g}$ are bases generators then there are linearly independent, hence $R = \text{col}(r_1, \ldots, r_{\mathtt{g}})$ will be full row rank. Therefore, it follows from Theorem 2.8 that R induces a minimal kernel representation.
(If) From Theorem 2.8 R inducing a minimal kernel representation implies that R is full rank, hence, the rows of R are linearly independent. Now since the rows of R are generators of the the left syzygy of M and are linearly independent then they are bases generators. □

We now consider a special type of behaviours called autonomous.

Definition 2.19. Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$. $\mathfrak{B}$ is called *autonomous* if for all $w_1, w_2 \in \mathfrak{B}$,

$$[w_1(t) = w_2(t) \ \forall \ t \leq 0] \Longrightarrow [w_1 = w_2] \,.$$

For autonomous behaviours future trajectories are completely determined by past trajectories. Autonomous behaviours admit a square and full row rank kernel representation and they are finite-dimensional subspaces of $\mathscr{L}^{\mathtt{w}}$.

Proposition 2.20. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$. Then the following conditions are equivalent:*

i. *$\mathfrak{B}$ is autonomous,*

ii. *$\mathfrak{B}$ is finite dimensional, and*

iii. *$\mathfrak{B}$ admits a kernel representation $R(\sigma)w = 0$ where $R \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\xi]$ and det$(R) \neq 0$.*

Proof. See Proposition 4.5 pp. 241-242 of [81]. □

2.5 Equivalent behaviours and annihilators

The equivalence between two behaviours with different kernel representations can be explained using *module theory* (see Appendix A). Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$, then the $\mathbb{R}[\xi]$-*submodule* *of* $\mathbb{R}^{1\times \mathtt{w}}[\xi]$ *generated by the rows of* R is defined by

$$< R >:= \left\{ r \in \mathbb{R}^{1\times \mathtt{w}}[\xi] \mid \exists\ \upsilon \in \mathbb{R}^{1\times \mathtt{g}}[\xi] \text{ such that } r = \upsilon R \right\}.$$

Proposition 2.21. *Let* $\mathfrak{B}_1 = \ker(R_1(\sigma))$ *and* $\mathfrak{B}_2 = \ker(R_2(\sigma))$, *where* $R_1, R_2 \in \mathbb{R}^{\bullet\times \mathtt{w}}[\xi]$. *Then the following statements are equivalent,*

i. $\mathfrak{B}_1 = \mathfrak{B}_2$, *and*

ii. $< R_1 >=< R_2 >$.

Proof. The proof follows Theorem 35 p. 45 of [9]. □

Let R_1 and R_2 in Proposition 2.21 induce minimal representations, then the following hold true.

Proposition 2.22. *Let* $\mathfrak{B}_1 = \ker(R_1(\sigma))$ *and* $\mathfrak{B}_2 = \ker(R_2(\sigma))$ *where both* R_1 *and* R_2 *are of full row rank. Then* $\mathfrak{B}_1 = \mathfrak{B}_2$ *if and only if there exist a unimodular matrix* $U \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ *such that* $R_1 = UR_2$.

Proof. See Proposition III.3 p. 263 of [82]. □

The following results holds true about image representations.

Proposition 2.23. *Let* $\mathfrak{B}_1, \mathfrak{B}_2 \in \mathscr{L}^{\mathtt{w}}_{contr}$. *Assume that* $M_1, M_2 \in \mathbb{R}^{\mathtt{w}\times\bullet}[\xi]$ *induce an observable image representation of* $\mathfrak{B}_1, \mathfrak{B}_2$, *respectively. Then* $\mathfrak{B}_1 = \mathfrak{B}_2$ *if and only if there exists a unimodular matrix* $W \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ *such that* $M_1 = M_2W$.

Proof. See Theorem 6 pp. 247-248 of [68]. □

In the case where $\mathfrak{B}_1 \subseteq \mathfrak{B}_2$ then the following holds true.

Proposition 2.24. *Let* $\mathfrak{B}_2 = \ker(R_2(\sigma))$ *and* $\mathfrak{B}_1 = \ker(R_1(\sigma))$ *where* $R_2, R_1 \in \mathbb{R}^{\bullet\times \mathtt{w}}[xi]$. *Then* $\mathfrak{B}_1 \subseteq \mathfrak{B}_2$ *if and only if there exist a matrix* $F \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ *such that* $R_2 = FR_1$.

Proof. See Theorem V.5 p. 291 of [82]. □

Let $n \in \mathbb{R}^{1\times \mathtt{w}}[\xi]$, then n is called an *annihilator* of $w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ if $n(\sigma)w = 0$. Moreover, let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ if $n(\sigma)w = 0$ for all $w \in \mathfrak{B}$ then n is called an *annihilator of* $\mathfrak{B}$, thus we write $n(\sigma)\mathfrak{B} = 0$. The set of annihilators of $\mathfrak{B}$ is defined by

$$\mathfrak{N}_{\mathfrak{B}} := \left\{ n \in \mathbb{R}^{1\times \mathtt{w}}[\xi] \mid n(\sigma)\mathfrak{B} = 0 \right\}.$$

Since $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$, i.e. $\mathfrak{B}$ is shift invariant and closed, then for all $n_1, n_2 \in \mathfrak{N}_{\mathfrak{B}}$ and $\alpha \in \mathbb{R}[\xi]$ it follows that $n_1 + n_2 \in \mathfrak{N}_{\mathfrak{B}}$ and $\alpha n_i \in \mathfrak{N}_{\mathfrak{B}}$, for $i = 1, 2$. Therefore, $\mathfrak{N}_{\mathfrak{B}}$ is a $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times\mathtt{w}}[\xi]$, henceforth we refer to $\mathfrak{N}_{\mathfrak{B}}$ as the *module of annihilators.* The set of annihilators of $\mathfrak{B}$ of degree (see Appendix A, Definition A.11) at most $j \in \mathbb{Z}_+$ is defined by

$$\mathfrak{N}^j_{\mathfrak{B}} := \left\{ n \in \mathbb{R}^{1\times\mathtt{w}}[\xi] \mid n \in \mathfrak{N}_{\mathfrak{B}} \text{ and } \deg(n) \leqslant j \right\}.$$

Proposition 2.25. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ and $R \in \mathbb{R}^{\bullet\times\mathtt{w}}[\xi]$. Denote by $\mathfrak{N}_{\mathfrak{B}}$ the module of annihilators of $\mathfrak{B}$ and by $< R >$ the $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times\mathtt{w}}[\xi]$ spanned by the rows of R. Then $< R > = \mathfrak{N}_{\mathfrak{B}}$ if and only if $\mathfrak{B} = \ker(R(\sigma))$.*

Proof. See p. 169 of [87]. □

2.6 Inputs, outputs and integer invariants

We consider the behavioural definition of *input/output* systems. We also cover some integer invariants of interest in this thesis.

2.6.1 Inputs and outputs

Let $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}_1} \times \mathbb{R}^{\mathtt{w}_2}, \mathfrak{B})$ and $\Pi_{w_1}(\mathfrak{B})$ be the projection of $\mathfrak{B}$ onto a subset w_1 of its variables.

Definition 2.26. Let $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}_1} \times \mathbb{R}^{\mathtt{w}_2}, \mathfrak{B})$. Then w_1 is called *free* if $\Pi_{w_1}(\mathfrak{B}) = (\mathbb{R}^{\mathtt{w}_1})^{\mathbb{Z}}$.

w_1 is free if it can take any value in $(\mathbb{R}^{\mathtt{w}_1})^{\mathbb{Z}}$.

Definition 2.27. Let $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}_1} \times \mathbb{R}^{\mathtt{w}_2}, \mathfrak{B}) \in \mathscr{L}^{\mathtt{w}_1+\mathtt{w}_2}$. Σ is a *input/output system*, $\Sigma_{I/O}$, if

i. w_1 is free;

ii. w_2 does not contain any free components;

If conditions (i)-(ii) above are satisfied, then w_1 is called the *input variable* and w_2 is called the *output variable.* Moreover, w_1 is said to be *maximally free.* We shall use the standard notation u for input variables and y for output variables.

To characterize input/output systems in terms of representations first consider the following definitions.

Definition 2.28. Let $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}_1} \times \mathbb{R}^{\mathtt{w}_1}, \mathfrak{B}) \in \mathscr{L}^{\mathtt{w}_1+\mathtt{w}_2}$. w_2 does not *anticipate* w_1 if

$$\begin{aligned}&\left[(w_1', w_2'), (w_1'', w_2'') \in \mathfrak{B}\right] \text{ and } \left[w_1'(t) = w_1''(t) \text{ for } t \leqslant 0\right] \Rightarrow \\ &\left[\exists\ w_2 \text{ such that } (w_1'', w_2) \in \mathfrak{B} \text{ and } w_2(t) = w_2'(t) \text{ for } t \leqslant 0\right].\end{aligned}$$

Non-anticipating means that it is only the past and not the future of w_1 that influences the past of w_2. A non-anticipating input/output system is defined as follows.

Definition 2.29. Let $\Sigma = (\mathbb{Z}, \mathbb{R}_1^{\mathtt{w}_1} \times \mathbb{R}_2^{\mathtt{w}_2}, \mathfrak{B})$ with w_1 free. Σ is a *non-anticipating input/output system* if w_2 does not anticipate w_1.

We illustrate the notion of non-anticipation and the fact that input/output partitions are not unique with an example below.

Example 2.1. Consider a dynamical system whose behaviour $\mathfrak{B}$ consists of trajectories satisfying the following difference equation

$$w_1(t+1) + w_1(t) + w_2(t) - w_3(t) = 0 \text{ for } t \in \mathbb{Z}_+,$$

Hence, the system admits a kernel representation

$$\underbrace{\begin{bmatrix}\sigma + 1 & 1 & -1\end{bmatrix}}_{R}\begin{bmatrix}w_1 \\ w_2 \\ w_3\end{bmatrix} = 0. \tag{2.5}$$

Now consider the following equation

$$w_3(t) = w_1(t+1) + w_1(t) + w_2(t)$$

such that the system admits a representation

$$w_3 = \begin{bmatrix}\sigma + 1 & 1\end{bmatrix}\begin{bmatrix}w_1 \\ w_2\end{bmatrix}.$$

Then both w_1 and w_2 are free since they are not restricted, hence, it follows from Definition 2.27 that such partition, $((w_1, w_2), w_3)$, is a valid input/output partition with w_3 as the output. Now notice that w_3 is influenced by the future of w_1 and the present of w_2, therefore this input/output defines an anticipating input/output partition.

Now consider

$$w_1(t+1) + w_1(t) = w_3(t) - w_2(t),$$

hence,

$$\begin{bmatrix}\sigma + 1\end{bmatrix} w_1 = \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}w_2 \\ w_3\end{bmatrix}.$$

Such partition, $(w_1,(w_2,w_3))$, is an input/output as well, furthermore, the $(w_1,(w_2,w_3))$ partition is non-anticipating since w_1 is influenced only by the present not the future of w_3 and w_2. □

Let $\mathtt{m},\mathtt{p}\in\mathbb{Z}_+$, in Theorem 2.30 below, we give a characterization of input/output systems, see Definition A.17 in Appendix A for the definition of *proper rational functions.*

Theorem 2.30. *Let $\Sigma=(\mathbb{Z},\mathbb{R}^{\mathtt{m}}\times\mathbb{R}^{\mathtt{p}},\mathfrak{B})\in\mathscr{L}^{\mathtt{m}+\mathtt{p}}$. Then Σ is an input/output dynamical system if and only if*

$$\mathfrak{B}=\{(u,y):\mathbb{Z}\to(\mathbb{R}^{\mathtt{m}}\times\mathbb{R}^{\mathtt{p}})\mid P(\sigma)y=Q(\sigma)u\}$$

with $Q\in\mathbb{R}^{\mathtt{p}\times\mathtt{m}}[\xi]$, $P\in\mathbb{R}^{\mathtt{p}\times\mathtt{p}}[\xi]$ and $det(P)\neq 0$. Moreover, it defines a non-anticipating dynamical system if and only if $P^{-1}Q\in\mathbb{R}^{\mathtt{p}\times\mathtt{m}}(\xi)$ is matrix of proper rational functions.

Proof. Follows from Proposition VIII.6 p. 272 of [82]. □

Consider $w=[w_1,w_2\dots,w_{\mathtt{w}}]$ a vector in $\mathbb{R}^{\mathtt{w}}$, partition $\mathbf{w}:=\{1,2,\dots,\mathtt{w}\}$ into $\mathbf{w}_1:=\{i_1,i_2\dots,i_{\mathtt{m}}\}$ and $\mathbf{w}_2:=\{j_1,j_2\dots,j_{\mathtt{p}}\}$ such that $\mathbf{w}_1\cap\mathbf{w}_2=\varnothing$ and $\mathbf{w}_1\cup\mathbf{w}_2=\mathbf{w}$. Define the vectors $w_1:=[w_{i_1},w_{i_2}\dots,w_{i_{\mathtt{m}}}]\in\mathbb{R}^{\mathtt{m}}$ and $w_2:=[w_{j_1},w_{j_2}\dots,w_{j_{\mathtt{p}}}]\in\mathbb{R}^{\mathtt{p}}$. Then $w=\mathrm{col}(w_1,w_2)$ is called a *componentwise partition* of w.

Theorem 2.31. *Let $\Sigma=(\mathbb{Z},\mathbb{R}^{\mathtt{w}},\mathfrak{B})\in\mathscr{L}^{\mathtt{w}}$. Then there exists a permutation matrix $\Pi\in\mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ such that $\Pi w=:\mathrm{col}(w_1,w_2)$ is a componentwise partition of w such that $\Sigma=(\mathbb{Z},\mathbb{R}^{\mathtt{m}}\times\mathbb{R}^{\mathtt{p}},\mathfrak{B})$ is a non-anticipating input/output system.*

Proof. See Theorem VIII.7 p. 272 of [82]. □

It follows from Theorems 2.30 and 2.31 that input/output dynamical systems are parametrized by a triple (P,Q,Π) where $\Pi\in\mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ is a permutation matrix that identifies which components of w are inputs and which are outputs, i.e.

$$\Pi w=\begin{bmatrix}u\\ y\end{bmatrix}.$$

Now let $\mathfrak{B}=\ker(R(\sigma))$ such that $R\in\mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ induces a minimal representation. In Algorithm 1 on p. 23, we show how to determine $Q\in\mathbb{R}^{\mathtt{p}\times\mathtt{m}}[\xi]$ and $P\in\mathbb{R}^{\mathtt{p}\times\mathtt{p}}[\xi]$ such that $P^{-1}Q$ is matrix of proper rational functions.

Algorithm 1: Computation of input/output partition

Input : Full row rank $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ with $\mathtt{w} \geq \mathtt{g}$

Output : $Q \in \mathbb{R}^{\mathtt{g}\times\mathtt{m}}[\xi]$, $P \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ such that $P^{-1}Q$ is matrix of proper rational functions

1 Choose P as a submatrix with maximal determinantal degree among all $\mathtt{g} \times \mathtt{g}$ nonsingular submatrices of R.

2 Define the components of w corresponding to the columns of P as outputs.

3 Define the remaining components of w as inputs and the corresponding columns of R as Q.

Example 2.2. We revisit Example 2.1. Notice that from (2.5) the 1×1 submatrix of R with maximal determinantal degree corresponds to w_1, hence, having w_1 as the output results in a non-anticipating system. □

2.6.2 Integer invariants

We discuss some integer invariants associated with $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$.

- $\mathtt{w}(\mathfrak{B})$: The *number of components* of w.
- $\mathtt{p}(\mathfrak{B})$: The *output cardinality*, i.e. the number of output variables of $\mathfrak{B}$. Let $\mathfrak{B} = \ker(R(\sigma))$ where $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ induces a minimal representation then $\mathtt{p}(\mathfrak{B}) = \operatorname{rank}(R) = \mathtt{g}$. It follows from Proposition 2.20 that if $\mathfrak{B}$ is autonomous then $\mathtt{g} = \mathtt{w}$, hence, $\mathtt{p}(\mathfrak{B}) = \operatorname{rank}(R) = \mathtt{w}$.
- $\mathtt{m}(\mathfrak{B})$: The *input cardinality*, i.e. the number of inputs of $\mathfrak{B}$ and is defined by $\mathtt{m}(\mathfrak{B}) := \mathtt{w} - \mathtt{p}(\mathfrak{B})$. If $\mathfrak{B}$ is autonomous then $\mathtt{p}(\mathfrak{B}) = \mathtt{w}$, consequently, $\mathtt{m}(\mathfrak{B}) = 0$.
- $\mathtt{L}(\mathfrak{B})$: The *lag* of $\mathfrak{B}$. The smallest integer L such that (2.1) holds. Equivalently, $\mathtt{L}(\mathfrak{B})$ is the smallest degree over all $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ such that $\mathfrak{B} = \ker(R(\sigma))$.

Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ and $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ be such that $\mathfrak{B} = \ker(R(\sigma))$. Write $R = \operatorname{col}(r_1, \dots, r_{\mathtt{g}})$ where $r_i \in \mathbb{R}^{1\times\mathtt{w}}[\xi]$ for $i = 1, \dots, \mathtt{g}$. Now, denote by d_i the degrees of r_i and assume that $r_i \neq 0$ and that the rows of R have been permuted so that $0 \leq d_1 \leq d_2 \leq \cdots \leq d_{\mathtt{g}}$. Then $(d_1, d_2, \dots, d_{\mathtt{g}})$ is called the *lag structure* of R and d_1 is the called the *minimal lag* associated with R. Now define a total ordering on all lag structures by

$$\begin{aligned}
&\left[(d_1, d_2, \dots, d_{\mathtt{g}}) \leq (d'_1, d'_2, \dots, d'_{\mathtt{g}'})\right] :\Leftrightarrow \\
&\left[(d_1, d_2, \dots, d_{\mathtt{g}}) = (d'_1, d'_2, \dots, d'_{\mathtt{g}'})\right] \text{ or } \left[\mathtt{g} < \mathtt{g}'\right] \text{ or} \\
&\left[\exists\ \mathtt{p}\ \leq \mathtt{g} = \mathtt{g}' \text{such that } d_{\mathtt{p}+1} < d'_{\mathtt{p}+1} \text{ and } d_i = d'_i \text{ for } i = 1, \dots, \mathtt{p}\right].
\end{aligned}$$

Definition 2.32. Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ with lag structure $(d_1, d_2, \ldots, d_{\mathtt{g}})$ be such that $\mathfrak{B} = \ker(R(\sigma))$. Then $(d_1, d_2, \ldots, d_{\mathtt{g}})$ is called the *shortest lag structure* of $\mathfrak{B}$ if

$$\left[R' \in \mathbb{R}^{\mathtt{g}'\times\mathtt{w}}[\xi] \text{ with lag structure } (d'_1, d'_2, \ldots, d'_{\mathtt{g}'})\right] \text{ and } \left[\mathfrak{B} = \ker(R'(\sigma))\right] \Rightarrow \left[(d_1, d_2, \ldots, d_{\mathtt{g}}) \leq (d'_1, d'_2, \ldots, d'_{\mathtt{g}'})\right].$$

The minimal lag of the shortest lag structure of $\mathfrak{B}$ is called the *shortest lag* and is denoted by $\mathtt{l}(\mathfrak{B})$. Notice that the degree of each row of any kernel representation of $\mathfrak{B}$ is at least $\mathtt{l}(\mathfrak{B})$.

In Theorem 2.33 below, we give necessary and sufficient conditions for a kernel representation to define a shortest lag structure of $\mathfrak{B}$, where *row properness* has been defined in Definition A.13 in Appendix A.

Theorem 2.33. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ be such that $\mathfrak{B} = \ker(R(\sigma))$. Write $R = \mathrm{col}(r_1, \ldots, r_{\mathtt{g}})$ such that $0 \leq d_1 \leq d_2 \leq \cdots \leq d_{\mathtt{g}}$ where $d_i = deg(r_i)$ for $i = 1, \ldots \mathtt{g}$. Then $(d_1, d_2, \ldots, d_{\mathtt{g}})$ is the shortest lag structure of $\mathfrak{B}$ if and only if the matrix R is row proper.*

Proof. Follows from Theorem 6, p. 570 of [79]. □

Another integer invariant of interest is associated with a special latent variable called the *state* variable. In order to discuss this integer invariant we introduce the notion of concatenation of two trajectories. Let $w_1, w_2 \in \mathfrak{B}$, the *concatenation* of w_1 and w_2 at $t' \in \mathbb{Z}$ is the trajectory $w_1 \wedge_{t'} w_2$ whose value at $t \in \mathbb{Z}$ equals

$$(w_1 \wedge_{t'} w_2)(t) = \begin{cases} w_1(t) & \text{for } t < t' \\ w_2(t) & \text{for } t \geq t'. \end{cases}$$

Definition 2.34. Let $\Sigma_s = (\mathbb{T}, \mathbb{W}, \mathbb{X}, \mathfrak{B}_s)$. Then x is a *state* variable if

$$[(w_1, x_1), (w_2, x_2) \in \mathfrak{B}_s] \text{ and } [t \in \mathbb{Z}] \text{ and } [x_1(t) = x_2(t)] \Rightarrow [(w_1, x_1) \wedge_t (w_2, x_2) \in \mathfrak{B}_s]. \tag{2.6}$$

If (2.6) holds, then Σ_s is called a *state space system* and (2.6) is called the *state property*. The value of the state variable x at time t determines whether $w_1, w_2 \in \mathfrak{B}_s$ are concatenable at t.

$\Sigma_s = (\mathbb{T}, \mathbb{W}, \mathbb{X}, \mathfrak{B}_s)$ induces a manifest dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ with the manifest behaviour defined by

$$\mathfrak{B} := \{w \mid \exists\ x \text{ such that } (w, x) \in \mathfrak{B}_s\}.$$

Σ_s is called a *state space realization* of Σ. It is worth noting that state space realization of Σ is not unique, this fact is articulated on pp. 565-566 of [79]. This leads us to the following definition of minimal state space realization for linear time invariant complete systems.

Definition 2.35. Let $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$ then $\Sigma_s = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathbb{R}^{\mathtt{n}}, \mathfrak{B}_s)$ is a *minimal state space realization* of Σ if for any other state space realization $\Sigma'_s = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathbb{R}^{\mathtt{n}'}, \mathfrak{B}'_s)$ of Σ, $\mathtt{n} \leq \mathtt{n}'$.

The integer $\mathtt{n}$ in Definition 2.35 associated with a minimal state space realization of $\mathfrak{B}$ is called the *McMillan degree* and is denoted by $\mathtt{n}(\mathfrak{B})$.

2.7 Persistency of excitation

In order to introduce the notion of *persistency of excitation*, we define the following terms.

Definition 2.36. The *Hankel matrix* of depth $L \in \mathbb{N}$ associated with a vector $w(1), w(2), \dots, w(T)$, where $T \in \mathbb{N} \geq L$, is defined by

$$\mathfrak{H}_L(w) := \begin{bmatrix} w(1) & w(2) & \dots & w(T-L+1) \\ w(2) & w(3) & \dots & w(T-L+2) \\ \vdots & \vdots & \dots & \vdots \\ w(L) & w(L+1) & \dots & w(T) \end{bmatrix}. \tag{2.7}$$

$\mathfrak{H}_L(w)$ has L rows and $T-L+1$ columns. The Hankel matrix with an infinite number of columns and rows associated with an infinite vector w is denoted by $\mathfrak{H}(w) \in \mathbb{R}^{\infty\times\infty}$.

Definition 2.37. Let T and L as in Definition 2.36. A vector $u = u(1), u(2), \dots, u(T)$ is *persistently exciting* of order L if the Hankel matrix $\mathfrak{H}_L(u)$ has full row rank.

Now, to state the "fundamental lemma" cf. of [89], consider $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$ with lag $\mathtt{L}(\mathfrak{B})$, and McMillan degree $\mathtt{n}(\mathfrak{B})$. Let $\Pi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ be such that $\Pi w = \mathrm{col}(u, y)$ for $w \in \mathfrak{B}$, where u are inputs and y the corresponding outputs. Finally, let $\Delta \in \mathbb{Z}_+$ and denote by $\mathfrak{N}^{\Delta}_{\mathfrak{B}}$ the module of annihilators of $\mathfrak{B}$ of degree at most Δ and by $\tilde{\mathfrak{N}}^{\Delta}_{\mathfrak{B}}$ the set whose elements are coefficients vectors (see Definition A.9 in Appendix A) of the elements of $\mathfrak{N}^{\Delta}_{\mathfrak{B}}$. That is, for all $r \in \mathfrak{N}^{\Delta}_{\mathfrak{B}}$ then $\tilde{r} \in \tilde{\mathfrak{N}}^{\Delta}_{\mathfrak{B}}$, note that $\tilde{r}$ has a finite number of entries equal to Δ.

Lemma 2.38. *Let*

1. $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$*,*
2. $\Pi\tilde{w} = (\tilde{u}, \tilde{y}) \in \mathfrak{B}_{|[1,T]}$
3. $\Delta \in \mathbb{Z}_+$ *be such that* $\Delta > \mathtt{L}(\mathfrak{B})$

If $\tilde{u}$ is persistently exciting of order $\Delta + \mathtt{n}(\mathfrak{B})$, then

$$\textit{col span}(\mathfrak{H}_\Delta(\tilde{w})) = \mathfrak{B}_{|[1,\Delta]}$$

and

$$\textit{leftkernel}(\mathfrak{H}_\Delta(\tilde{w})) = \tilde{\mathfrak{N}}_{\mathfrak{B}}^{\Delta}.$$

Proof. Follows from the proof of Theorem 1 in [89] pp. 327-328. □

Under the assumption of Lemma 2.38, any Δ samples long trajectory $\tilde{w}' \in \mathfrak{B}_{|[1,\Delta]}$ can be written as a linear combination of the columns of $\mathfrak{H}_\Delta(\tilde{w})$. Consequently, any element of the subspace spanned by the columns of $\mathfrak{H}_\Delta(\tilde{w})$ is a trajectory of $\mathfrak{B}_{|[1,\Delta]}$. From Lemma 2.38 since leftkernel$(\mathfrak{H}_\Delta(\tilde{w})) = \tilde{\mathfrak{N}}_{\mathfrak{B}}^{\Delta}$ it follows that under the assumptions of the fundamental lemma we can recover from $\tilde{w}$ the laws of the system $\mathfrak{B}$ that generated $\tilde{w}$, but first we discuss the notion of identifiability as follows.

2.7.1 Identifiability

Clearly from Lemma 2.38 not every $\tilde{w} \in \mathfrak{B}_{|[1,T]}$ can be used to recover the laws of $\mathfrak{B}_{|[1,T]}$ that generated $\tilde{w}$. In fact take for example the trajectory $\tilde{w} = 0 \in \mathfrak{B}_{|[1,T]}$, it is obvious that such the trajectory can not be used to recover any laws of $\mathfrak{B}_{|[1,T]}$.

Let $N \in \mathbb{Z}_+$, and define the data set $\mathscr{D} := \{w^1, \ldots, w^N\}$ where $w^i \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ for $i = 1, \ldots, N$. In order to formally state the possibility of using $\mathscr{D}$ to find $\mathfrak{B}$, we define identifiability as follows.

Definition 2.39. Let $\mathfrak{B} \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$. $\mathfrak{B}$ is *identifiable* from the data set $\mathscr{D}$ if the following conditions holds:

1. $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$,
2. $\mathscr{D} \in \mathfrak{B}$, and
3. any $\mathfrak{B}' \in \mathscr{L}^{\mathtt{w}}$ and $\mathscr{D} \in \mathfrak{B}' \Longrightarrow \mathfrak{B} \subseteq \mathfrak{B}'$.

In the literature, $\mathfrak{B}$ satisfying conditions $(1) - (3)$ in Definition 2.39 is called the *most powerful unfalsified model (MPUM)* for $\mathscr{D}$ in $\mathscr{L}^{\mathtt{w}}$. This is, because from condition (2) $\mathfrak{B}$ is unfalsified by $\mathscr{D}$ and from condition (3) $\mathfrak{B}$ restricts more than any $\mathfrak{B}' \in \mathscr{L}^{\mathtt{w}}$ which is unfalisified by $\mathscr{D}$. It has been shown in Theorem 13 p.677 of [80] that MPUM exits for a given data set.

In the following result we give sufficient conditions for identifiability for controllable systems.

Theorem 2.40. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ and $\tilde{w} = (\tilde{u}, \tilde{y}) \in \mathfrak{B}$ where $\tilde{u}$ is an input trajectory and $\tilde{y}$ is the corresponding output trajectory. If $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $\tilde{u}$ is persistently exciting of order at least $\mathtt{L}(\mathfrak{B}) + \mathtt{n}(\mathfrak{B})$ then $\mathfrak{B}$ is identifiable from $\tilde{w}$.*

Proof. See Theorem 8.16 p. 121 of [43]. □

Following from Lemma 2.38 and Theorem 2.40 we state the following definition which will be used throughout this thesis.

Definition 2.41. Let $\tilde{w} \in \mathfrak{B}_{|[1,T]}$ and $\Delta \in \mathbb{Z}_+$ be such that $\Delta \geq \mathtt{L}(\mathfrak{B})$. $\tilde{w}$ is called *sufficiently informative* about $\mathfrak{B}$ if

1. col span$(\mathfrak{H}_\Delta(\tilde{w})) = \mathfrak{B}_{|[1,\Delta]}$,
2. leftkernel$(\mathfrak{H}_\Delta(\tilde{w})) = \tilde{\mathfrak{N}}^{\Delta}_{\mathfrak{B}}$.

Now, denote by $\tilde{\mathfrak{N}}_{\mathfrak{B}}$ the set whose elements are coefficient vectors of the elements of $\mathfrak{N}_{\mathfrak{B}}$. Notice that in this case the elements of $\tilde{\mathfrak{N}}_{\mathfrak{B}}$ have infinite number of entries which are zero everywhere except a finite number of entries. Sufficiently informative is defined as follows for $\tilde{w} \in \mathfrak{B}$.

Definition 2.42. Let $\tilde{w} \in \mathfrak{B}$. $\tilde{w}$ is called *sufficiently informative* about $\mathfrak{B}$ if

1. col span$(\mathfrak{H}(\tilde{w})) = \mathfrak{B}$, and
2. leftkernel$(\mathfrak{H}(\tilde{w})) = \tilde{\mathfrak{N}}_{\mathfrak{B}}$.

2.7.2 An algorithm for finding annihilators from data

In the literature there are several algorithms for computing a left kernel of the Hankel matrix, but in this thesis we shall adopt Algorithm 2 p. 679 of [80]. This is because such algorithm computes shortest lag structure annihilators of $\mathfrak{B}$, and this is necessary for our results in Chapter 5. For complete readability the algorithm is given on p. 28.

Theorem 2.43. *Let $\tilde{w} \in \mathfrak{B}$. From Algorithm 2 arrange $l_1, l_2, \dots, l_{\mathtt{p}}$ in a non-decreasing order and define $R :=$ col$(r_1, r_2, \dots, r_{\mathtt{p}})$. Then $(deg(r_1), deg(r_2), \dots, deg(r_{\mathtt{p}}))$ is the shortest lag structure of $\mathfrak{B}$. Furthermore, $j'_1, j'_2, \dots, j'_{\mathtt{p}}$th and $j_1, j_2, \dots, j_{\mathtt{m}}$th component of w are outputs and inputs variables, respectively.*

Proof. Follows from the proof of Theorem 14 p. 679 of [80]. □

Algorithm 2: Algorithm for computing shortest lag basis generators of annihilators of $\mathfrak{B}$

Input : $\tilde{w}$ (of dimension $\mathtt{w}$)
Output : $r_1, r_2, \ldots, r_p$ shortest lag basis generators of $\mathfrak{N}_{\mathfrak{B}}$
Assumptions: Theorem 2.40

1. Build the Hankel matrix $\mathfrak{H}(\tilde{w})$

2. Determination of dependency

i. Starting from the first row determine which rows of $\mathfrak{H}(\tilde{w})$ are linearly dependent on the preceding rows.

ii. Step (i) above results in an infinite column vector $\mathbf{d}$ whose elements are $*$ and $\circ$. A $*$ signifies that the i^{th} row of $\mathfrak{H}(\tilde{w})$ is linearly independent of preceding rows and a $\circ$ in the i^{th} element of $\mathbf{d}$ signifies that the i^{th} row of $\mathfrak{H}(\tilde{w})$ is linearly dependent.

3. Determination of input and output variables
If i^{th} row of $\mathbf{d}$ has $\circ$ then there will be $\circ$ in the $(i + n\mathtt{w})^{th}$ row of $\mathbf{d}$ for all $n \in \mathbb{Z}_+$.

i. Choose $j \in \mathbf{w}$, where $\mathbf{w}$ denotes the set $\{1, 2, \ldots, \mathtt{w}\}$. If there is a $*$ on each $(j + n\mathtt{w})^{th}$ row of $\mathbf{d}$ then the j^{th} component of w is an input, otherwise it is an output.

ii. Denote the indices of the input variables by $1 \leq j_1 \leq j_2 \leq \cdots \leq j_{\mathtt{m}} \leq \mathtt{w}$ and the output variable indices by $1 \leq j'_1 \leq j'_2 \leq \cdots \leq j'_{\mathtt{p}} \leq \mathtt{w}$.

4. Determination of the recursion

i. Choose $i \in \mathbf{w}$ such that the i^{th} variable of $\mathtt{w}$ is an output.

ii. Let $l_i := \min\{k \in \mathbb{Z}_+ | \text{the } (i + k\mathtt{w})^{th} \text{ element of } \mathbf{d} \text{ is a } \circ\}$.

iii. $(i + l_i\mathtt{w})^{th}$ row of $\mathfrak{H}(\tilde{w})$ can be written as a linear combination of the rows preceding it, i.e there exists a coefficient vector $\tilde{r}_{i,0}, \tilde{r}_{i,1}, \ldots, \tilde{r}_{i,l_i} \in \mathbb{R}^{1\times\mathtt{w}}$ such that

$$\tilde{w}_i(t + l_i) = \sum_{k=0}^{l_i} \tilde{r}_{i,k}\tilde{w}_i(t + k) \ t \in \mathbb{Z}$$

5. Specification of $r_1, r_2, \ldots, r_p$

i. Define $r_i(\xi) := [0, \ldots, 0, 1, 0, \ldots, 0](\xi^{l_i}) - \sum_{k=0}^{l_i} \tilde{r}_{i,k}\xi^k, \ i \in \{j'_1, j'_1, \ldots, j'_{\mathtt{p}}\}$

2.8 Summary

We have discussed some concepts of the behavioural framework that will be used throughout the thesis. In particular we defined linear difference system, their representations, controllable systems and the notions of persistency of excitation and identifiability.

Chapter 3

Control as interconnection

3.1 Introduction

The classical approach to control is the so called *intelligent control* (depicted in Figure 3.1). Under intelligent control a controller is viewed as a feedback signal processor, with measurements of the to-be-controlled system sensors being fed as inputs to the controller and outputs of the controller being fed as inputs to the to-be-controlled system actuators.

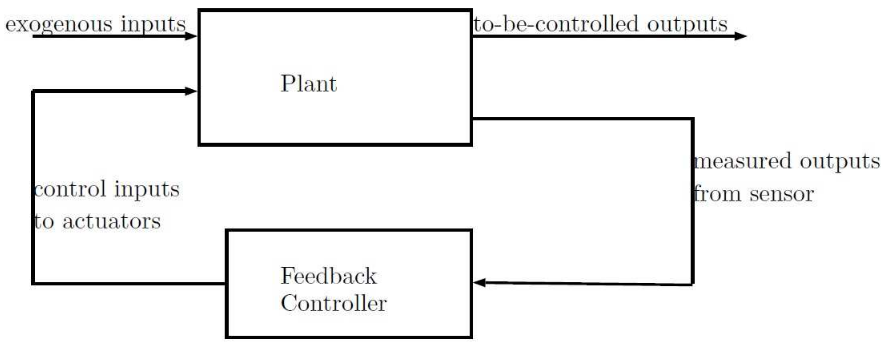

Figure 3.1: Intelligent control

Despite having its merits and many applications in the literature (see for example [10, 11, 63]), the fact that the intelligent controller imposes an input/output structure gives a restrictive view of control (see also Remark 1.1 in Chapter 1 why imposing input/output on the to-be-controlled systems is not always a good idea). Consequently, some control problems do not adhere to the intelligent control paradigm, for eaxmple see p. 329 of [84], where the door closing mechanism is used to show that imposing input/output structure is not the most natural framework for control design. More examples of control problem that do not conform to input/output structure are covered on pp. 5-8 of [2].

Since the 1990's there has been a paradigm shift, whereby the concepts of the behavioural approach to dynamical systems, covered in Chapter 2 and references within, are applied to control problems. This has resulted in the study of control problems without the need to impose an input/output structure on the to-be-controlled systems. Consequently, control is viewed as the interconnection of the to-be-controlled system and a controller. This is called *control as interconnection*. Under the control as interconnection paradigm, control is the restriction of the to-be-controlled system trajectories to a subset trajectories which meets the control objectives. It has be shown that under suitable conditions the feedback structure of the intelligent control paradigm is a special case of control as interconnection, see [84]. Ibid also shows how the door closing mechanism adhere to the interconnection paradigm.

Since its introduction, the interconnection paradigm has been studied by several authors and used to solve many control problems including, but not limited to, the following: In [2, 3, 4] stabilization and pole placement problems are studied in case where the system variables are partitioned into *control variables* and *to-be-controlled variables*; A solution to H_∞ control in the behavioural context is presented [69]; [13] the authors presents a solution to the problem of asymptotic tracking and regulation; In [75] the idea of the canonical controller is first introduced and is further studied in [26, 88]. For the first time, in this thesis control as interconnection will be used to solve data-driven control problems and control of flat systems.

3.2 Interconnection paradigm

Let $\Sigma \in \mathscr{L}^\bullet$ be the to-be-controlled system which consists of two types of variables. The first variables are called the *to-be-controlled*, denoted by w, these are the variables we intend to influence. The second variables are *interconnecting variables*, i.e. these are the variables used to interconnect a to-be-designed controller with the to-be-controlled system, in order to influence the to-be-controlled variables. These variables are called *control variables* and are denoted by c, see Figure 3.2.

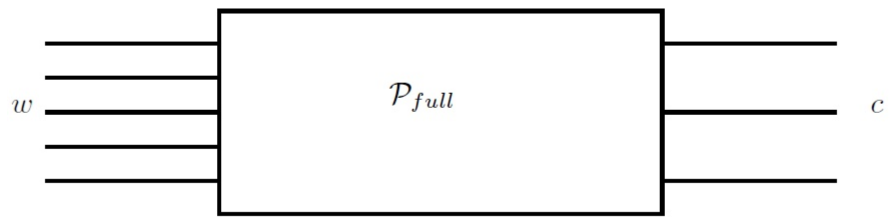

Figure 3.2: To-be-controlled system

Define the to-be-controlled system *full behaviour* by

$$\mathcal{P}_{full} := \{(w, c) : \mathbb{Z} \to (\mathbb{R}^{\mathtt{w}} \times \mathbb{R}^{\mathtt{c}}) \mid (w, c) \text{ satisfy system equations}\}$$

and the *manifest behaviour*, obtained by eliminating the control variables, by

$$\mathcal{P} := \{w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid \exists\, c \text{ such that } (w, c) \in \mathcal{P}_{full}\}\,.$$

Define a to-be designed *control behaviour* (depicted in Figure 3.3) by

$$\mathcal{C} := \{c : \mathbb{Z} \to \mathbb{R}^{\mathtt{c}} \mid c \text{ satisfy controller equations}\}\,.$$

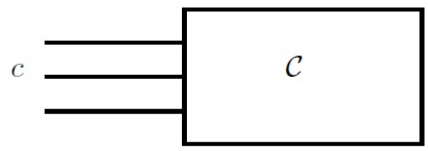

Figure 3.3: Controller behaviour

The interconnection of the controller and the to-be-controlled system through the control variables, called *partial interconnection* (shown in Figure 3.4), results in the *full controlled behaviour* defined by

$$\mathcal{K}_{full} := \{(w, c) : \mathbb{Z} \to (\mathbb{R}^{\mathtt{w}} \times \mathbb{R}^{\mathtt{c}}) \mid (w, c) \in \mathcal{P}_{full} \text{ and } c \in \mathcal{C}\} = \mathcal{P}_{full} \wedge_c \mathcal{C}.$$

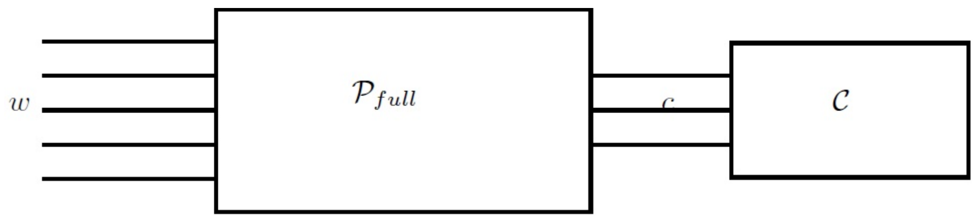

Figure 3.4: Partial interconnections

$\mathcal{K}_{full}$ induces the *manifest controlled behaviour* defined by

$$\mathcal{K} := \{w \mid \exists\, c \in \mathcal{C} \text{ such that } (w, c) \in \mathcal{K}_{full}\}\,. \tag{3.1}$$

$\mathcal{K}$ defined in (3.1) is obtained by the projection of $\mathcal{K}_{full}$ onto the to-be-controlled variables, i.e. $\mathcal{K} = \Pi_w(\mathcal{K}_{full})$. In this case, $\mathcal{C}$ is said to *implement* $\mathcal{K}$ or $\mathcal{K}$ is *implementable*.

In Theorem 3.2 we give necessary and sufficient conditions for the existence of a controller $\mathcal{C}$ such that $\mathcal{C}$ implement $\mathcal{K}$ but first consider the following definition of the hidden behaviours.

Definition 3.1. Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$. The *hidden behaviour* is defined by

$$\mathcal{N} := \{w \in \mathcal{P} \mid (w, 0) \in \mathcal{P}_{full}\}\,. \tag{3.2}$$

$\mathcal{N}$ consists of the to-be-controlled variables trajectories that occur when the control variable trajectories are restricted to zero. Therefore, the hidden behaviour is a sub-behaviour of the to-be-controlled system behaviour consisting of all trajectories $(w, 0)$.

Theorem 3.2. *Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$, $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}$ and $\mathcal{N} \in \mathscr{L}^{\mathtt{w}}$ as before. A controller $\mathcal{C} \in \mathscr{L}^{\mathtt{c}}$ exists, implementing $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}$ if and only if*

$$\mathcal{N} \subseteq \mathcal{K} \subseteq \mathcal{P}.$$

Proof. See Theorem 1 p. 55 of [90]. □

The inclusion $\mathcal{K} \subseteq \mathcal{P}$ means that the controlled behaviour $\mathcal{K}$ must be a subset of the to-be-controlled system behaviour $\mathcal{P}$ since the interconnection aims at restricting the system to a certain desired behaviour, therefore that behaviour must be a sub-behaviour of $\mathcal{P}$. Consequently, $\mathcal{K} \nsubseteq \mathcal{P}$ is not implementable by any controller. The inclusion $\mathcal{N} \subseteq \mathcal{K}$ is also important since it implies that whatever controller $\mathcal{C}$ is chosen it must be able to achieve the controlled behaviour of the system when $c = 0$.

A case of interest in this thesis arises when $\mathcal{N} = \{0\}$, see p. 55 of [90]. This means that the to-be-controlled variables w are observable from the control variables c. This is called *full information control*. Under full information control, any sub-behaviour of $\mathcal{P}$ is implementable.

Full information control is characterized as follows.

Proposition 3.3. *Let $\mathcal{P}_{full}$ and $\mathcal{N}$ in Theorem 3.2 and let $(w, c) \in \mathcal{P}_{full}$. Then the following statements are equivalent,*

1. *w is observable from c, and*
2. *$\mathcal{N} = \{0\}$.*

Proof. See Proposition 3 p. 72 of [70]. □

3.3 Full interconnection

Let $\mathcal{N} = \{0\}$, and any sub-behaviour of $\mathcal{P}$ be implementable. In this section we deal with a special interconnection whereby there is no separation between the control variables and the to-be-controlled variables, that is $w = c$, hence both $\mathcal{P}, \mathcal{C} \in \mathscr{L}^{\mathtt{w}}$. This interconnection is called the *full interconnection.* Define

$$\mathcal{P} := \{w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid w \ \text{ satisfies the system equations}\}$$

and

$$\mathcal{C} := \{w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid w \ \text{ satisfies the controller equations}\}\,.$$

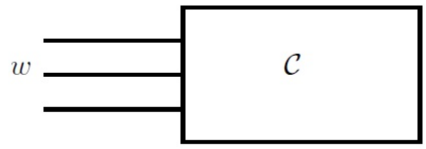

Figure 3.5: Full interconnection controller

Under full interconnection, the interconnection of the to-be-controlled system and the controller through the to-be-controlled variables results in the to-be-controlled variables obeying both the laws of the system and the controller, see Figure 3.6. Hence, the controlled behaviour is defined by

$$\mathcal{K} := \{w : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}} \mid w \ \in \mathcal{P} \text{ and } w \in \mathcal{C}\} = \mathcal{P} \wedge \mathcal{C}.$$

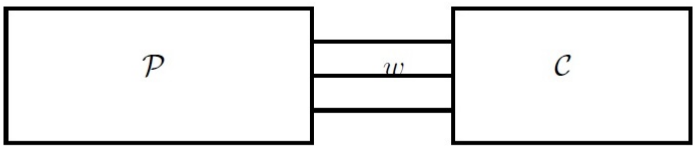

Figure 3.6: Full interconnection

Under full interconnection $\mathcal{K} = \mathcal{P} \cap \mathcal{C}$ as we now state using representations.

Proposition 3.4. *Let $\mathcal{K}, \mathcal{P}$ and $\mathcal{C}$ as already defined. Assume that $R, C \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$ are such that $\mathcal{P} = \ker(R(\sigma))$ and $\mathcal{C} = \ker(C(\sigma))$, respectively. If $\mathcal{C}$ implements $\mathcal{K}$ then $D :=$ col(R, C) induces a kernel representation of $\mathcal{K}$.*

Proof. Follows from the proof of Theorem 6 pp. 332-333 of [84]. □

In the proof of Theorem 6 of [84], a procedure for finding C in the case where R and any kernel representation of $\mathcal{K}$ is known is presented. The procedure uses the *Smith form* (see pp. 390-392 of [27]). Another procedure for computing C when both R and a kernel represenation of $\mathcal{K}$ are known has been covered in Algorithm 3.2 p. 2183 of [71]. The solution of [71] uses the `Polynomial Toolbox` in `Matlab`. In Chapter 5, section 5.3 of this thesis we shall focus on finding C when both R and D are unknown. In order to do that, under suitable conditions which will be specified, we shall use observed trajectories from both $\mathcal{P}$ and $\mathcal{K}$.

It is worth noting that Theorem 6 of [84] and Algorithm 3.2 of [71] are based on the assumption that $\mathcal{C}$ regularly implements $\mathcal{K}$.

Definition 3.5. Let $\mathcal{P}, \mathcal{C} \in \mathscr{L}^{\mathtt{w}}$ and $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}$ be the controlled behaviour obtained by the full interconnection of $\mathcal{P}$ and $\mathcal{C}$. Then the interconnection of $\mathcal{P}$ and $\mathcal{C}$ is called *regular* if

$$\mathtt{p}(\mathcal{K}) = \mathtt{p}(\mathcal{P}) + \mathtt{p}(\mathcal{C}).$$

We give a characterization of regular interconnections as follows, these results we will be useful in the data-driven results developed in this thesis.

Theorem 3.6. *Let $\mathcal{P} = \ker(R(\sigma))$ and $\mathcal{C} = \ker(C(\sigma))$, where $R, C \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$. Assume that $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}$ is the controlled behaviour obtained by the full interconnection of $\mathcal{P}$ and $\mathcal{C}$. Then $\mathcal{C}$ regularly implements $\mathcal{K}$ if and only if $< R > \cap < C >= \{0\}$.*

Proof. By the assumption that $\mathcal{C}$ implements $\mathcal{K}$ via full interconnection then $\mathrm{col}(R, C)$ induces a kernel representation of $\mathcal{K}$.
(Only If) Assume that $\mathtt{p}(\mathcal{K}) = \mathtt{p}(\mathcal{P}) + \mathtt{p}(\mathcal{C})$, then $\mathrm{rank}(\mathrm{col}(R, C)) = \mathrm{rank}(R) + \mathrm{rank}(C)$. Recall that $\mathrm{rank}(R)$ and $\mathrm{rank}(C)$ are equal to the number of bases generators of $< R >$ and $< C >$, respectively. Moreover, $\mathrm{rank}(\mathrm{col}(R, C))$ will be equal to number of bases generators of $< \mathrm{col}(R, C) >$. Therefore, the number of bases generators of $< \mathrm{col}(R, C) >$ being equals to the sum of the number of generators of $< R >$ and $< C >$ implies that $< R > \cap < C >= \{0\}$.
(If) Assume that $< R > \cap < C >= \{0\}$. Then $\mathrm{rank}(\mathrm{col}(R, C)) = \mathrm{rank}(R) + \mathrm{rank}(C)$. From the definition of output cardinality then $\mathtt{p}(\mathcal{P}) = \mathrm{rank}(R)$, $\mathtt{p}(\mathcal{C}) = \mathrm{rank}(C)$ and $\mathtt{p}(\mathcal{K}) = \mathrm{rank}(\mathrm{col}(R, C))$, hence, $\mathtt{p}(\mathcal{K}) = \mathtt{p}(\mathcal{P}) + \mathtt{p}(\mathcal{C})$. □

In Proposition 3.7, we give necessary and sufficient conditions for $\mathrm{col}(R, C)$ in Proposition 3.4 to induce a minimal representation of the controlled system.

Proposition 3.7. *Let $\mathcal{P} = \ker(R(\sigma))$ and $\mathcal{C} = \ker(C(\sigma))$, with $R, C \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$. Assume that $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}$ is the controlled behaviour obtained by the full interconnection of $\mathcal{P}$ and $\mathcal{C}$. Then $\mathrm{col}(R, C)$ induces a minimal representation of $\mathcal{K}$ if and only if $< R > \cap < C >= \{0\}$ and R, C are of full row rank.*

Proof. (If) Assume that $< R > \cap < C >= \{0\}$ and R, C are full row rank. Then $\mathrm{col}(R, C)$ is full row rank, therefore, it follows from Theorem 2.8 that $\mathrm{col}(R, C)$ induces

a minimal kernel representation $\mathcal{K}$.
(Only if) Assume that col(R, C) is full row rank. Then R and C are full row rank, moreover, rank(col(R, C)) = rank(R) + rank(C). Therefore, it follows from Theorem 3.6 that $< R > \cap < C >= \{0\}$. $\square$

3.4 Summary

In this chapter, we reviewed some concepts of control as interconnection, covering both full and partial interconnections. We covered the notion of regular interconnection for full interconnection. We gave a characterization of regular full interconnection. Using the material covered in this chapter and Chapter 2, in the next chapter we develop a solution to our first data-driven control problem.

Chapter 4

Data-driven control: prescribed path

In this chapter, we formulate and develop a solution to our first data-driven control problem. Using partial interconnection, we consider a case where one is given: a trajectory (w, c) from the to-be-controlled system; a to-be-controlled system hybrid representation; a *desired behaviour* (i.e. the desired controlled behaviour, denoted by $\mathcal{D}$) kernel representation and a desired prescribed path w_{pre} specified on the time interval t_0 to t_1. The problem is to find a control variable trajectory such that the to-be-controlled variables trajectory follows the given prescribed path w_{pre} for the specified time interval t_0 to t_1.

In order to find a solution, we first verify that the conditions of Theorem 3.2 are satisfied using the given representations. Then we use the given trajectories and the desired behaviour representation to compute a trajectory w'_d of the desired behaviour. Then we state sufficient conditions under which w'_d is sufficiently informative about the desired behaviour. Under such conditions, we then compute a trajectory w_d of the desired behaviour which is equal to the prescribed path on the specified time interval. Finally, we find a trajectory c_d corresponding to w_d. Our solution is summarised by a step-by-step Algorithm 3 on p. 46 and demonstrated using a simulation example.

4.1 Problem statement

Consider a to-be-controlled linear difference system with full behaviour $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}\times\mathtt{c}}$ and manifest behaviour $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}$. Assume that $R_1 \in \mathbb{R}^{\mathtt{p}\times\mathtt{w}}[\xi], M_1 \in \mathbb{R}^{\mathtt{p}\times\mathtt{c}}[\xi]$ and $R_2 \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ are such that

$$\mathcal{P}_{full} = \{(w, c) \mid R_1(\sigma)w = M_1(\sigma)c\} \quad (4.1)$$

and

$$\mathcal{P} = \{w \mid R_2(\sigma)w = 0\} \tag{4.2}$$

respectively. Now let $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}$ be the desired controlled behaviour and assume that $D_1 \in \mathbb{R}^{\mathtt{t}\times\mathtt{w}}[\xi]$ is such that

$$\mathcal{D} = \{w \mid D_1(\sigma)w = 0\}\,. \tag{4.3}$$

The prescribed path problem is formally stated as follows.

Problem 4.1. "Prescribed path" problem. *Assume that*

1. *c is observable from w; and*
2. *$\mathcal{D}, \mathcal{P} \in \mathscr{L}^{\mathtt{w}}_{contr}$;*

Given

- *$(w, c) \in \mathcal{P}_{full}$;*
- *a prescribed trajectory $w_{pre} \in \mathcal{D}_{|[t_0,t_1]}$ with $t_0, t_1 \in \mathbb{N}$, $t_0 \leqslant t_1$; and*
- *R_1, M_1 in (4.1) and D_1 in (4.3) such that they induce minimal representations of their respective behaviours.*

Find, if it exists, a control variable trajectory c_d, such that there exists $w_d : \mathbb{Z} \to \mathbb{R}^{\mathtt{w}}$ such that

a. $(w_d, c_d) \in \mathcal{P}_{full}$,

b. $w_{d_{|[t_0,t_1]}} = w_{pre}$.

Remark 4.2. We assume that given trajectories are infinitely long, in practical applications the given trajectories have finite length. The problem of consistency (see [37]), i.e. the convergence of the identified system to the "true system" as the length of observed trajectories tends to infinity, is of paramount importance. This is a matter for future research. □

4.2 Prescribed path solution

In this section we present our solution to Problem 4.1. In order to find c_d in Problem 4.1 we first need to verify that $\mathcal{D}$ is implementable, i.e. check if conditions of Theorem 3.2 are satisfied. Such verification is outlined in the next subsection. If a controller, $\mathcal{C} \in \mathscr{L}^{\mathtt{c}}$, exists implementing $\mathcal{D}$ then Problem 4.1 has a solution otherwise there is no solution, moreover, we shall show later that c_d a solution to Problem 4.1 belongs to $\mathcal{C}$ that implements $\mathcal{D}$.

4.2.1 Implementability

By the assumption that c is observable from w then $\mathcal{N} = \{0\}$, therefore, we only need to verify $\mathcal{D} \subseteq \mathcal{P}$, as we now explain. Recall from Proposition 2.24 that $\mathcal{D} \subseteq \mathcal{P}$ if and only if there exists $F \in \mathbb{R}^{\mathtt{g}\times\mathtt{t}}[\xi]$ such that $R_2 = FD_1$. Therefore, we need to prove the existence of F such that $R_2 = FD_1$, otherwise there is no solution to Problem 4.1. Notice that to compute R_2 from (4.1), one can use the elimination procedure stated in Theorem 2.12.

4.2.2 To-be-controlled variable trajectory

We show how to find w_d such that $w_{d_{|[t_0,t_1]}} = w_{pre}$.

Theorem 4.3. *Let $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and D_1 in (4.3) induces a minimal kernel representation. Then there exists a matrix $Q \in \mathbb{R}^{\mathtt{w}\times\mathtt{t}}[\xi]$ such that $D_1Q = I_{\mathtt{t}}$. Moreover, im$((I_{\mathtt{w}} - QD_1)(\sigma)) = \ker(D_1(\sigma))$. Define w'_d by*

$$w'_d := (I_{\mathtt{w}} - QD_1)(\sigma)w, \tag{4.4}$$

where $w \in \mathcal{P}$, then $w'_d \in \mathcal{D}$.

Proof. The existence of Q such that $D_1Q = I_{\mathtt{t}}$ follows from the fact that D_1 induces a minimal kernel representation and that $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$, consequently, $D_1(\lambda)$ is full row rank for all $\lambda \in \mathbb{C}$. Therefore, D_1 is left prime, consequently, D_1 admits a right inverse Q. To show $\mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) = \ker(D_1(\sigma))$, we start with the inclusion $\mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) \subseteq \ker(D_1(\sigma))$. For w define

$$w' := (I_{\mathtt{w}} - QD_1)(\sigma)w.$$

Now compute

$$\begin{aligned} D_1(\sigma)w' &= D_1(\sigma)((I_{\mathtt{w}} - QD_1)(\sigma)w) \\ &= D_1(\sigma)w - D_1QD_1(\sigma)w. \end{aligned}$$

Since $D_1Q = I_{\mathtt{t}}$ it follows that

$$\begin{aligned} D_1(\sigma)w' &= D_1(\sigma)w - D_1(\sigma)w \\ &= 0. \end{aligned}$$

Hence $\mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) \subseteq \ker(D_1(\sigma))$. To prove the converse inclusion, assume by contradiction that there exists $w' \in \mathcal{D}$ such that $w' \notin \mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma))$. Now

$$\begin{aligned}(I_{\mathtt{w}} - QD_1)(\sigma)w' &= w' - (QD_1)(\sigma)w' \\ &= w',\end{aligned}$$

which implies that $w' \in \mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma))$. Therefore, $\mathrm{im}((I - QD_1)(\sigma)) = \ker(D_1(\sigma))$. To prove $w'_d \in \mathcal{D}$, notice that since $\mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) = \ker(D_1(\sigma))$ and $\mathcal{D} = \ker(D_1(\sigma))$ then $w'_d \in \mathcal{D}$. □

Note that even though $\mathcal{P} \supseteq \mathcal{D}$, $\mathcal{P}$ and $\mathcal{D}$ need not necessarily have the same input/output structure, as we now show in Lemma 4.4 below.

Lemma 4.4. *Let $\mathcal{P} = \ker(R_2(\sigma))$, where $R_2 \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ induces a minimal representation and $\mathcal{D} \subseteq \mathcal{P}$. Then there exists $D' \in \mathbb{R}^{(\mathtt{t}-\mathtt{g})\times\mathtt{w}}[\xi]$ such that $D_1 := \mathrm{col}(R_2, D')$ induces a minimal representation of $\mathcal{D}$. Moreover, $\mathtt{p}(\mathcal{D}) \geq \mathtt{p}(\mathcal{P})$.*

Proof. Let $\mathfrak{N}_{\mathcal{D}}$ and $\mathfrak{N}_{\mathcal{P}}$ denote the module of annihilators of $\mathcal{D}$ and $\mathcal{P}$, respectively. Since $\mathcal{D} \subseteq \mathcal{P}$ then $\mathfrak{N}_{\mathcal{P}} \subseteq \mathfrak{N}_{\mathcal{D}}$ (this fact follows from the proof of Proposition 2.24). Write $R_2 = \mathrm{col}(r_1, \dots, r_{\mathtt{g}})$, since R_2 is minimal then $r_1, \dots, r_{\mathtt{g}}$ are bases generators of $\mathfrak{N}_{\mathcal{P}}$. Moreover, since $\mathfrak{N}_{\mathcal{P}} \subseteq \mathfrak{N}_{\mathcal{D}}$ then there exists $r'_{\mathtt{g}+1}, \dots r'_{\mathtt{t}}$ such that $r_1, \dots, r_{\mathtt{g}}, r'_{\mathtt{g}+1}, \dots r'_{\mathtt{t}}$ are bases generators of $\mathfrak{N}_{\mathcal{D}}$. Define $D' := \mathrm{col}(r'_{\mathtt{g}+1}, r'_{\mathtt{g}+2}, \dots r'_{\mathtt{t}})$. Now the rows of $D_1 = \mathrm{col}(R_2, D')$ span $\mathfrak{N}_{\mathcal{D}}$ and form a basis of $\mathfrak{N}_{\mathcal{D}}$, hence, D_1 induces a minimal representation of $\mathcal{D}$. Now notice that $\mathtt{p}(\mathcal{P}) = \mathtt{g}$ and $\mathtt{p}(\mathcal{D}) = \mathtt{g} + (\mathtt{t} - \mathtt{g}) = \mathtt{t}$, hence, $\mathtt{t} > \mathtt{g}$ means that $\mathcal{D}$ has more output variables. □

It follow from Lemma 4.4 that we can define $\mathrm{col}(u, y) := \Pi w$ and $\mathrm{col}(u', y') := \Pi' w'_d$, where $\Pi, \Pi' \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ and u, u' are inputs. Partition $\Pi' = \mathrm{col}(\Pi'_u, \Pi'_y)$ compatibly with the partition of $\Pi' w'_d = \mathrm{col}(u', y')$. Then the following result holds true about the polynomial operator in the shift $\Pi'_u(I_{\mathtt{w}} - QD_1)(\sigma)$.

Lemma 4.5. *Let $\Pi' = \mathrm{col}(\Pi'_u, \Pi'_y)$ as above and $(I_{\mathtt{w}} - QD_1)$ as in Theorem 4.3. Then $\Pi'_u(I_{\mathtt{w}} - QD_1)$ is full row rank.*

Proof. Since $\mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) = \ker(D_1(\sigma))$, see Theorem 4.3, and u' is an input variable for $\ker(D_1(\sigma))$ then $\Pi'_u \mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) = (\mathbb{R}^{\mathtt{m}(\mathcal{D})})^{\mathbb{Z}}$. Now since $\Pi'_u \mathrm{im}((I_{\mathtt{w}} - QD_1)(\sigma)) = \mathrm{im}(\Pi'_u(I_{\mathtt{w}} - QD_1)(\sigma))$ then $\Pi'_u(I_{\mathtt{w}} - QD_1)$ is full row rank. Consequently, $\Pi'_u(I_{\mathtt{w}} - QD_1)(\sigma)$ is surjective see Proposition 2.5. □

To show sufficient conditions for w'_d to be sufficiently infomative about $\mathcal{D}$, first define $F_u := \Pi'_u - \Pi'_u QD_1$ and denote by $\tilde{F}_u$ the coefficient matrix of F_u with a finite number of block-columns, see Appendix A Definition A.10 for the definition of the *coefficient* *matrix* of a polynomial matrix and the right shift of a coefficient matrix. Now, denote by $< F_u >$ the $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times\bullet}[\xi]$ generated by the rows of F_u, and by $\mathfrak{N}_{\mathcal{P}}$ the module of annihilators of $\mathcal{P}$.

Theorem 4.6. *Let $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $w \in \mathcal{P}$ be sufficiently informative about $\mathcal{P}$. If $< F_u > \cap\, \mathfrak{N}_{\mathcal{P}} = \{0\}$ and u is persistently exciting of order at least $\mathtt{L}(\mathcal{P}) + \mathtt{n}(\mathcal{P})$ then u' is persistently exciting of order at least $\mathtt{L}(\mathcal{D}) + \mathtt{n}(\mathcal{D})$.*

Proof. Let $L \in \mathbb{Z}_+$ be such that $L-\deg(F_u) \geqslant \mathtt{L}(\mathcal{D})+\mathtt{n}(\mathcal{D})$. Denote by $\tilde{F}_u$ the coefficient matrix of F_u with a finite number L of block-columns. Since u' is the input variable of $\ker(D_1(\sigma))$ then it follows from (4.4) that

$$u' = F_u(\sigma)w.$$

Therefore, the following equality holds

$$\mathfrak{H}_{L-\deg(F_u)}(u') = \mathrm{col}(\sigma_R^k \tilde{F}_u)_{k=0,\ldots,L-1-\deg(F_u)} \mathfrak{H}_L(w).$$

Now assume by contradiction that u' is not persistently exciting, then there exists a non-zero vector $\tilde{\alpha} \in \mathbb{R}^{1\times(L-\deg(F_u))\mathtt{m}(\mathcal{D})}$ such that $\tilde{\alpha}\mathfrak{H}_{L-\deg(F_u)}(u') = 0$. Consequently, $\tilde{\alpha}\mathrm{col}(\sigma_R^k \tilde{F}_u)_{k=0,\ldots,L-1-\deg} \in \mathrm{leftkernel}(\mathfrak{H}_L(w))$. Now let $\alpha \in \mathbb{R}^{1\times\bullet}[\xi]$ to be the polynomial vector whose coefficient matrix is $\tilde{\alpha}$. Since u is persistently exciting and $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}_{contr}$ then $\mathrm{leftkernel}(\mathfrak{H}(w)) = \tilde{\mathfrak{N}}_{\mathcal{P}}$. Therefore, $\alpha F_u \in \mathfrak{N}_{\mathcal{P}}$, moreover, $\alpha F_u \neq 0$ (see Lemma 4.5), which contradicts $< F_u > \cap\, \mathfrak{N}_{\mathcal{P}} = \{0\}$. □

Remark 4.7. The lags $\mathtt{L}(\mathcal{D})$, $\mathtt{L}(\mathcal{P})$ and McMillan degrees $\mathtt{n}(\mathcal{D})$, $\mathtt{n}(\mathcal{P})$ are not known a priori. Therefore, all trajectories must be generated with input variable trajectories that are persistently exciting of some sufficiently high order. Moreover, in the rest of this thesis any number L greater than $\mathtt{L}(\mathcal{D})$ or $\mathtt{L}(\mathcal{P})$ shall be chosen to be "sufficiently large". □

We now proceed to show how to find $w_d \in \mathcal{D}$ such that $w_{d_{|[t_0,t_1]}} = w_{pre}$ under the assumptions of Theorem 4.6. Let $L > \mathtt{L}(\mathcal{D})$ and recall that if w'_d is sufficiently informative then for all $w' \in \mathcal{D}_{|[1,L]}$ there exists $\upsilon^\top \in \mathbb{R}^{1\times\infty}$ such that $w' = \mathfrak{H}_L(w'_d)\upsilon$ (see Lemma 2.38). Therefore, given $w_{pre} \in (\mathbb{R}^{\mathtt{w}})^{[t_0,t_1]}$ with $0 \leqslant t_0 \leqslant t_1 \leqslant L$, the computation of w_d such that $w_{d_{|[t_0,t_1]}} = w_{pre}$ amounts to finding υ if it exists such that $w_d = \mathfrak{H}_L(w'_d)\upsilon$. Define $H := \mathfrak{H}_{L,J}(w'_d) \in \mathbb{R}^{L\times J}$ with $J \in \mathbb{Z}_+$ such that $J \gg L$. Define H_1 as the block partition of the rows of H from row $\mathtt{w}t_0$ to row $\mathtt{w}t_1$ and solve for υ in

$$H_1\upsilon = w_{pre}. \tag{4.5}$$

If (4.5) has no solution then $w_{pre} \notin \mathcal{D}_{|[t_0,t_1]}$, therefore, we can not compute $w_d \in \mathcal{D}$ such that $w_{d_{|[t_0,t_1]}} = w_{pre}$. Otherwise, $w_d \in \mathcal{D}$ such that $w_{d_{|[t_0,t_1]}} = w_{pre}$ is defined by

$$w_d := \mathfrak{H}(w'_d)\upsilon \tag{4.6}$$

where $\mathfrak{H}(w'_d) \in \mathbb{R}^{\infty \times J}$.

Remark 4.8. Since $J \gg L$ then H_1 has more columns that rows, if υ exists such that (4.5) holds then it is not unique. Let A be the matrix whose columns are a basis of $\ker(H_1)$ and $\bar{\upsilon}$ be a particular solution of (4.5). Then the set of all possible solutions for (4.5) is defined by

$$\mathscr{S} := \{\bar{\upsilon} + \mathsf{A}v | v \in \mathbb{R}^{\mathtt{g}}\}$$

where $\mathtt{g}$ is the number of columns of A. □

Theorem 4.9. *Let $w \in \mathcal{P}$ be sufficiently informative about $\mathcal{P}$, and $< F_u > \cap \mathfrak{N}_{\mathcal{P}} = \{0\}$. Then w'_d in (4.4) is sufficiently informative about $\mathcal{D}$. Moreover, if $w_{pre} \in \mathcal{D}_{|[t_0,t_1]}$ then w_d defined in (4.6) belongs to $\mathcal{D}$ with w_{pre} as the prescribed path.*

Proof. The fact that w'_d in (4.4) is sufficiently informative about $\mathcal{D}$ follows from Theorem 4.6, therefore, col span$(\mathfrak{H}(w'_d)) = \mathcal{D}$. Now since $w_{pre} \in \mathcal{D}_{|[t_0,t_1]}$ then $\upsilon \in \mathscr{S}$ exists such that (4.5) holds, therefore, $\mathfrak{H}(w'_d)\upsilon = w_d$. Let H_1 as in (4.5) and $\mathfrak{H}(w'_d)_{|[t_0,t_1]}$ be the block rows of $\mathfrak{H}(w'_d)$ from row $\mathtt{w}t_0$ to row $\mathtt{w}t_1$. Then $\mathfrak{H}(w'_d)_{|[t_0,t_1]} = H_1$ which implies that $w_{d_{|[t_0,t_1]}} = w_{pre}$. □

4.2.3 Control variable trajectory

Under the assumptions of Theorem 4.9, we now compute a control variable trajectory corresponding to w_d defined in (4.6).

Since c is observable from w, then it follows from section 2.3 in Chapter 2 that there exists $O \in \mathbb{R}^{\mathtt{c}\times\mathtt{w}}[\xi]$ such that

$$(w, c) \in \mathcal{P}_{full} \Rightarrow c = O(\sigma)w. \tag{4.7}$$

To find O in (4.7) we use the given hybrid representation as follows.

Since M_1 and R_1 in (4.1) induces a minimal hybrid presentation and c is observable from w, then $M_1(\lambda)$ is full column rank for all $\lambda \in \mathbb{C}$, hence M_1 is right prime. Let a left inverse of M_1 be $K \in \mathbb{R}^{\mathtt{c}\times\mathtt{p}}[\xi]$ and define

$$O := KR_1,$$

then O satisfies (4.7). Consequently, c_d corresponding to w_d is defined by

$$c_d := O(\sigma)w_d. \tag{4.8}$$

Furthermore, if $\mathcal{C}$ implements $\mathcal{D}$ then $c_d \in \mathcal{C}$ as we show in Lemma 4.11. First we need the following result.

Proposition 4.10. *Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$, $\mathcal{C} \in \mathscr{L}^{\mathtt{c}}$ and $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}$ as before. Assume that c is observable from w. If $\mathcal{C}$ implements $\mathcal{D}$ via partial interconnection through c with respect to $\mathcal{P}_{full}$ then for all $(w, c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$ then $c \in \mathcal{C}$.*

Proof. By the assumption that $\mathcal{C}$ implements $\mathcal{D}$ then

$$\mathcal{P}_{full} \wedge_c \mathcal{C} = \{w \mid \exists\ c \in \mathcal{C} \text{ such that } (w, c) \in \mathcal{P}_{full}\} = \mathcal{D}.$$

Let $(w, c), (w, c') \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$, and assume by contradiction that $c \in \mathcal{C}$ and $c' \notin \mathcal{C}$. Then by linearity $(0, c - c') = (w, c) - (w, c') \in \mathcal{P}_{full}$ now by the assumption that c is observable from w then $(0, c - c')$ implies that $c = c'$, hence, contradiction. □

It follows from Proposition 4.10 that under observability assumption a controller that implements a given desired behaviour is unique.

Lemma 4.11. *Let w_d defined in (4.6) satisfying conditions of Theorem 4.9 and c_d in (4.8). Under the assumption that c is observable from w, if a controller $\mathcal{C}$ implements $\mathcal{D}$ then c_d belongs to $\mathcal{C}$. Moreover, c_d imposes the prescribed path w_{pre} on the to-be-controlled variable trajectory for the time interval $[t_0, t_1]$.*

Proof. Since $w_d \in \mathcal{D}$ and by the observability assumption c_d is the control variable trajectory such that $(w_d, c_d) \in \mathcal{P}_{full}$. Moreover, since $w_d \in \mathcal{D}$ then it follows from Proposition 4.10 that $c_d \in \mathcal{C}$. c_d imposing w_{pre} follows from the fact that c_d corresponds to w_d such that $w_{d_{[t_0,t_1]}} = w_{pre}$. □

Remark 4.12. From Lemma 4.11 to use c_d to find $\mathcal{C}$ that interconnects with $\mathcal{P}_{full}$ there are three alternatives to consider. Firstly, in this thesis we covered conditions for identifiability in the case when the input variable trajecories are persistently exciting. However, in the case that c_d has structured inputs, i.e. the input variable trajectories of c_d is not persistently exciting, one can still find the MPUM as explained on p. 288 of [82]. Secondly, if c_d is a vector exponential trajectory then the MPUM can also be computed as shown on p. 288 of [82]. Finally, if $\mathcal{D}$ is implemtatable then it can be implemented by the canonical controller, see [26, 88] in order to realise w_d. □

4.3 Algorithm

The prescribed path solution is summarised in Algorithm 3 below.

Algorithm 3: Solution for Problem 4.1

Input : R_1,M_1,D_1, $(w, c) \in \mathcal{P}_{full}$, t_0, t_1 and w_{pre}

Output : c_d

Assumptions: Theorem 4.9

1 Verify $\mathcal{D} \subseteq \mathcal{P}$ (See subsection 4.2.1). If $\mathcal{D} \not\subseteq \mathcal{P}$, stop. Otherwise go to step 2.

2 Compute Q such that $D_1 Q = I_{\mathtt{t}}$.

3 Define $w'_d := (I_{\mathtt{w}} - QD_1)(\sigma)w$.

4 Choose L and J such that $L > \mathtt{L}(\mathcal{D})$ (see remark 4.7) and $J \gg L$.

5 Define $H := \mathfrak{H}_{L,J}(w'_d)$ and H_1 as a partition of rows of H from row $\mathtt{w}t_0$ to row $\mathtt{w}t_1$.

6 Solve $H_1 \upsilon = w_{pre}$ for υ.

7 **if** *no solution for* υ **then**

8 $\quad$ $w_{pre} \notin \mathcal{D}_{|[t_0,t_1]}$ [No Solution for c_d]. Stop.

9 **else**

10 $\quad$ Build $\mathfrak{H}(w'_d) \in \mathbb{R}^{\infty \times J}$;

11 $\quad$ Define $w_d := \mathfrak{H}(w'_d)\upsilon$;

12 $\quad$ Compute K such that $KM_1 = I_{\mathtt{c}}$;

13 $\quad$ Define $O := KR_1$;

14 $\quad$ Compute $c_d = O(\sigma)w_d$.

4.4 Example

In this section we shall demonstrate Algorithm 3 with a simulation example. The data generated is finite length 5000, see Remark 4.2. Consider the to be-controlled system whose full behaviour $\mathcal{P}_{full}$ is induced by

$$\underbrace{\begin{bmatrix} \sigma + \frac{1}{2} & 1 & 0 & 1 \\ 0 & \sigma + \frac{1}{3} & 1 & 0 \\ 0 & 0 & \sigma + \frac{1}{4} & 1 \\ 0 & 0 & 0 & \sigma + \frac{1}{5} \end{bmatrix}}_{R_1} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}_{M_1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{4.9}$$

To eliminate the control variables we pre-multiply (4.9) by

$$U = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}.$$

Hence, $\mathcal{P}$ admits a kernel representation

$$\underbrace{\begin{bmatrix} \sigma+\frac{1}{2} & 1 & -\sigma-\frac{1}{4} & 0 \\ 0 & \sigma+\frac{1}{3} & 1 & -\sigma-\frac{1}{5} \end{bmatrix}}_{R_2} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0.$$

The desired behaviour $\mathcal{D}$ chosen is induced by

$$\underbrace{\begin{bmatrix} \sigma+\frac{1}{2} & 1 & -\sigma-\frac{1}{4} & 0 \\ 0 & \sigma+\frac{1}{3} & 1 & -\sigma-\frac{1}{5} \\ 0 & 0 & \sigma+\frac{1}{6} & 1 \end{bmatrix}}_{D_1} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0. \tag{4.10}$$

Define

$$F := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

then $R_2 = FD_1$, therefore, $\mathcal{D}$ is implementable.

Using `rightInverse` command in `Singular`[1] a right inverse of D_1 is

$$Q = \begin{bmatrix} -\frac{240}{137}\xi + \frac{6}{137} & \frac{3870}{3151}\xi + \frac{24273}{12604} & \frac{5004}{3151}\xi - \frac{9471}{31510} \\ \frac{120}{137}\xi^2 + \frac{44}{137}\xi + \frac{124}{137} & -\frac{1935}{3151}\xi^2 - \frac{3792}{3151}\xi - \frac{8409}{12604} & -\frac{2502}{3151}\xi^2 - \frac{2532}{15755}\xi + \frac{5793}{31510} \\ -\frac{120}{137}\xi - \frac{40}{137} & \frac{1935}{3151}\xi + 74556302 & \frac{2502}{3151}\xi + \frac{423}{3151} \\ \frac{120}{137}\xi^2 + \frac{60}{137}\xi + \frac{20}{411} & -\frac{1935}{3151}\xi^2 - \frac{4050}{3151}\xi - \frac{2485}{12604} & -\frac{2502}{3151}\xi^2 - \frac{840}{3151}\xi + \frac{6161}{6302} \end{bmatrix}$$

and the matrix QD_1 is

$$\left[\begin{matrix} -1.752\xi^2 - 0.832\xi + 0.022 & 1.228\xi^2 + 0.583\xi + 0.686 \\ 0.876\xi^3 + 0.759\xi^2 + 1.066\xi + 0.453 & -0.614\xi^3 - 0.532\xi^2 - 0.747\xi + 0.683 \\ -0.876\xi^2 - 0.73\xi - 0.146 & 0.614\xi^2 + 0.512\xi + 0.102 \\ 0.876\xi^3 + 0.876\xi^2 + 0.268\xi + 0.024 & -0.614\xi^3 - 0.614\xi^2 - 0.188\xi - 0.017 \end{matrix}\right.$$
$$\left.\begin{matrix} 3.34\xi^2 + 1.5864\xi + 1.865 & -1.228\xi^2 - 0.583\xi - 0.686 \\ -1.67\xi^3 - 1.447\xi^2 - 2.03\xi - 0.863 & 0.614\xi^3 + 0.532\xi^2 + 0.747\xi + 0.317 \\ 1.67\xi^2 + 1.392\xi + 1.28 & -0.614\xi^2 - 0.512\xi - 0.102 \\ -1.67\xi^3 - 1.67\xi^2 - 0.510\xi - 0.046 & 0.614\xi^3 + 0.614\xi^2 + 0.188\xi + 1.02 \end{matrix}\right]$$

[1] `Singular` is a free and open source computer algebra system for polynomial computations, see [19].

We generate (w, c) by simulating $\mathcal{P}_{full}$ in (4.9) in `Matlab`. To do the simulation, we find a non-anticipating input/output partition of $\mathcal{P}_{full}$ using Algorithm 1 as

$$P^{-1}Q = \begin{bmatrix} \frac{24\xi^2+14\xi+26}{24\xi^3+26\xi^2+9\xi+1} & \frac{-240\xi^2-124\xi-136}{120\xi^4+154\xi^3+71\xi^2+14\xi+1} \\ \frac{-12}{12\xi^2+7\xi+1} & \frac{60\xi^2+27\xi+63}{60\xi^3+47\xi^2+12\xi+1} \\ \frac{4}{4\xi+1} & \frac{-20}{20\xi^2+9\xi+1} \\ 0 & \frac{5}{5\xi+1} \end{bmatrix} \tag{4.11}$$

with c_1, c_2 as inputs and w_1, w_2, w_3, w_4 as outputs. Then we generate $(w, c) \in \mathcal{P}_{full}$ of length 50000 with c_1, c_2 a realization of *white Gaussian noise* process in order to guarantee persistency of excitation.

Let $\widetilde{QD}_1$ be the coefficient matrix of QD_1 with 4 block columns. Then w'_d is given by

$$\mathfrak{H}_1(w'_d) = \mathfrak{H}_1(w) - \widetilde{QD}_1\mathfrak{H}_1(w).$$

Choose $L = 100$, $J = 49900$ and define $H = \mathfrak{H}_{L,J}(w)$. Let $w_{pre} \in \mathcal{D}$ with $t_0 = 0, t_1 = 7$ be chosen as

$$w_{pre} = \begin{bmatrix} 0 & 0 & -0.3090 & -0.4256 & -0.7408 & -0.7841 & -0.8812 & -0.8157 \\ 0 & 0.1545 & 0.2733 & 0.5267 & 0.6490 & 0.7386 & 0.7449 & 0.6821 \\ 0 & 0 & -0.1545 & -0.2681 & -0.3598 & -0.4156 & -0.4307 & -0.4037 \\ 0 & 0.1545 & 0.2939 & 0.4045 & 0.4755 & 0.5000 & 0.4755 & 0.4045 \end{bmatrix}.$$

Define H_1 as the first 32 rows of H and solve a linear system of equations

$$H_1 \upsilon = w_{pre}$$

for υ. Then $w_d = \mathfrak{H}(w'_d)\upsilon$ where $\mathfrak{H}(w'_d) \in \mathbb{R}^{45000\times 49900}$.

Now to find c_d, the observability map O is computed by finding a left inverse of M_1 as

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Hence,

$$O := KR_1 = \begin{bmatrix} \xi + \frac{1}{2} & 1 & 0 & 1 \\ 0 & \xi + \frac{1}{3} & 1 & 0 \end{bmatrix}. \tag{4.12}$$

Let $\tilde{O}$ be the coefficient matrix of O with 2 block columns then $\mathfrak{H}_1(c_d) = \tilde{O}\mathfrak{H}_1(w_d)$ and $c_{d|[0,7]}$ is computed as

$$c_{d|[0,7]} = \begin{bmatrix} 0 & 0 & -0.0129 & -0.0223 & -0.0300 & -0.0346 & -0.0359 & -0.0336 \\ 0.1545 & 0.3248 & 0.4633 & 0.5564 & 0.5951 & 0.5755 & 0.4996 & 0.3748 \end{bmatrix}.$$

To verify that the computed c_d impose the prescribed path w_{pre}, we simulate (4.11) with c_d as the input and the output $w_{|[0,7]}$ is

$$w_{|[0,7]} = \begin{bmatrix} 0 & 0 & -0.3090 & -0.4256 & -0.7408 & -0.7841 & -0.8812 & -0.8157 \\ 0 & 0.1545 & 0.2733 & 0.5267 & 0.6490 & 0.7386 & 0.7449 & 0.6821 \\ 0 & 0 & -0.1545 & -0.2681 & -0.3598 & -0.4156 & -0.4307 & -0.4037 \\ 0 & 0.1545 & 0.2939 & 0.4045 & 0.4755 & 0.5000 & 0.4755 & 0.4045 \end{bmatrix},$$

which is exactly the same as w_{pre}.

4.5 Summary

We developed a solution for a special data-driven control problem where both system data and representation are used find control variable trajectories such that the to-be-controlled system follow a given prescribed path within a given time interval. The problem solved in this chapter, as already outlined, uses both data and representations; in the next chapter we shall address completely data-driven problems.

Chapter 5

Data-driven control: general case

In this chapter, we develop solutions for two data-driven problems. In the first problem, we consider given $(\tilde{w}, \tilde{c})$ and $\tilde{w}_d$ from the to-be-controlled system and the desired controlled system, respectively. The problem is to find a controller that implements the desired controlled system via partial interconnection. To find a solution we state conditions under which the given trajectories can be used to find a controller.

The second problem is the full interconnection case. Consider given $\tilde{w}$ and and $\tilde{w}_d$ of the to-be-controlled system and the desired controlled system, respectively. In this case, the problem is to find a controller that implement the desired controlled system via full interconnection. We shall demonstrate how one can find such a controller using only trajectories $\tilde{w}$ and $\tilde{w}_d$ under suitable conditions.

5.1 Problem statements

Let the to-be-controlled system full behaviour be $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$, the desired behaviour be $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}$ and finally, a to-be-designed controller behaviour be $\mathcal{C} \in \mathscr{L}^{\mathtt{c}}$. For the data-driven partial interconnection problem we consider a case where sufficiently informative (see Definition 2.42) $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ and $\tilde{w}_d \in \mathcal{D}$ are given. Furthermore, in finding a controller that implements $\mathcal{D}$, no representations of $\mathcal{P}_{full}$ and $\mathcal{D}$ are required nor will $(\tilde{w}, \tilde{c})$ and $\tilde{w}_d$ be used to find such representations. The data-driven partial interconnection problem is formally stated as follows.

Problem 5.1. Partial interconnection case. *Assume that*

1. *c is observable from w.*

Given sufficiently informative

i. *$(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ and*

ii. *$\tilde{w}_d \in \mathcal{D}$.*

Find, if it exists, a controller $\mathcal{C}$ that implements $\mathcal{D}$ via partial interconnection with $\mathcal{P}_{full}$.

To solve Problem 5.1 above, we state sufficient conditions under which $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ can be used to find $\tilde{c}_d \in \mathcal{C}$ corresponding to the given $\tilde{w}_d \in \mathcal{D}$. Under such conditions we then find $\tilde{c}_d \in \mathcal{C}$ and determine sufficient conditions under which $\tilde{c}_d$ is sufficiently informative about $\mathcal{C}$ such that standard procedures can be applied to find a representation of $\mathcal{C}$.

Now assume that there is no separation between the to-be-controlled variables and control variables, i.e. full interconnection. Let the to-be-controlled system behaviour be $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}$, the desired behaviour $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}$ and a to-be-designed controller behaviour be $\mathcal{C} \in \mathscr{L}^{\mathtt{w}}$. In common with the partial interconnection, in the full interconnection case we aim to use only particular trajectories from $\mathcal{P}$ and $\mathcal{D}$ to find $\mathcal{C}$ that implements $\mathcal{D}$. The data-driven full interconnection problem is formally stated as follows.

Problem 5.2. Full interconnection case. *Given*

i. *$\tilde{w} \in \mathcal{P}$,*

ii. *$\tilde{w}_d \in \mathcal{D}$.*

Find, if it exists, a controller $\mathcal{C}$ that implements $\mathcal{D}$ via full interconnection with $\mathcal{P}$.

Let $\mathfrak{N}_{\mathcal{C}}$ be the module of annihilators of $\mathcal{C}$. To solve Problem 5.2 above, we aim to use $\tilde{w}$ and $\tilde{w}_d$ to find a set of generators for $\mathfrak{N}_{\mathcal{C}}$ under suitable conditions.

5.2 Partial interconnection solution

In this section, we develop a solution to Problem 5.1. The solution is summarised in Algorithm 4 on p. 60. We start off by verifying that $\mathcal{D} \subseteq \mathcal{P}$ using $(\tilde{w}, \tilde{c})$ and $\tilde{w}_d$. Since $(\tilde{w}, \tilde{c})$ and $\tilde{w}_d$ are sufficiently informative about their respective behaviours then it follows from Definition 2.42 that the uniqueness of $\mathcal{P}_{full}$ and $\mathcal{D}$, is guaranteed, therefore, we use the given trajectories to verify that $\mathcal{D} \subseteq \mathcal{P}$, as we proceed to do in the next subsection.

5.2.1 Implementability

To verify that $\mathcal{D}$ is implementable using observed trajectories we cover some concepts of angle between behaviours, see section 3 pp. 199-201 of [59].

Define the map

$$\langle \cdot, \cdot \rangle : \mathbb{R}^{\mathtt{w}} \times \mathbb{R}^{\mathtt{w}} \to \mathbb{R}$$

by

$$\langle w, w' \rangle := w^\top w',$$

for $w, w' \in l_2^{\mathtt{w}}$.

Definition 5.3. The *angle*, Θ, between $w, w' \in l_2^{\mathtt{w}}$ is defined by

$$\Theta\left(w, w'\right) := \begin{cases} 0 & \text{if } w = 0 \text{ and } w' = 0, \\ \arccos\left(\dfrac{|\langle w, w' \rangle|}{\|w\| \|w'\|}\right) & \text{if } w \neq 0 \text{ and } w' \neq 0, \\ \dfrac{\pi}{2} & \text{if } \left(w = 0 \text{ and } w' \neq 0\right) \text{ or } \left(w \neq 0 \text{ and } w' = 0\right). \end{cases}$$

From Definition 5.3 if $\Theta(w, w') = \dfrac{\pi}{2}$ and both w, w' are non-zero, then $\langle w, w' \rangle = 0$. Hence, w, w' are said to be *orthogonal*. Define

$$w \perp w' :\Leftrightarrow \langle w, w' \rangle = 0,$$

we shall denote by $w \perp w'$ two orthogonal trajectories $w, w' \in l_2^{\mathtt{w}}$.

Definition 5.4. Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $w' \in l_2^{\mathtt{w}}$. The *angle* between w' *and* $\mathfrak{B}$ is defined by

$$\Theta\left(w', \mathfrak{B}\right) := \min_{w \in \mathfrak{B} \cap l_2^{\mathtt{w}}} \Theta\left(w', w\right).$$

From the above definition the angle between w' and $\mathfrak{B}$ can be computed by the *orthogonal projection* of w' onto $\mathfrak{B}$. Moreover, if $\Theta(w', \mathfrak{B}) = \dfrac{\pi}{2}$ and w' is non-zero then w' is orthogonal to all non-zero $w \in \mathfrak{B}$.

Now let $\mathfrak{B}, \mathfrak{B}' \in \mathscr{L}^{\mathtt{w}}_{contr}$, then the angle between $\mathfrak{B}$ and $\mathfrak{B}'$ is defined by

$$\Theta\left(\mathfrak{B}, \mathfrak{B}'\right) := \max\left\{\sup_{w \in \mathfrak{B} \cap l_2^{\mathtt{w}}} \Theta\left(w, \mathfrak{B}'\right), \sup_{w' \in \mathfrak{B}' \cap l_2^{\mathtt{w}}} \Theta\left(w', \mathfrak{B}\right)\right\} \tag{5.1}$$

Notice that angles $\Theta(w, \mathfrak{B}')$ and $\Theta(w', \mathfrak{B})$ are bounded between 0 and $\dfrac{\pi}{2}$, hence, $0 \leq \Theta(\mathfrak{B}, \mathfrak{B}') \leq \dfrac{\pi}{2}$. If $\Theta(\mathfrak{B}, \mathfrak{B}') = \dfrac{\pi}{2}$ then $\langle w, w' \rangle = 0$ for all $w \in \mathfrak{B} \cap l_2^{\mathtt{w}}$ and $w' \in \mathfrak{B}' \cap l_2^{\mathtt{w}}$. This leads us to the definition of the *orthogonal complement* of $\mathfrak{B}$ as follows,

$$\mathfrak{B}^{\perp} := \left\{ w' \in l_2^{\mathtt{w}} \mid w \perp w' \text{ for all } w \in \mathfrak{B} \cap l_2^{\mathtt{w}} \right\}. \tag{5.2}$$

We shall denote the orthogonal complement of $\mathfrak{B}$ by $\mathfrak{B}^{oc}$.

Proposition 5.5. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$. If $\mathfrak{B}^{oc}$ is the orthogonal complement of $\mathfrak{B}$ then*

1. $\Theta\left(\mathfrak{B}, \mathfrak{B}^{oc}\right) = \dfrac{\pi}{2}$,
2. $\mathfrak{B}^{oc} \in \mathscr{L}^{\mathtt{w}}_{contr}$,
3. $\mathscr{L}^{\mathtt{w}} = \mathfrak{B} \oplus \mathfrak{B}^{oc}$.

Proof. See Proposition 2.8 p. 196 of [59] and Lemma 3.2 p. 201 of [59]. □

Let $L \in \mathbb{N}$ greater than both $\mathtt{L}(\mathcal{P})$ and $\mathtt{L}(\mathcal{D})$. Let the Hankel matrices of depth L associated with $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$ be $\mathfrak{H}_L(\tilde{w})$ and $\mathfrak{H}_L(\tilde{w}_d)$, respectively. We now apply the concepts just introduced to the case of interest in this section, that is to determine if the desired controlled behaviour is implementable. Recall that if $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$ are sufficiently informative then $\text{col span}(\mathfrak{H}_L(\tilde{w})) = \mathcal{P}_{|[1,L]]}$ and $\text{col span}(\mathfrak{H}_L(\tilde{w}_d) = \mathcal{D}_{|[1,L]]}$. Therefore, to verify $\mathcal{D}_{|[1,L]} \subseteq \mathcal{P}_{|[1,L]}$ one can always show that $\text{col span}(\mathfrak{H}_L(\tilde{w}_d)) \subseteq \text{col span}(\mathfrak{H}_L(\tilde{w}))$.

Let $\tilde{w}', \tilde{w}'' \in \mathbb{R}^L$ and define the angle between $\tilde{w}', \tilde{w}''$ by

$$\theta\left(\tilde{w}', \tilde{w}''\right) := \begin{cases} 0 & \text{if } \tilde{w}' = 0 \text{ and } \tilde{w}'' = 0, \\ \arccos\left(\dfrac{|\langle \tilde{w}', \tilde{w}'' \rangle|}{\|\tilde{w}'\| \|\tilde{w}''\|}\right) & \text{if } \tilde{w}' \neq 0 \text{ and } \tilde{w}'' \neq 0, \\ \dfrac{\pi}{2} & \text{if } \left(\tilde{w}' = 0 \text{ and } \tilde{w}'' \neq 0\right) \text{ or } \left(\tilde{w}' \neq 0 \text{ and } \tilde{w}'' = 0\right). \end{cases}$$

Now define the angle between $\tilde{w}'$ and $\text{col span}(\mathfrak{H}_L(\tilde{w}''))$, where $\tilde{w}'' \in \mathbb{R}^{\mathbb{T}}$ with $T > L$, by

$$\theta\left(\tilde{w}', \text{col span}(\mathfrak{H}_L(\tilde{w}''))\right) := \min_{\tilde{w}_1 \in \text{col span}(\mathfrak{H}_L(\tilde{w}''))} \theta\left(\tilde{w}_1, \tilde{w}'\right).$$

Finally define the angle between $\text{col span}(\mathfrak{H}_L(\tilde{w}_d))$ and $\text{col span}(\mathfrak{H}_L(\tilde{w}))$ by

$$\theta\left(\text{col span}(\mathfrak{H}_L(\tilde{w})), \text{col span}(\mathfrak{H}_L(\tilde{w}_d))\right) :=$$
$$\max\left\{ \sup_{\tilde{w}' \in \text{col span}(\mathfrak{H}_L(\tilde{w}))} \theta\left(\tilde{w}', \text{col span}(\mathfrak{H}_L(\tilde{w}_d))\right), \sup_{\tilde{w}' \in \text{col span}(\mathfrak{H}_L(\tilde{w}_d))} \theta\left(\tilde{w}', \text{col span}(\mathfrak{H}_L(\tilde{w}))\right) \right\}$$

To verify $\mathcal{D} \subseteq \mathcal{P}$ using $\tilde{w}$ and $\tilde{w}_d$, we consider the following result. We denote by $\dim(\text{col span}(\mathfrak{H}_L(\tilde{w})))$ the dimensions of $\text{col span}(\mathfrak{H}_L(\tilde{w}))$.

Theorem 5.6. *Let $\tilde{w}$ and $\tilde{w}_d$ be sufficiently informative about $\mathcal{P}$ and $\mathcal{D}$, respectively. Assume dim(col span($\mathfrak{H}_L(\tilde{w}_d)$)) $\leq$ dim(col span($\mathfrak{H}_L(\tilde{w})$)). Then the following statements are equivalent,*

1. *col span($\mathfrak{H}_L(\tilde{w}_d)$) $\subseteq$ col span($\mathfrak{H}_L(\tilde{w})$), and*

2. $\theta\,(\text{col span}(\mathfrak{H}_L(\tilde{w})), \text{col span}(\mathfrak{H}_L(\tilde{w}_d))) = 0.$

Proof. Since $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$ are sufficiently informative then col span($\mathfrak{H}_L(\tilde{w})$) $=$ $\mathcal{P}_{|[1,L]]}$ and col span($\mathfrak{H}_L(\tilde{w}_d)$ $= \mathcal{D}_{|[1,L]]}$. Therefore, col span($\mathfrak{H}_L(\tilde{w}_d)$) $\subseteq$ col span($\mathfrak{H}_L(\tilde{w})$) implies that $\mathcal{D}_{|[1,L]} \subseteq \mathcal{P}_{|[1,L]}$.
(1) $\implies$ (2) Assume that col span($\mathfrak{H}_L(\tilde{w}_d)$) $\subseteq$ col span($\mathfrak{H}_L(\tilde{w})$). Then for all $\tilde{w}' \in$ col span($\mathfrak{H}_L(\tilde{w}_d)$) then $\tilde{w}' \in$ col span($\mathfrak{H}_L(\tilde{w})$), hence,

$$\theta\left(\tilde{w}', \text{col span}(\mathfrak{H}_L(\tilde{w}))\right) = 0 = \theta\left(\tilde{w}', \text{col span}(\mathfrak{H}_L(\tilde{w}_d))\right).$$

Therefore, $\theta\,(\text{col span}(\mathfrak{H}_L(\tilde{w})), \text{col span}(\mathfrak{H}_L(\tilde{w}_d))) = 0$.
(2) $\implies$ (1) Assume $\theta\,(\text{col span}(\mathfrak{H}_L(\tilde{w})), \text{col span}(\mathfrak{H}_L(\tilde{w}_d))) = 0$, from the definition of $\theta\,(\text{col span}(\mathfrak{H}_L(\tilde{w})), \text{col span}(\mathfrak{H}_L(\tilde{w}_d)))$ then

$$\sup_{\tilde{w}'' \in \text{col span}(\mathfrak{H}_L(\tilde{w}_d))} \theta\left(\tilde{w}'', \text{col span}(\mathfrak{H}_L(\tilde{w}))\right) = 0$$

and

$$\sup_{\tilde{w}' \in \text{col span}(\mathfrak{H}_L(\tilde{w}))} \theta\left(\tilde{w}', \text{col span}(\mathfrak{H}_L(\tilde{w}_d))\right) = 0.$$

Since dim(col span($\mathfrak{H}_L(\tilde{w}_d)$)) $\leq$ dim(col span($\mathfrak{H}_L(\tilde{w})$)), then for all $\tilde{w}'' \in$ col span($\mathfrak{H}_L(\tilde{w}_d)$) there exist $\tilde{w}' \in$ col span($\mathfrak{H}_L(\tilde{w})$) such that $\tilde{w}' = \tilde{w}''$. Therefore, col span($\mathfrak{H}_L(\tilde{w}_d)$) $\subseteq$ col span($\mathfrak{H}_L(\tilde{w})$) □

It follows from Theorem 5.6 that if $\theta\,(\text{col span}(\mathfrak{H}_L(\tilde{w})), \text{col span}(\mathfrak{H}_L(\tilde{w}_d))) = 0$ then the desired behaviour is implementable, therefore, we can find a solution to Problem 5.1. Otherwise, there is no solution to Problem 5.1.

Remark 5.7. For easy verification of col span($\mathfrak{H}_L(\tilde{w}')$) $\subseteq$ col span($\mathfrak{H}_L(\tilde{w})$) one can compute *principal angles* between subspaces, see Appendix A section A.3 □

5.2.2 Control variable trajectories

Let $L \in \mathbb{Z}_+$. Under the assumption that c is observable from w, a $L \times L$ real matrix O such that

$$\mathfrak{H}_L(\tilde{c}) = \mathsf{O}\mathfrak{H}_L(\tilde{w}), \tag{5.3}$$

for all $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$, is called an *empirical observability map.* In the following discussion we show how to compute O.

Let $\mathcal{D}$ be implementable. To find a control variable trajectory that corresponds to a given $\tilde{w}_d \in \mathcal{D}$, we find O using the given $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$. First in the following result, we prove necessary and sufficient conditions that given $w \in \mathcal{D}$ one can find c such that $(w, c) \in \mathcal{P}_{full}$ and c belongs a controller $\mathcal{C}$ that implements $\mathcal{D}$.

Theorem 5.8. *Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$, $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}$ and $\mathcal{C} \in \mathscr{L}^{\mathtt{c}}$. Assume that $R, D \in \mathbb{R}^{\bullet\times\mathtt{w}}[\xi]$, $M \in \mathbb{R}^{\bullet\times\mathtt{c}}[\xi]$ and $C \in \mathbb{R}^{\bullet\times\mathtt{c}}[\xi]$ are such that $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R(\sigma) & -M(\sigma)\end{bmatrix}\right)$, $\mathcal{D} = \ker(D(\sigma))$ and $\mathcal{C} = \ker(C(\sigma))$. Furthermore, assume that $\mathcal{C}$ implements $\mathcal{D}$ via partial interconnection through c with respect to $\mathcal{P}_{full}$ and that c is observable from w. Let $N \in \mathbb{R}^{\mathtt{c}\times\bullet}[\xi]$ be such that $NM = I_{\mathtt{c}}$. The following statements are equivalent:*

1. *there exists $G \in \mathbb{R}^{\mathtt{c}\times\bullet}[\xi]$ right prime such that $Y := NR+GD$ induces a polynomial operator in the shift $Y(\sigma)$ such that $c = Y(\sigma)w$ for all $(w, c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$,*
2. *there exists $F \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ such that $NM = I_{\mathtt{c}} + FC$.*

Proof. The existence of N follows form the assumption that c is observable from w, consequently, M admits a left inverse. Furthermore, let $(w, c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$ then under the observability assumption it follows from Proposition 4.10 that if $\mathcal{C}$ implements $\mathcal{D}$ then $c \in \mathcal{C}$.
2) $\implies$ 1) Let $c \in \mathcal{C}$, assume that F exists such that $NM = I_{\mathtt{c}} + FC$, then

$$(NM)(\sigma)c = (I_{\mathtt{c}} + FC)(\sigma)c.$$

By the assumption that $\mathcal{C} = \ker(C(\sigma))$ then $(FC)(\sigma)c = 0$. Hence,

$$(NM)(\sigma)c = c.$$

Now, by the assumption that $\mathcal{C}$ implements $\mathcal{D}$ then there exists $w \in \mathcal{D}$ such that $(w, c) \in \mathcal{P}_{full}$, furthermore, $R(\sigma)w = M(\sigma)c$. Hence,

$$c = (NR)(\sigma)w + 0,$$

since $D(\sigma)w = 0$ this implies that G exists such that

$$c = (NR)(\sigma)w + (GD)(\sigma)w.$$

Therefore, $c = Y(\sigma)w$. G being left prime follows from; for $w \notin \mathcal{D}$ then $D(\sigma)w \neq 0$, therefore, $(GD)(\sigma)w \neq 0$.
To prove 1) $\implies$ 2), let $(w, c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$. By the assumptions that c is

observable from w and Y induces an observability map, it follows that $c = (NR)(\sigma)w + (GD)(\sigma)w$. Since $w \in \mathcal{D}$ then $D(\sigma)w = 0$. Hence, $c = (NR)(\sigma)w$. Now since $(w, c) \in \mathcal{P}_{full}$ then $R(\sigma)w = M(\sigma)c$. It follows that

$$c = (NR)(\sigma)w = (NM)(\sigma)c,$$

hence, $c = (NM)(\sigma)c$. Consequently,

$$(NM - I)(\sigma)c = 0.$$

Now recall that the controller $\mathcal{C} = \ker(C(\sigma))$ implements $\mathcal{D}$ and that $(w, c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$ then $c \in \mathcal{C}$, therefore, $C(\sigma)c = 0$. Since $(NM - I)(\sigma)c = 0$ and $C(\sigma)c = 0$, then F exists such that $NM - I = FC$. □

We now show how to find an empirical observability map $\mathsf{O} \in \mathbb{R}^{\bullet\times\bullet}$ using $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$.

Lemma 5.9. *Let* $\mathrm{col}(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ *be sufficiently informative about* $\mathcal{P}_{full}$. *Let* $Y \in \mathbb{R}^{\mathtt{c}\times\mathtt{w}}[\xi]$ *satisfy the conditions of Theorem 5.8 and* $L \in \mathbb{Z}_+$ *satisfy* $L > \mathtt{L}(\mathcal{P})$, $L > \mathtt{L}(\mathcal{D})$ *and* $L \gg deg(Y)$. *Denote by* $\mathfrak{H}_L(\tilde{c})$, $\mathfrak{H}_L(\tilde{w})$ *the Hankel matrices associated with* $\tilde{w}$ *and* $\tilde{c}$, *respectively, both with* L *block rows and an infinite number of columns. Then a solution for* $\mathsf{O} \in \mathbb{R}^{L\times L}$ *in*

$$\mathfrak{H}_L(\tilde{c}) = \mathsf{O}\mathfrak{H}_L(\tilde{w}) \tag{5.4}$$

induces an empirical observability map.

Proof. Denote by $\tilde{Y}$ the coefficient matrix of Y with finite number L of block-columns. Under the assumption that Y induces an observability map, then $\bar{\mathsf{O}} := \mathrm{col}(\sigma_R^k\tilde{Y})_{k=0,\dots,L-1}$ is a solution of (5.4), therefore, $\mathfrak{H}_L(\tilde{c}) = \bar{\mathsf{O}}\mathfrak{H}_L(\tilde{w})$. Now since $\mathrm{col}(\tilde{w}, \tilde{c})$ is sufficiently informative then $\mathrm{leftkernel}(\mathfrak{H}_L(\tilde{w})) \neq 0$. Therefore, (5.4) has infinitely many solutions. Let $\mathsf{K} \in \mathbb{R}^{L\times\bullet}$ be a matrix whose columns are a basis of $\mathrm{leftkernel}(\mathfrak{H}_L(\tilde{w}))$. Then the set of all solutions of (5.4) is defined by

$$\mathcal{G} := \left\{\bar{\mathsf{O}} + \mathsf{K}\mathsf{T} | \mathsf{T} \in \mathbb{R}^{\bullet\times L}\right\}.$$

Let $\mathsf{O} \in \mathcal{G}$, then $\mathsf{O} = \bar{\mathsf{O}} + \mathsf{K}\mathsf{T}$ for some K and T. Compute

$$\begin{aligned}\mathsf{O}\mathfrak{H}_L(\tilde{w}) &= (\bar{\mathsf{O}} + \mathsf{K}\mathsf{T})\mathfrak{H}_L(\tilde{w}) \\ &= \bar{\mathsf{O}}\mathfrak{H}_L(\tilde{w}) + \mathsf{K}\mathsf{T}\mathfrak{H}_L(\tilde{w}).\end{aligned}$$

Notice that $\mathsf{K}\mathsf{T}\mathfrak{H}_L(\tilde{w}) = 0$. Therefore,

$$\begin{aligned}\mathsf{O}\mathfrak{H}_L(\tilde{w}) &= \bar{\mathsf{O}}\mathfrak{H}_L(\tilde{w}) + 0 \\ &= \mathfrak{H}_L(\tilde{c}).\end{aligned}$$

Hence, O induce an empirical observability map. □

It follows from Lemma 5.9 that the Hankel matrix associated with the control variable trajectory $\tilde{c}_d$ corresponding to the $\tilde{w}_d \in \mathcal{D}$ is given by

$$\mathfrak{H}_L(\tilde{c}_d) = \mathsf{O}\mathfrak{H}_L(\tilde{w}_d). \tag{5.5}$$

To show that $\tilde{c}_d$ belongs a controller that implement $\mathcal{D}$ first we state the following straightforward result.

Lemma 5.10. *Let $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$ be sufficiently informative about their respective behaviours. Let $L > \mathtt{L}(\mathcal{P})$ and $L > \mathtt{L}(\mathcal{D})$. Then leftkernel($\mathfrak{H}_L(\tilde{w})$) $\subseteq$ leftkernel($\mathfrak{H}_L(\tilde{w}_d)$) if and only if $\mathcal{D} \subseteq \mathcal{P}$.*

Proof. Let $\mathfrak{N}_\mathcal{P}$ and $\mathfrak{N}_\mathcal{D}$ be the module of annihilators of $\mathcal{P}$ and $\mathcal{D}$, respectively. By the assumption that $\tilde{w}$ are $\tilde{w}_d$ are sufficiently informative about their respective behaviours then leftkernel$(\mathfrak{H}_L(\tilde{w})) = \tilde{\mathfrak{N}}^L_\mathcal{P}$ and leftkernel$(\mathfrak{H}_L(\tilde{w}_d)) = \tilde{\mathfrak{N}}^L_\mathcal{D}$. Therefore, leftkernel$(\mathfrak{H}_L(\tilde{w})) \subseteq$ leftkernel$(\mathfrak{H}_L(\tilde{w}_d))$ which implies that $\tilde{\mathfrak{N}}^L_\mathcal{P} \subseteq \tilde{\mathfrak{N}}^L_\mathcal{D}$. Hence, it is suffice to prove that $\mathfrak{N}_\mathcal{P} \subseteq \mathfrak{N}_\mathcal{D}$ if and only if $\mathcal{D} \subseteq \mathcal{P}$. The rest of the proof follow from the proof of Proposition 2.24. □

Proposition 5.11. *Let* $\mathrm{col}(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ *and $\tilde{w}_d \in \mathcal{D}$ be sufficiently informative about their respective behaviours. Let O satisfying the conditions of Lemma 5.9. Assume that $\mathcal{D} \subseteq \mathcal{P}$ and that c is observable from w. If $\mathcal{C}$ implements $\mathcal{D}$ then the control variable trajectory $\tilde{c}_d$ in* (5.5) *belongs to $\mathcal{C}$.*

Proof. Let $\bar{\mathsf{O}}$ and K as in Lemma 5.9 and Y satisfying the conditions of Theorem 5.8. Now since O satisfies conditions of Lemma 5.9 then $\mathsf{O} = \bar{\mathsf{O}} + \mathsf{KT}$ where $\mathsf{T} \in \mathbb{R}^{\bullet \times L}$. Compute

$$\mathsf{O}\mathfrak{H}_L(\tilde{w}_d) = \bar{\mathsf{O}}\mathfrak{H}_L(\tilde{w}_d) + \mathsf{KT}\mathfrak{H}_L(\tilde{w}_d),$$

since $\mathcal{D} \subseteq \mathcal{P}$ then leftkernel$(\mathfrak{H}_L(\tilde{w})) \subseteq$ leftkernel$(\mathfrak{H}_L(\tilde{w}_d))$ see Lemma 5.10. Therefore, $\mathsf{KT} \in$ leftkernel$(\mathfrak{H}_L(\tilde{w}_d))$. Consequently,

$$\begin{aligned}\mathsf{O}\mathfrak{H}_L(\tilde{w}_d) &= \bar{\mathsf{O}}\mathfrak{H}_L(\tilde{w}_d) + 0\\ &= \mathfrak{H}_L(\tilde{c}_d).\end{aligned}$$

Under the observability assumption, since $\bar{\mathsf{O}}$ is defined by $\bar{\mathsf{O}} := \mathrm{col}(\sigma_R^k \tilde{Y})_{k=0,\dots,L-1}$, where Y induces an observability map and $\tilde{w}_d \in \mathcal{D}$, then it follow from Proposition 4.10 that $\tilde{c}_d$ belongs $\mathcal{C}$. □

5.2.3 Controller identification

In the following result we prove sufficient conditions for $\tilde{c}_d$ to be sufficiently informative about $\mathcal{C}$. Let $\Pi \in \mathbb{R}^{\mathtt{c}\times\mathtt{c}}$,$\Pi_1 \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ be such that $(\tilde{c}_u, \tilde{c}_y) = \Pi\tilde{c}_d$ and $(\tilde{w}_u, \tilde{w}_y) = \Pi_1\tilde{w}_d$ where $\tilde{c}_u$ and $\tilde{w}_u$ are inputs. Define a partition of Π by $\Pi =: \mathrm{col}(\Pi_u, \Pi_y)$, compatibly with partitions of $\tilde{c}_d = (\tilde{c}_u, \tilde{c}_y)$. Now let Y satisfying the conditions of Theorem 5.8 and define $Y_u := \Pi_u Y$. Denote by $< Y_u >$ the $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times\bullet}[\xi]$ generated by the rows of Y_u and by $\mathfrak{N}_\mathcal{D}$ the module of annihilators of $\mathcal{D}$. Finally, let $\tilde{Y}$ be the coefficient of matrix of Y with finite number L block-columns where $L > \mathtt{L}(\mathcal{C})$.

Theorem 5.12. *Let a controller $\mathcal{C}$ implements $\mathcal{D}$. Assume that c is observable from w and let O in (5.4) induces an empirical observability map. Now assume $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$, and let $\tilde{w}_d \in \mathcal{D}$ and $\tilde{c}_d \in \mathcal{C}$ whose Hankel matrix is defined in (5.5). If $< Y_u > \cap\, \mathfrak{N}_\mathcal{D} = \{0\}$ and $\tilde{w}_u$ is persistently exciting of order at least $\mathtt{L}(\mathcal{D})+\mathtt{n}(\mathcal{D})$ then $\tilde{c}_u$ is persistently exciting of order at least $\mathtt{L}(\mathcal{C}) + \mathtt{n}(\mathcal{C})$.*

Proof. Since O induce an empirical observability map and $\tilde{c}_d \in \mathcal{C}$ then from Lemmas 5.9 and 5.11

$$\mathfrak{H}_L(\tilde{c}_d) = \mathsf{O}\mathfrak{H}_L(\tilde{w}_d) = \bar{\mathsf{O}}\mathfrak{H}_L(\tilde{w}_d) + 0$$

where $\bar{\mathsf{O}} := \mathrm{col}(\sigma_R^k \tilde{Y})_{k=0,\dots,L-1}$. Let $J \in \mathbb{Z}_+$ and define $\mathsf{O}_u \in \mathbb{R}^{J\mathtt{m}(\mathcal{C})\times L}$ by $\mathsf{O}_u := \mathrm{col}(\sigma_R^k \tilde{Y}_u)_{k=0,\dots,L-1}$, then $\mathfrak{H}_{J\mathtt{m}(\mathcal{C})}(\tilde{c}_u) = \mathsf{O}_u\mathfrak{H}_L(\tilde{w}_d)$. Now, assume to the contrary that $\tilde{c}_u$ is not persistently exciting, then there exists $\tilde{\alpha} \in \mathbb{R}^{1\times J\mathtt{m}(\mathcal{C})}$ such that $\tilde{\alpha}\mathfrak{H}_{J\mathtt{m}(\mathcal{C})}(\tilde{c}_u) = 0$. Therefore, $\tilde{\alpha}\mathsf{O}_u \in \mathrm{leftkernel}(\mathfrak{H}_L(\tilde{w}_d))$. Now since $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $\tilde{w}_u$ is persistently exciting then $\mathrm{leftkernel}(\mathfrak{H}_L(\tilde{w}_d)) = \tilde{\mathfrak{N}}^L_\mathcal{D}$, hence, $\tilde{\alpha}\mathsf{O}_u \in \tilde{\mathfrak{N}}^L_\mathcal{D}$. Let $\alpha \in \mathbb{R}^{1\times\bullet}[\xi]$ be the polynomial vector whose coefficient matrix is $\tilde{\alpha}$ then $\alpha Y_u \in \mathfrak{N}_\mathcal{D}$. Since c_u is the input variable then it is free, therefore, $\mathrm{im}(Y_u(\sigma)) = (\mathbb{R}^{\mathtt{m}(\mathcal{C})})^\mathbb{Z}$, hence, Y_u is full row rank, which implies that $\alpha Y_u \neq 0$. Consequently, $\alpha Y_u \in \mathfrak{N}_\mathcal{D}$ and $\alpha Y_u \neq 0$, hence, a contradiction. □

Remark 5.13. Note that it is not straightforward to verify the assumption of Theorem 5.12 from data. Therefore verifying that $\tilde{c}_u$ is persistently exciting can be done by determining which rows of $\mathfrak{H}_L(\tilde{c}_d)$ corresponds to the input variables (see steps (1)-(3) of Algorithm 2 on p. 28). Let $\mathfrak{H}_{J\mathtt{m}(\mathcal{C})}(\tilde{c}_u)$ be the rows of $\mathfrak{H}_L(\tilde{c}_d)$ corresponding to the input variables, if $\mathfrak{H}_{J\mathtt{m}(\mathcal{C})}(\tilde{c}_u)$ is full row rank then $\tilde{c}_u$ is persistently exciting and vice versa. □

5.2.4 Algorithm

The partial interconnection solution is summarized in the algorithm below.

Algorithm 4: Solution of Problem 5.1

Input : $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ and $\tilde{w}_d \in \mathcal{D}$.
Output : A representation of $\mathcal{C}$.
Assumptions: Theorem 5.12.

Choose L to be sufficiently large (see Remark 4.7).
Build the Hankel matrices: $\mathfrak{H}_L(\tilde{w}), \mathfrak{H}_L(\tilde{c}), \mathfrak{H}_L(\tilde{w}_d)$.
Verify $\mathcal{D} \subseteq \mathcal{P}$ (see subsection 5.2.1)
if $\mathcal{D} \subseteq \mathcal{P}$ **then**
  Solve $\mathfrak{H}_L(\tilde{c}) = \mathsf{O}\mathfrak{H}_L(\tilde{w})$ for O;
  Compute $\mathfrak{H}_L(\tilde{c}_d) = \mathsf{O}\mathfrak{H}_L(\tilde{w}_d)$.
else
  $\mathcal{D} \nsubseteq \mathcal{P}$ {No solution for $\tilde{c}_d$}.
Using $\tilde{c}_d$ and Algorithm 2 compute a representation for $\mathcal{C}$.

5.2.5 Example

Consider a system with a representation in (4.9) and a transfer function in (4.11) and the desired behaviour induced by (4.10) with a transfer function

$$P_2^{-1}Q_2 = \begin{bmatrix} \dfrac{360\xi^2 + 171\xi + 201}{180\xi^3 + 180\xi^2 + 55z + 5} \\ \dfrac{-90\xi^2 - 33\xi - 93}{90\xi^2 + 45\xi + 5} \\ \dfrac{6}{6\xi + 1} \end{bmatrix}. \tag{5.6}$$

Generate $(\tilde{w}, \tilde{c})$ and $\tilde{w}_d$ both of length $T = 50000$ (see Remark 4.2) by simulation of (4.11) and (5.6) in `Matlab`, with inputs (c_1, c_2 and w_4 in (4.9) and (4.10), respectively) a realization of white Gaussian noise process to guarantee persistency of excitation.

Choose $L = 100$ and build the Hankel matrices $\mathfrak{H}_L(\tilde{w}), \mathfrak{H}_L(\tilde{c})$ and $\mathfrak{H}_L(\tilde{w}_d)$. Compute the largest principal angle between the col span($\mathfrak{H}_L(\tilde{w})$) and col span($\mathfrak{H}_L(\tilde{w}_d)$). Which is 1.2363×10^{-14} and is approximately zero, therefore $\mathcal{D} \subseteq \mathcal{P}$. Solve $\mathfrak{H}_L(\tilde{c}) = \mathsf{O}\mathfrak{H}_L(\tilde{w})$ for O and compute $\mathfrak{H}_L(\tilde{c}_d) = \mathsf{O}\mathfrak{H}_L(\tilde{w}_d)$. Under the assumption of the Algorithm 4, $\tilde{c}_d$ is sufficiently informative and can be used to find representations of $\mathcal{C}$.

Find $l \in \mathbb{N}$ the lag of $\mathcal{C}$ as follows. Compute

$$\begin{aligned}\rho_t &= \text{rank}(\mathfrak{H}_t(\tilde{c}_d)) - \text{rank}(\mathfrak{H}_{t-1}(\tilde{c}_d)) \\ \rho_0 &= \text{rank}(\mathfrak{H}_0(\tilde{c}_d))\end{aligned}$$

with $t \in \mathbb{Z}_+$. Carry out the computation until ρ_t is a non-increasing sequence of non-negative integers. Then find t' such that $\rho_t = \rho_{t'}$ for $t \geq t'$ then the lag is t'. For this example $l = 2$.

To find a representation of $\mathcal{C}$, build $\mathfrak{H}_{l+1}(\tilde{c}_d)$. Then compute the *singular value decomposition* (SVD) of $\mathfrak{H}_{l+1}(\tilde{c}_d) := U\Sigma V^\top$. Let r be the rank of $\mathfrak{H}_{l+1}(\tilde{c}_d)$. Partition U into $\begin{bmatrix} U_1 & U_2 \end{bmatrix}$ where U_1 has r columns then $U_2^\top$ is the left kernel of $\mathfrak{H}_{l+1}(\tilde{c})$, and we obtain a kernel representation of $\mathcal{C}$ as

$$\underbrace{\begin{bmatrix} -0.9356\sigma^2 - 0.3430\sigma - 0.0312 & -0.0780 \end{bmatrix}}_{C_1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0. \tag{5.7}$$

For comparison, we use polynomial operations to compute a controller representation from (4.9) and (4.10). This is done by computing the syzygy of $\text{col}(R_1, D_1')$ where $D_1' = [0\ 0\ \xi + \frac{1}{6}\ 1]$. We obtain

$$\underbrace{\begin{bmatrix} -\sigma^2 - 0.3667\sigma - 0.0333 & -0.0833 \end{bmatrix}}_{C_2} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0. \tag{5.8}$$

Equations (5.7) and (5.8) represents the same behaviour since there exists a unimodular matrix $U \in \mathbb{R}^{\bullet \times \bullet}[\xi]$ such that $C_2 = UC_1$, see Proposition 2.22. In this case $U = [1.0688]$.

5.2.6 Prescribed path problem revisited

In this section, we revisit the prescribed path problem covered in Chapter 4. We consider a completely data-driven version of Problem 4.1 stated as follows.

Problem 5.14. "Complete data-driven prescribed path" case. *Assume that*

1. c is observable from w.

Given

i. sufficiently informative infinite $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ and $\tilde{w}' \in \mathcal{D}$;

ii. a prescribed trajectory $\tilde{w}_{pre} \in \mathcal{D}_{|[t_0,t_1]}$ with $t_0, t_1 \in \mathbb{N}$, $t_0 \leqslant t_1$.

Find, if it exists, a control variable trajectory $\tilde{c}_d$, such that there exists $\tilde{w}_d$ such that

a. $(\tilde{w}_d, \tilde{c}_d) \in \mathcal{P}_{full}$

b. $\tilde{w}_{d_{|[t_0,t_1]}} = \tilde{w}_{pre}$.

We present a solution to Problem 5.14 in Algorithm 5 since all the necessary conditions have been covered in Chapter 4 and this chapter.

Algorithm 5: Solution for Problem 5.14

Input : $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}, \tilde{w}' \in \mathcal{D}$, t_0, t_1 and $\tilde{w}_{pre}$
Output : $\tilde{c}_d$
Assumptions: Theorem 4.9, Lemma 5.9

```
1  Choose L to be sufficiently large (see Remark 4.7) and J ≫ L.
2  Build the Hankel matrices: 𝔋_L(w̃), 𝔋_L(c̃), 𝔋_{L,J}(w̃').
3  Verify 𝒟 ⊆ 𝒫 (see subsection 5.2.1). If 𝒟 ⊈ 𝒫, stop. Otherwise go to step 4.
4  Define H_1 as a partition of rows of 𝔋_L(w̃') from row wt_0 to row wt_1
5  Solve H_1 υ = w̃_pre for υ.
6  if no solution for υ then
7      w_pre ∉ 𝒟_|[t_0,t_1] [No Solution for c_d]. Stop.
8  else
9      Build 𝔋(w̃') ∈ ℝ^{∞×J};
10     Define w̃_d := 𝔋(w̃')υ;
11     Solve 𝔋_L(c̃) = O𝔋_L(w̃) for O;
12     Compute 𝔋_L(c̃_d) = O𝔋_L(w̃_d).
```

In the following result we prove the correctness of Algorithm 5.

Proposition 5.15. *Let $(\tilde{w}, \tilde{c}) \in \mathcal{P}_{full}$ and $\tilde{w}' \in \mathcal{D}$ be sufficiently informative about their respective behaviours. Assume $\mathcal{D} \subseteq \mathcal{P}$ and that O satisfy the conditions of Lemma 5.9. Under the observability assumption if a controller $\mathcal{C}$ implements $\mathcal{D}$ then the control variable trajectory $\tilde{c}_d$ in Algorithm 5 belongs to $\mathcal{C}$. Moreover, $\tilde{c}_d$ imposes $\tilde{w}_{pre}$ on $\tilde{w}_d$ between the time interval t_0 and t_1.*

Proof. The fact that $\tilde{c}_d \in \mathcal{C}$ follows from the proof of Proposition 5.11, and $\tilde{c}_d$ imposing $\tilde{w}_{pre}$ on $\tilde{w}_d$ between the time interval t_0, t_1 follows from the proof of Lemma 4.11 □

Remark 5.16. Note that $\tilde{c}_d$ in Algorithm 5 cannot be used to find a representation for $\mathcal{C}$. This is because there are no conditions under which $\tilde{w}_d$ such that $\tilde{w}_{d_{|[t_0,t_1]}} = \tilde{w}_{pre}$ is sufficiently informative about $\mathcal{D}$. □

Remark 5.17. In the case where t_0, t_1 are taken to be $t_0 = 0$ and t_1 sufficiently large then Algorithm 5 can be suitably adjusted that the result is a tracking algorithm, whereby we find $\tilde{c}_d$ such the to-be-controlled system tracks a given trajectory $\tilde{w}_{pre}$. □

Remark 5.18. In our solutions to the partial interconnection problems, i.e. Problems 4.1 and 5.1, we assume that the control variables are observable from the to-be-controlled variables. This assumption allows for the computation of control viariable trajectories from to-be-controllable trajectories. If we consider the case where such an assumption is dropped then we would have to develop a procedure for computing control variable trajectories from the to-be-controlled variable trajectories. The development of a procedure for computing control variable trajectories from the to-be-controlled variable trajectories , in the case of c not observable from w, is still an open research direction, see [74] for the design of estimators in the behavioural setting. □

5.3 Full interconnection solution

In this section, we develop a solution to Problem 5.2. We follow the same procedure described in subsection 5.2.1 to verify that $\mathcal{D} \subseteq \mathcal{P}$ using $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$.

5.3.1 Controller identification

Let $\mathfrak{N}_\mathcal{P}$,$\mathfrak{N}_\mathcal{C}$ and $\mathfrak{N}_\mathcal{D}$ be the modules of annihilators of $\mathcal{P}, \mathcal{C}$ and $\mathcal{D}$, respectively. To find a set of generators of $\mathfrak{N}_\mathcal{C}$ using $\tilde{w}$ and $\tilde{w}_d$, consequently, finding $\mathcal{C}$ we need the necessary and sufficient conditions in Theorem 5.20. First, we state the following results.

Lemma 5.19. *Let $r_1, \ldots, r_\mathtt{p}$ and $c_1, \ldots, c_\mathtt{c}$ be generators of $\mathfrak{N}_\mathcal{P}$ and $\mathfrak{N}_\mathcal{C}$, respectively. Then $r_1, \ldots, r_\mathtt{g}, c_1, \ldots, c_\mathtt{c}$ are a bases generators of $\mathfrak{N}_\mathcal{D}$ if and only if $\mathfrak{N}_\mathcal{P} \cap \mathfrak{N}_\mathcal{C} = \{0\}$ and $r_1, \ldots, r_\mathtt{p}$ and $c_1, \ldots, c_\mathtt{c}$ are bases generators of their respective module of annihilators.*

Proof. Follows the same arguments of the proof of Proposition 3.7 □

Theorem 5.20. *Assume that full row rank $C \in \mathbb{R}^{\mathtt{c}\times\mathtt{w}}[\xi]$, $R \in \mathbb{R}^{\mathtt{p}\times\mathtt{w}}[\xi]$ and $D \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ are such that $\mathcal{C} = \ker(C(\sigma))$, $\mathcal{P} = \ker(R(\sigma))$ and $\mathcal{D} = \ker(D(\sigma))$. Furthermore, assume that $\mathfrak{N}_\mathcal{P} \cap \mathfrak{N}_\mathcal{C} = \{0\}$. Then $\mathcal{C}$ implements $\mathcal{D}$ via full interconnection if and only if $\mathfrak{N}_\mathcal{P} + \mathfrak{N}_\mathcal{C} = \mathfrak{N}_\mathcal{D}$.*

Proof. Write $C = \text{col}(c_1, \dots, c_{\mathtt{c}})$ and $R = \text{col}(r_1, \dots, r_{\mathtt{p}})$ and $D = \text{col}(d_1, \dots, d_{\mathtt{g}})$. Since $\mathcal{C} = \ker(C(\sigma))$, $\mathcal{P} = \ker(R(\sigma))$ and $\mathcal{D} = \ker(D(\sigma))$ then $c_1, \dots, c_{\mathtt{c}}$, $r_1, \dots, r_{\mathtt{p}}$ and $d_1, \dots, d_{\mathtt{g}}$ are generators of $\mathfrak{N}_{\mathcal{C}}$,$\mathfrak{N}_{\mathcal{P}}$ and $\mathfrak{N}_{\mathcal{D}}$, respectively.

(Only if) Assume that $\mathcal{C}$ implements $\mathcal{D}$, then $\mathcal{C} \cap \mathcal{P} = \ker\left(\begin{bmatrix} C(\sigma) \\ R(\sigma) \end{bmatrix}\right)$. Now since C and R are full row rank and $\mathfrak{N}_{\mathcal{P}} \cap \mathfrak{N}_{\mathcal{C}} = \{0\}$ then $\text{col}(C, R)$ induces a minimal representation. Since $\mathcal{D} = \ker(D(\sigma))$ then

$$\ker(D(\sigma)) = \ker\left(\begin{bmatrix} C(\sigma) \\ R(\sigma) \end{bmatrix}\right).$$

Since D is full row rank and $\text{col}(C, R)$ induces a minimal representation then it follows from Proposition 2.22 that there exists a unimodular matrix $F \in \mathbb{R}^{\mathtt{g} \times \mathtt{g}}[\xi]$, where $\mathtt{c} + \mathtt{p} = \mathtt{g}$, such that

$$D = F \begin{bmatrix} C \\ R \end{bmatrix}.$$

To show the inclusion $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} \supseteq \mathfrak{N}_{\mathcal{D}}$, partition the columns of $F = \begin{bmatrix} F_1 & F_2 \end{bmatrix}$ accordingly with respect to the rows of C and R. Then $D = F_1 C + F_2 R$ implies that we can write $d_1, \dots, d_{\mathtt{g}}$ as a linear combination of $c_2, \dots, c_{\mathtt{c}}$ and $r_1, \dots, r_{\mathtt{p}}$. Therefore, $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} \supseteq \mathfrak{N}_{\mathcal{D}}$. To show the converse inclusion, notice that since F is unimodular then it admits an inverse, i.e. $F^{-1}F = I_{\mathtt{g}}$. Therefore,

$$F^{-1}D = \begin{bmatrix} C \\ R \end{bmatrix}$$

which implies that $c_2, \dots, c_{\mathtt{c}}$ and $r_1, \dots, r_{\mathtt{p}}$ can be written as a linear combination of $d_1, \dots, d_{\mathtt{g}}$. Hence, $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} \subseteq \mathfrak{N}_{\mathcal{D}}$, therefore, $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} = \mathfrak{N}_{\mathcal{D}}$,

(If) Assume that $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} = \mathfrak{N}_{\mathcal{D}}$, then for $i = 1, \dots, \mathtt{g}$ there exist $m_{1_i}, \dots, m_{\mathtt{c}_i}, m'_{1_i}, \dots, m'_{\mathtt{p}_i} \in \mathbb{R}[\xi]$ such that $d_i = m_{1_i} c_1 + \dots + m_{\mathtt{c}_i} c_{\mathtt{c}} + m'_{1_i} r_1 + \dots + m'_{\mathtt{p}_i} r_{\mathtt{p}}$. Define $f_{1_i} := [m_{1_i} \ \dots \ m_{\mathtt{c}_i}]$, $f_{2_i} := [m'_{1_i} \ \dots \ m'_{\mathtt{p}_i}]$, $F_1 := \text{col}(f_{1_1}, \dots, f_{1_{\mathtt{c}}})$ and $F_2 := \text{col}(f_{2_1}, \dots, f_{2_{\mathtt{p}}})$. Recall that $C = \text{col}(c_1, \dots, c_{\mathtt{c}})$, $R = \text{col}(r_1, \dots, r_{\mathtt{p}})$ and $D = \text{col}(d_1, \dots, d_{\mathtt{g}})$, therefore, $F_1 C + F_2 R = D$, consequently,

$$D = \begin{bmatrix} F_1 & F_2 \end{bmatrix} \begin{bmatrix} C \\ R \end{bmatrix}$$

with $\begin{bmatrix} F_1 & F_2 \end{bmatrix} \in \mathbb{R}^{\mathtt{g} \times (\mathtt{c}+\mathtt{p})}[\xi]$. Similarly, since $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} = \mathfrak{N}_{\mathcal{D}}$, then there exist $m_{1_i}, \dots, m_{\mathtt{g}_i} \in \mathbb{R}[\xi]$ for $i = 1, \dots, (\mathtt{c} + \mathtt{p})$ such that $c_j = m_{1_i} d_1 + \dots + m_{\mathtt{g}_i} d_{\mathtt{g}}$ for $j = 1, \dots, \mathtt{c}$ and $r_{j'} = m_{1_i} d_1 + \dots + m_{\mathtt{g}_i} d_{\mathtt{g}}$ for $j' = 1, \dots, \mathtt{p}$. Define $f_i := [m_{1_i} \ \dots \ m_{\mathtt{g}_i}]$, and $F' := \text{col}(f_1, \dots, f_{(\mathtt{c}+\mathtt{p})})$. Then

$$F'D = \begin{bmatrix} C \\ R \end{bmatrix}.$$

Now, since D, R, C are of full row rank, and $\mathfrak{N}_{\mathcal{P}} \cap \mathfrak{N}_{\mathcal{C}} = \{0\}$ then $\mathtt{c} + \mathtt{p} = \mathtt{g}$, and matrices

F' and $[F_1\ F_2]$ are unimodular. Consequently, $\ker(D(\sigma)) = \ker(\begin{bmatrix} C(\sigma) \\ R(\sigma) \end{bmatrix})$ which implies that $\mathcal{D} = \mathcal{P} \cap \mathcal{C}$. □

To find a solution to Problem 5.2, under assumption of Theorem 5.20, we find bases generators $r_1, \dots r_{\mathtt{p}}$ and $a_1, \dots a_{\mathtt{g}}$ of $\mathfrak{N}_{\mathcal{P}}$ and $\mathfrak{N}_{\mathcal{D}}$, respectively, such that $(\deg(r_1), \dots \deg(r_{\mathtt{p}}))$ and $(\deg(a_1), \dots \deg(a_{\mathtt{t}}))$ are shortest lag structure of $\mathcal{P}$ and $\mathcal{D}$, respectively. Then under the assumptions of Lemma 5.19, we compute basis generators for $\mathfrak{N}_{\mathcal{C}}$.

5.3.2 Algorithm

Our solution to Problem 5.2 is summarised in Algorithm 6 on p. 66. Note that in Algorithm 6 we denote by $\mathfrak{N}^n_{\mathcal{C}}$ a set of annihilators of $\mathcal{C}$ of degree n.

We now prove the correctness of Algorithm 6 as follows.

Proposition 5.21. *Let $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$ be sufficiently informative about their respective behaviours. Let $r_1, \dots, r_{\mathtt{g}}$ and $a_1, \dots, a_{\mathtt{t}}$ in Algorithm 6 be such that $(deg(r_1), \dots deg(r_{\mathtt{p}}))$ and $(deg(a_1), \dots deg(a_{\mathtt{t}}))$ are shortest lag structure of $\mathcal{P}$ and $\mathcal{D}$, respectively. Let $\mathcal{C}$ implement $\mathcal{D}$, then $\mathfrak{N}_{\mathcal{C}}$ in Algorithm 6 is the module of annihilators of $\mathcal{C}$.*

Proof. The fact that $(\deg(r_1), \dots \deg(r_{\mathtt{p}}))$ and $(\deg(a_1), \dots \deg(a_{\mathtt{t}}))$ in Algorithm 6 are shortest lag structure of $\mathcal{P}$ and $\mathcal{D}$, respectively, follows from Theorem 2.43. Denote by $\mathfrak{N}^n_{\mathcal{D}}$, $\mathfrak{N}^n_{\mathcal{P}}$ and $\mathfrak{N}^n_{\mathcal{C}}$ the module annihilators of $\mathcal{D}, \mathcal{P}$ and $\mathcal{C}$ of degree n, receptively. From Algorithm 6 if $k = q$ then $a_{l'_1}, \dots, a_{l'_q} \in \mathfrak{N}^n_{\mathcal{P}}$. Hence, $\mathfrak{N}^n_{\mathcal{C}} = \{0\}$. Now, if $k = 0$ and $q \neq 0$, then $a_{l'_1}, \dots, a_{l'_q} \in \mathfrak{N}^n_{\mathcal{D}}$ such that $a_{l'_1}, \dots, a_{l'_q} \notin \mathfrak{N}^n_{\mathcal{P}}$ implies that $a_{l'_1}, \dots, a_{l'_q} \in \mathfrak{N}^n_{\mathcal{C}}$, therefore $\mathfrak{N}^n_{\mathcal{C}} = \{a_{l'_1}, \dots, a_{l'_q}\}$. Finally, $k < q$ means $\mathfrak{N}^n_{\mathcal{D}}$ has more generators of degree n than $\mathfrak{N}^n_{\mathcal{P}}$, therefore, some of them belong to $\mathfrak{N}^n_{\mathcal{C}}$. Denote by $\tilde{\mathfrak{N}}^n_{\mathcal{P}}$ and $\tilde{\mathfrak{N}}^n_{\mathcal{D}}$ the sets whose elements are $\tilde{a}_{l'_1}, \dots, \tilde{a}_{l'_q}$ and $\tilde{r}_{l_1}, \dots, \tilde{r}_{l_k}$, respectively. Since $\mathfrak{N}_{\mathcal{C}} \cap \mathfrak{N}_{\mathcal{P}} = \{0\}$ then $\mathfrak{N}^n_{\mathcal{C}} \cap \mathfrak{N}^n_{\mathcal{P}} = \{0\}$, moreover, $\tilde{\mathfrak{N}}^n_{\mathcal{C}} \cap \tilde{\mathfrak{N}}^n_{\mathcal{P}} = \{0\}$. Now under the assumption of Theorem 5.20 $\mathfrak{N}_{\mathcal{P}} + \mathfrak{N}_{\mathcal{C}} = \mathfrak{N}_{\mathcal{D}}$ then $\mathfrak{N}^n_{\mathcal{P}} + \mathfrak{N}^n_{\mathcal{C}} = \mathfrak{N}^n_{\mathcal{D}}$. Consequently, $\tilde{\mathfrak{N}}^n_{\mathcal{P}} + \tilde{\mathfrak{N}}^n_{\mathcal{C}} = \tilde{\mathfrak{N}}^n_{\mathcal{D}}$. Now since $\tilde{\mathfrak{N}}^n_{\mathcal{C}} \cap \tilde{\mathfrak{N}}^n_{\mathcal{P}} = \{0\}$, $\tilde{\mathfrak{N}}^n_{\mathcal{P}} + \tilde{\mathfrak{N}}^n_{\mathcal{C}} = \tilde{\mathfrak{N}}^n_{\mathcal{D}}$, and $\tilde{r}_{l_1}, \dots, \tilde{r}_{l_k}$ are bases of $\tilde{\mathfrak{N}}^n_{\mathcal{P}}$ then the projection matrix P exists and $\tilde{u}_1^\top, \dots \tilde{u}_x^\top$ are the coefficients vectors of annihilators of $\mathcal{C}$ of lag n. Hence, $\mathfrak{N}^n_{\mathcal{C}} = \{u_1, \dots, u_x\}$. □

5.3.3 Example

We consider an example of *power factor rectification* (see pp.53-54 of [25]) using passive devices like capacitors and inductor. The problem of reducing the reactive power which arises because of the the phase-shift between the source voltage and current.

Algorithm 6: Solution for Problem 5.2

Input : $\tilde{w} \in \mathcal{P}$ and $\tilde{w}_d \in \mathcal{D}$
Output : $\mathfrak{N}_{\mathcal{C}}$
Assumptions: Theorem 5.20 and Lemma 5.19

1 Determinations of bases generators of $\mathfrak{N}_{\mathcal{P}}$ and $\mathfrak{N}_{\mathcal{D}}$

i. Using Algorithm 2 determine shortest lag bases $r_1, \dots, r_{\mathtt{g}}$ and $a_1, \dots, a_{\mathtt{t}}$ of $\mathfrak{N}_{\mathcal{P}}$ and $\mathfrak{N}_{\mathcal{D}}$, respectively.

ii. Define $\tilde{r}_1, \dots, \tilde{r}_{\mathtt{g}}$ and $\tilde{a}_1, \dots, \tilde{a}_{\mathtt{t}}$ as the coefficients of $r_1, \dots, r_{\mathtt{g}}$ and $a_1, \dots, a_{\mathtt{t}}$ respectively.

iii. Define $d_m := \deg(a_m)$ for $m = 1, \dots, \mathtt{t}$, $\mathtt{t} := \{1, 2, \dots, \mathtt{t}\}$ and $\mathtt{g} := \{1, 2, \dots, \mathtt{g}\}$. Let $d = \max(d_1, \dots d_m)$.

2 Compute steps 3-4 recursively starting from $n = 0$ to d.

3 Classifying $\tilde{r}_1, \dots, \tilde{r}_{\mathtt{g}}$ and $\tilde{a}_1, \dots, \tilde{a}_{\mathtt{t}}$ by their lags

i. choose $l_1, \dots l_k \in \mathtt{g}$ such that $\tilde{r}_{l_1}, \dots, \tilde{r}_{l_k}$ are all of lag n. If there is no $\tilde{r}_{l_1}, \dots, \tilde{r}_{l_k}$ of lag n set $k = 0$.

ii. choose $l'_1, \dots l'_q \in \mathtt{t}$ such that $\tilde{a}_{l'_1}, \dots, \tilde{a}_{l'_q}$ are all of lag n. If there is no $\tilde{a}_{l'_1}, \dots, \tilde{a}_{l'_q}$ of lag n set $q = 0$.

4 Compute coefficients of the elements of $\mathfrak{N}_{\mathcal{C}}^n$ as follows

if $k = q$ **then**
  $\mathfrak{N}_{\mathcal{C}}^n := \{0\}$
else if $k = 0$ *and* $q \neq 0$ **then**
  $\tilde{a}_{l'_1}, \dots, \tilde{a}_{l'_q}$ are the coefficients of annihilators of $\mathcal{C}$ of degree n, therefore define $\mathfrak{N}_{\mathcal{C}}^n := \{a_{l'_1}, \dots, a_{l'_q}\}$.
else if $k < q$ **then** Define the matrix A whose columns are $\tilde{r}_{l_1}, \dots, \tilde{r}_{l_k}$ as

$$A := \begin{bmatrix} \tilde{r}_{0_{l_1}} & \dots & \tilde{r}_{0_{l_k}} \\ \vdots & \dots & \vdots \\ \tilde{r}_{n_{l_1}} & \dots & \tilde{r}_{n_{l_k}} \end{bmatrix};$$

Define $P := A(A^\top A)^{-1} A^\top$;
Define $H := [\tilde{a}_{l'_1} - P\tilde{a}_{l'_1}, \dots, \tilde{a}_{l'_q} - P\tilde{a}_{l'_q}]$;
Compute x rank of H and compute the SVD of $H = U \Sigma V^\top$;
Partition $U = [U_1 \; U_2]$ where U_1 has x columns;
The columns of U_1, $\tilde{u}_1^\top, \dots \tilde{u}_x^\top$, are the coefficients of annihilators of $\mathcal{C}$ of degree n, therefore, define $\mathfrak{N}_{\mathcal{C}}^n := \{u_1, \dots, u_x\}$;
end ;

5 Specification of $\mathfrak{N}_{\mathcal{C}}$

i. Define $\mathfrak{N}_{\mathcal{C}} := \bigcup_{k=0}^{d} \mathfrak{N}_{\mathcal{C}}^k$

Let the circuit in Figure 5.1 be the to-be-controlled system with $w = \text{col}(i_z, i_v, i_s, v)$, $\mathtt{m}(\mathcal{P}) = 2$ and $\mathtt{p}(\mathcal{P}) = 2$. The input/output variables are i_s, v and i_z, i_v, respectively. To generate $\tilde{w} \in \mathcal{P}$ with $T = 200000$ samples, the circuit in Figure 5.1 is simulated in `Matlab Simulink` with sampling rate of 50µs. To guarantee that $\tilde{w}$ is sufficiently informative i_s and v are generated by current and voltage sources which are driven by a random number generator so that both are persistently exciting of sufficiently high order. The values of R, L and C are 100Ω,0.01H and 0.001C, respectively.

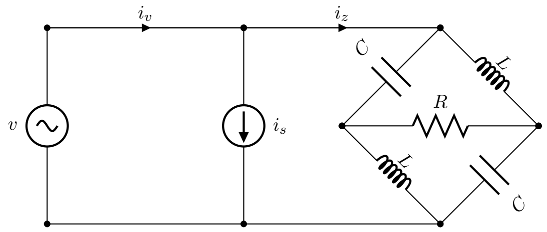

Figure 5.1: To-be controlled system

The controlled system, i.e. circuit with the correct power factor, is chosen as in Figure 5.2. $w_d = \text{col}(i_z, i_v, i_s, v)$, with $\mathtt{m}(\mathcal{D}) = 1$ and $\mathtt{p}(\mathcal{D}) = 3$. To generate $\tilde{w}_d \in \mathcal{D}$ the circuit in Figure 5.2 is simulated like the one above but this time with only v generated by voltage sources which is driven by a random number generator. The values of R, L and C are the same while $L_f = 0.001$H and $R_f = 0.4252\Omega$.

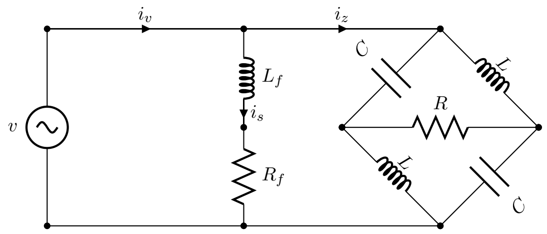

Figure 5.2: Example of controlled system

Using Algorithm 6, $\mathfrak{N}_{\mathcal{P}}$ has 2 basis generators, one of degree 0 and the other 2. $\mathfrak{N}_{\mathcal{D}}$ has 3 basis generators, one of degree 0, and others of degree 1 and 2. Hence, the generator of $\mathfrak{N}_{\mathcal{D}}$ of degree 1 belongs to $\mathfrak{N}_{\mathcal{C}}$. Consequently, a controller representation is

$$\begin{bmatrix} -\frac{322\,\sigma}{1611}+\frac{266}{1155} & \frac{322\,\sigma}{1611}-\frac{226}{1155} & -\frac{965\,\sigma}{2414}+\frac{452}{1155} & -\frac{151\,\sigma}{10180}-\frac{151}{10180} \end{bmatrix} \begin{bmatrix} i_z \\ i_v \\ i_s \\ v \end{bmatrix} = 0 \tag{5.9}$$

We verify the controller above by interconnecting it with the to-be-controlled systems then comparing the impulse response with that of the controlled system. The impulse responses coincide as shown in Figure 5.3.

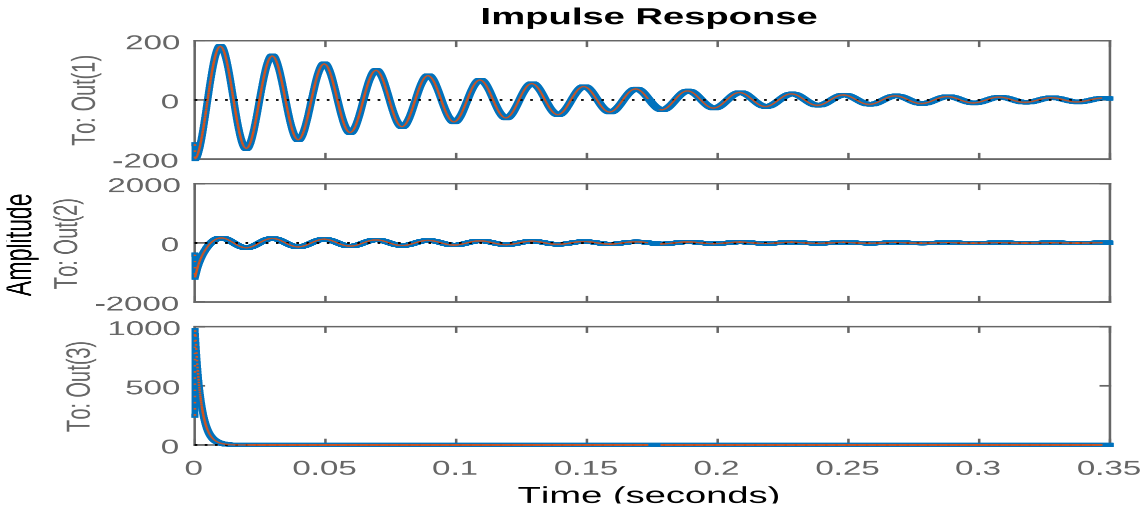

Figure 5.3: Impulse responses of the to-be-controlled system interconnected with computed controller (blue) and that of the controlled system(red)

5.4 Summary

We have presented solutions for two completely data-driven control problems. In the first solution we present sufficient conditions and an algorithm for find a controller when the plant and the controller interconnect through the control variables and the second algorithm is when the control variable and to-be-controlled variables coincide. In both cases we assume that the given data is generated with all input variable trajectories persistently exciting of sufficiently high order.

Chapter 6

Data-driven dissipativity

In this chapter, we consider a problem of determining whether a linear difference system is dissipative using sufficiently informative (see Definition 2.41) trajectories from such a system. In particular, given a sufficiently informative trajectory from a system and a supply rate, we state a necessary and sufficient condition such that the trajectory can be used to determine if the system is dissipative with respect to the given supply rate.

First we study *quadratic difference forms*, QdF. We discuss some theory and notations of quadratic difference forms and the theory of dissipativity using QdF's. The material presented in this section, like in the previous section, is based on polynomial rather than *Laurent polynomials*, see Appendix A Definition A.1.

6.1 Basics

Let $\Phi \in \mathbb{R}^{\mathtt{w}_1 \times \mathtt{w}_2}[\zeta, \eta]$ be a two variable polynomial matrix in the indeterminates ζ and η with real coefficients, then

$$\Phi(\zeta, \eta) := \sum_{k,j=0}^{N} \Phi_{k,j} \zeta^k \eta^j,$$

where $N \in \mathbb{N}$ and $\Phi_{k,j} \in \mathbb{R}^{\mathtt{w}_1 \times \mathtt{w}_2}$. For $w_i \in (\mathbb{R}^{\mathtt{w}_i})^{\mathbb{Z}}$, with $i = 1, 2$, $\Phi(\zeta, \eta)$ induces a *bilinear difference form*

$$L_\Phi : (\mathbb{R}^{\mathtt{w}_1})^{\mathbb{Z}} \times (\mathbb{R}^{\mathtt{w}_2})^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$$

defined by

$$L_\Phi(w_1, w_2)(t) := \sum_{k,j=0}^{N} w_1(t+k)^\top \Phi_{kj} w_2(t+j).$$

Now, let $\mathtt{w}_1 = \mathtt{w}_2 = \mathtt{w}$ and $\Phi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$. Then for $w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$, $\Phi(\zeta,\eta)$ induces a *quadratic difference form* (QdF)

$$Q_\Phi : (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$$

defined by

$$Q_\Phi(w)(t) := \sum_{k,j=0}^{N} w(t+k)^\top \Phi_{kl} w(t+j).$$

$\Phi(\zeta,\eta)$ is called *symmetric* if $\Phi(\zeta,\eta) = \Phi(\zeta,\eta)^\top$ and we denote by $\mathbb{R}_s^{\mathtt{w}_1\times\mathtt{w}_2}[\zeta,\eta]$ the ring of symmetric two variable polynomial matrices in the indeterminates ζ and η with real coefficients. $\Phi(\zeta,\eta)$ is closely associated with the *coefficient matrix*

$$\tilde{\Phi} := \begin{bmatrix} \Phi_{0,0} & \Phi_{0,1} & \dots & \Phi_{0,N} & \dots \\ \Phi_{1,0} & \Phi_{1,1} & \dots & \Phi_{1,N} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Phi_{N,0} & \Phi_{N,1} & \dots & \Phi_{N,N} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

in the sense that

$$\Phi(\zeta,\eta) = \begin{bmatrix} I_{\mathtt{w}} & \zeta I_{\mathtt{w}} & \dots & \zeta^N I_{\mathtt{w}} & \dots \end{bmatrix} \tilde{\Phi} \begin{bmatrix} I_{\mathtt{w}} \\ \eta I_{\mathtt{w}} \\ \vdots \\ \eta^N I_{\mathtt{w}} \\ \vdots \end{bmatrix}.$$

Notice that $\tilde{\Phi}$ is an infinite matrix with only a finite block $N \times N$ nonzero entries. Furthermore, $\Phi(\zeta,\eta)$ is symmetric if and only if its coefficient matrix is symmetric.

Even though $\tilde{\Phi}$ is infinite, the largest powers of ζ and η in $\Phi(\zeta,\eta)$ are finite. Therefore, we define the *effective size* of $\tilde{\Phi}$ by

$$N := \min\left\{N' \mid \Phi_{k,j} = 0 \text{ for all } k,j > N'\right\}.$$

We shall denote by $\tilde{\Phi}_N$ the coefficient matrix with effective size N.

Define the map

$$\nabla : \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta] \to \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$$

by

$$\nabla\Phi(\zeta,\eta) := (\zeta\eta - 1)\Phi(\zeta,\eta).$$

$\nabla\Phi$ induces a QdF $Q_{\nabla\Phi}$ which defines the *rate of change* of Q_Φ, i.e.

$$Q_{\nabla\Phi}(w)(t) := Q_\Phi(w)(t+1) - Q_\Phi(w)(t),$$

for all $t \in \mathbb{Z}$ and $w \in (\mathbb{R}^w)^\mathbb{Z}$.

6.1.1 Nonnegativity and positivity of QdFs

In this subsection we introduce concepts of positivity and nonnegativity.

Definition 6.1. Let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$, a QdF Q_Φ is called *nonnegative* if $Q_\Phi(w)(t) \geq 0$ for all $w \in (\mathbb{R}^\mathtt{w})^\mathbb{Z}$ and $t \in \mathbb{Z}$. If $Q_\Phi(w)(t) \geq 0$ and if $Q_\Phi(w)(t) = 0$ implies that $w = 0$ for all $w \in (\mathbb{R}^\mathtt{w})^\mathbb{Z}$ then we call Q_Φ *positive.*

We denote nonnegativity and positivity by $Q_\Phi \geq 0$ and $Q_\Phi > 0$, respectively.

Proposition 6.2. *Let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ with coefficient matrix $\tilde{\Phi}$. Then $Q_\Phi \geq 0$ if and only if $\tilde{\Phi} \geq 0$.*

Proof. See Proposition 2.1 p. 33 of [30]. □

Now let $\mathfrak{B} \in \mathscr{L}^\mathtt{w}$. We define positivity and nonnegativity along $\mathfrak{B}$ as follows.

Definition 6.3. Q_Φ is called *nonnegative along* $\mathfrak{B}$ if $Q_\Phi(w)(t) \geq 0$ for all $w \in \mathfrak{B}$. Furthermore, if $Q_\Phi(w)(t) \geq 0$ and if $Q_\Phi(w)(t) = 0$ implies that $w = 0$ for all $w \in \mathfrak{B}$ then Q_Φ is *positive along* $\mathfrak{B}$.

Nonnegativity and positivity along $\mathfrak{B}$ are denoted by $Q_\Phi \overset{\mathfrak{B}}{\geq} 0$ and $Q_\Phi \overset{\mathfrak{B}}{>} 0$, respectively.

6.1.2 Average positivity and average nonnegativity

We now introduce the concepts of average positivity and average nonnegativity.

Definition 6.4. Let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$. Then Q_Φ is called

1. *average nonnegative* if $\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) \geq 0$ for all $w \in l_2^\mathtt{w}$ and for all $t \in \mathbb{Z}$.

2. *average positive* if $\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) > 0$ for all nonzero $w \in l_2^\mathtt{w}$ and for all $t \in \mathbb{Z}$.

In Proposition 6.5 we state sufficient conditions for average positivity and average non-negativity.

Proposition 6.5. *Let* $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$. *Then*

1. $\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) \geq 0$ *for all* $w \in l_2^{\mathtt{w}}$ *if and only if* $\Phi(e^{-j\omega}, e^{j\omega}) \geq 0$ *for all* $\omega \in [0, 2\pi)$.
2. $\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) > 0$ *for all nonzero* $w \in l_2^{\mathtt{w}} \cap (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ *if and only if* $\Phi(e^{-j\omega}, e^{j\omega}) \geq 0$ *for all* $\omega \in [0, 2\pi)$ *and det*$(\Phi(\xi^{-1}, \xi)) \neq 0$

Proof. see Proposition 3.1 p. 35 of [30]. □

6.2 Dissipative systems

The study of how dynamical systems interact with environment in terms of energy supplied, stored and dissipated was first introduced in [78]. *Dissipativity* means that an increase in the stored energy cannot exceed the supplied energy. Dissipativity is useful for various problems in systems theory including stability, stabilisability, LQ control, system identification and model reduction (see [17, 21, 31, 55, 86, 83]). We introduce some concepts of dissipativity.

Definition 6.6. Let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ and $\mathfrak{B} \in \mathscr{L}_{contr}^{\mathtt{w}}$. Then $\mathfrak{B}$ is called

1. $\Phi-$*dissipative* if for all $w \in l_2^{\mathtt{w}} \cap \mathfrak{B}$ it holds that
$$\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) \geq 0.$$
2. $\Phi-$*half-line dissipative* if for every $T \in \mathbb{Z}$ and for all $w \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[-\infty,T]}$ it holds that
$$\sum_{t=-\infty}^{T} Q_\Phi(w)(t) \geq 0.$$

Consider a case when some of the supplied energy is stored in the system. The stored energy is also associated with a QdF defined as follows.

Definition 6.7. Let $\Psi, \Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ and $\mathfrak{B} \in \mathscr{L}_{contr}^{\mathtt{w}}$ be $\Phi-$dissipative. The QdF Q_Ψ induced by $\Psi(\zeta,\eta)$ is called a *storage function* for $\mathfrak{B}$ with respect to Q_Φ if for all $w \in \mathfrak{B}$ and all $t \in \mathbb{Z}$, it holds that
$$Q_{\nabla\Psi}(w)(t) \leqslant Q_\Phi(w)(t). \tag{6.1}$$

Equation (6.1) is called the *dissipation inequality*, and explains the property that the rate of change of the stored energy is less than or equal the supplied energy. To account for the portion of the supplied energy which is not stored consider the following definition.

Definition 6.8. Let $\Delta, \Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ and $\mathfrak{B} \in \mathscr{L}_{contr}^{\mathtt{w}}$ be $\Phi-$dissipative. Q_Δ is called a *dissipation function* for $\mathfrak{B}$ with respect to Q_Φ if $Q_\Delta \geq 0$, and for all $w \in l_2^{\mathtt{w}} \cap \mathfrak{B}$ and all $t \in \mathbb{Z}$

$$\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) = \sum_{t=-\infty}^{\infty} Q_\Delta(w)(t).$$

Let $\mathfrak{B} \in \mathscr{L}_{contr}^{\mathtt{w}}$ admit an observable image representation

$$w = M(\sigma)\ell \tag{6.2}$$

where $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ and $\ell \in (\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}}$. Now let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ such that

$$\sum_{t=-\infty}^{\infty} Q_\Phi(w)(t) \geq 0 \tag{6.3}$$

for all $w \in l_2^{\mathtt{w}} \cap \mathfrak{B}$ and all $t \in \mathbb{Z}$. By substituting (6.2) in (6.3) we get

$$\sum_{t=-\infty}^{\infty} Q_\Phi(M(\sigma)\ell)(t) \geq 0$$

for all $\ell \in (\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}} \cap l_2^{\mathtt{l}}$. Therefore, instead of studying dissipativity along manifest trajectories we can always study in terms unconstrained latent variable trajectories. Define

$$\Phi' := M(\zeta)^\top \Phi(\zeta,\eta) M(\eta),$$

then

$$Q_\Phi(M(\sigma)\ell) = Q_{\Phi'}(\ell).$$

We relate supply rate, storage and dissipation functions as follows.

Proposition 6.9. *Let $\mathfrak{B} \in \mathscr{L}_{contr}^{\mathtt{w}}$ and let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$. Assume that $\mathfrak{B} = im(M(\sigma))$, where $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ is full column rank for all $\lambda \in \mathbb{C}$. Then the following statements are equivalent:*

1. *$\mathfrak{B}$ is $\Phi-$dissipative;*
2. *there exists a storage function for $\mathfrak{B}$ with respect to Q_Φ;*
3. *there exists a dissipation function for $\mathfrak{B}$ with respect to Q_Φ;*
4. *$M(e^{-j\omega})^\top \Phi(e^{-j\omega}, e^{j\omega}) M(e^{j\omega}) \geq 0$ for all $\omega \in [0, 2\pi)$.*

Moreover, let $\Psi, \Delta \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ *be such that* Q_Ψ *and* Q_Δ *induce a storage and dissipation function for* $\mathfrak{B}$ *with respect to* Φ*. Then there is a one-one relationship between* Q_Ψ, Q_Δ Q_Φ *described by*

$$Q_{\nabla\Psi}(w)(t) = Q_\Phi(w)(t) - Q_\Delta(w)(t), \tag{6.4}$$

for all $t \in \mathbb{Z}$ *and* $w \in \mathfrak{B}$*, equivalently,*

$$\Psi(\zeta,\eta) = \frac{\Phi(\zeta,\eta) - \Delta(\zeta,\eta)}{\zeta\eta - 1}.$$

Proof. See Proposition 3.3, p. 39 of [30]. □

Equation (6.4) is called the *dissipation equality.*

6.3 Dissipativity from data

In this section we develop our results on data-driven dissipativity.

Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ and $\tilde{w} \in \mathfrak{B}_{|[0,\infty]}$ be sufficiently informative (see Definition 2.41) about $\mathfrak{B}$ then for all $\tilde{w}' \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$ there exists $\alpha \in \mathbb{R}^\infty$ such that $\tilde{w}' = \mathfrak{H}(\tilde{w})\alpha$. Let the Hankel matrix associated with $\tilde{w}' \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$ be

$$\mathfrak{H}(\tilde{w}') = \underbrace{\begin{bmatrix} \tilde{w}(0) & \tilde{w}(1) & \tilde{w}(2) & \dots \\ \tilde{w}(1) & \tilde{w}(2) & \tilde{w}(3) & \dots \\ \tilde{w}(2) & \tilde{w}(3) & \tilde{w}(4) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}}_{\mathfrak{H}(\tilde{w})} \underbrace{\begin{bmatrix} \alpha(0) & 0 & 0 & 0 & \dots \\ \alpha(1) & \alpha(0) & 0 & 0 & \dots \\ \alpha(2) & \alpha(1) & \alpha(0) & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}}_{\mathcal{A}(\alpha)}. \tag{6.5}$$

Now, let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ and denote by $\tilde{\Phi}$ the coefficient matrix of Φ. Compute

$$\mathfrak{H}(w')^\top \tilde{\Phi} \mathfrak{H}(w') = \begin{bmatrix} Q_\Phi(\tilde{w}')(0) & L_\Phi(\tilde{w}', \sigma\tilde{w}')(0) & L_\Phi(\tilde{w}', \sigma^2 w')(0) & \dots \\ L_\Phi(\sigma\tilde{w}', \tilde{w}')(0) & Q_\Phi(\sigma\tilde{w}', \sigma\tilde{w}')(0) & L_\Phi(\sigma\tilde{w}', \sigma^2\tilde{w}')(0) & \dots \\ L_\Phi(\sigma^2\tilde{w}', \tilde{w}')(0) & L_\Phi(\sigma^2\tilde{w}', \sigma\tilde{w}')(0) & Q_\Phi(\sigma^2\tilde{w}', \sigma^2\tilde{w}')(0) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{6.6}$$

and the trace of $\mathfrak{H}(w')^\top \tilde{\Phi} \mathfrak{H}(w')$ as

$$tr(\mathfrak{H}(\tilde{w}')^\top \tilde{\Phi} \mathfrak{H}(\tilde{w}')) = \sum_{t=0}^{\infty} Q_\Phi(\tilde{w}')(t) \tag{6.7}$$

From (6.5) $\mathfrak{H}(\tilde{w}') = \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)$, therefore,

$$\operatorname{tr}(\mathfrak{H}(\tilde{w}')^\top \tilde{\Phi} \mathfrak{H}(\tilde{w}')) = \operatorname{tr}(\mathcal{A}(\alpha)^\top \mathfrak{H}(\tilde{w})^\top \tilde{\Phi} \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)). \tag{6.8}$$

From (6.7) and (6.8) $\tilde{w} \in \mathfrak{B}_{|[0,\infty]}$ can be used to determine if $\mathfrak{B}$ is half-line dissipative with respect to Φ as we state in the following result.

Theorem 6.10. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $\Phi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}_s[\zeta,\eta]$ with the coefficient matrix denoted by $\tilde{\Phi}$, and $\tilde{w} \in \mathfrak{B}_{|[0,\infty]}$. Assume that $\tilde{w}$ is sufficiently about $\mathfrak{B}$. The following statements are equivalent:*

1. *$\mathfrak{B}$ is half-line dissipative with respect to Φ,*
2. *tr$(\mathcal{A}(\alpha)^\top \mathfrak{H}(\tilde{w})^\top \tilde{\Phi} \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)) \geq 0$, for all $\alpha^\top \in \mathbb{R}^{1\times\infty}$ such that $\mathfrak{H}(\tilde{w}') = \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)$ and $\tilde{w}' \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$.*

Proof. We proceed to prove (1) $\implies$ (2) as follows. Assume that $\mathfrak{B}$ is $\Phi-$half-line dissipative and recall that $\operatorname{tr}(\mathcal{A}(\alpha))^\top \mathfrak{H}(\tilde{w})^\top \tilde{\Phi} \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)) = Q_\Phi(\tilde{w}')(0) + Q_\Phi(\tilde{w}')(1) + \cdots = \sum_{t=0}^{\infty} Q_\Phi(\tilde{w}')(t)$. Now by the assumption that $\mathfrak{B}$ is $\Phi-$half-line dissipative it follows that $\sum_{t=0}^{\infty} Q_\Phi(\tilde{w}')(t) \geq 0$ for all $\tilde{w}' \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$ which implies for all α such that $\tilde{w}' = \mathfrak{H}(\tilde{w})\alpha \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$ then $\operatorname{tr}(\mathcal{A}(\alpha))^\top \mathfrak{H}(\tilde{w})^\top \tilde{\Phi} \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha))) \geq 0$.
(2) $\implies$ (1) Assume that $\operatorname{tr}(\mathcal{A}(\alpha)^\top \mathfrak{H}(\tilde{w})^\top \tilde{\Phi} \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)) \geq 0$ for all α such that $\tilde{w}' = \mathfrak{H}(\tilde{w})\alpha \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$. Now since $\operatorname{tr}(\mathcal{A}(\alpha))^\top \mathfrak{H}(\tilde{w})^\top \tilde{\Phi} \mathfrak{H}(\tilde{w})\mathcal{A}(\alpha)) = Q_\Phi(\tilde{w}')(0) + Q_\Phi(\tilde{w}')(1) + \cdots = \sum_{t=0}^{\infty} Q_\Phi(\tilde{w}')(t)$ then it follows that $\sum_{t=0}^{\infty} Q_\Phi(\tilde{w}')(t) \geq 0$, therefore, $\sum_{t=0}^{\infty} Q_\Phi(\tilde{w}')(t) \geq 0$ for all α such that $\tilde{w}' = \mathfrak{H}(\tilde{w})\alpha \in (l_2^{\mathtt{w}} \cap \mathfrak{B})_{|[0,\infty]}$ which implies that $\mathfrak{B}$ is half-line dissipative with respect to the supply rate Q_Φ. □

Since we are dealing with data, in practical terms data is normally of finite length, therefore, we need to develop dissipativity condition on finite data. First consider the following definition.

Definition 6.11. Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$, $\Phi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}_s[\zeta,\eta]$ and $L \in \mathbb{Z}_+$. $\mathfrak{B}$ is called *$\Phi-L$-dissipative* if $\sum_{t=0}^{L} Q_\Phi(w)(t) \geq 0$ for all $w \in \mathfrak{B}_{|[0,L]}$.

Now, let $\alpha^\top \in \mathbb{R}^{1\times L}$ and define the matrix $\mathcal{A}_{2L}(\alpha) \in \mathbb{R}^{2L\times L}$ by

$$\mathcal{A}_{2L}(\alpha) := \begin{bmatrix} \alpha(0) & 0 & 0 & \dots & 0 \\ \alpha(1) & \alpha(0) & 0 & \dots & 0 \\ \alpha(2) & \alpha(1) & \alpha(0) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha(L) & \alpha(L-1) & \alpha(L-2) & \dots & \alpha(0) \\ 0 & \alpha(L) & \alpha(L-1) & \dots & \alpha(1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha(L) \end{bmatrix},$$

Let $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$, and assume that the effective size of $\tilde{\Phi}$ is $N \in Z_+$. Choose $L \in \mathbb{Z}_+$ such that $L \geq N$ denote by $\tilde{\Phi}_L$ the $L \times L$ matrix

$$\tilde{\Phi}_L := \begin{bmatrix} \Phi_{0,0} & \Phi_{0,1} & \dots & \Phi_{0,N} & 0_{\mathtt{w}\times(L-N)\mathtt{w}} \\ \Phi_{1,0} & \Phi_{1,1} & \dots & \Phi_{1,N} & 0_{\mathtt{w}\times(L-N)\mathtt{w}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Phi_{N,0} & \Phi_{N,1} & \dots & \Phi_{N,N} & 0_{\mathtt{w}\times(L-N)\mathtt{w}} \\ 0_{(L-N)\mathtt{w}\times\mathtt{w}} & 0_{(L-N)\mathtt{w}\times\mathtt{w}} & 0_{(L-N)\mathtt{w}\times\mathtt{w}} & 0_{(L-N)\mathtt{w}\times\mathtt{w}} & 0_{(L-N)\mathtt{w}\times\mathtt{w}} \end{bmatrix},$$

Finally, let $T \in \mathbb{N}$, $\tilde{w} \in \mathfrak{B}_{|[0,T]}$ be sufficiently informative about $\mathfrak{B}$ and denote the Hankel matrix of depth L and finite number of columns $2L$ associated with $\tilde{w}$ by $\mathfrak{H}_{L,2L}(\tilde{w})$.

Theorem 6.12. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$. Furthermore, let $L \in \mathbb{Z}_+$ $\tilde{w} \in \mathfrak{B}_{|[0,T]}$, α and $\tilde{\Phi}_L$ as above. $\mathfrak{B}$ is $\Phi - L$-dissipative if and only if for all α,*
$tr(\mathcal{A}_{2L}(\alpha)^\top \mathfrak{H}_{L\times 2L}(\tilde{w})^\top \tilde{\Phi}_L \mathfrak{H}_{L\times 2L}(\tilde{w})\mathcal{A}_{2L}(\alpha)) \geq 0$.

Proof. Follows the same argument in the proof of Theorem 6.10. □

In the following Proposition we give the relationship between $\Phi - L$-dissipative and Φ-dissipative.

Proposition 6.13. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $L \in \mathbb{Z}_+$, $\Phi \in \mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$. If for all $w \in l_2^{\mathtt{w}} \cap \mathfrak{B}$ and $t_0 \in \mathbb{Z}$ there exist $w' \in \mathfrak{B}_{|[t_0,t_0+L]}$ satisfying $w(t) = w'(t)$ for $t_0 \leqslant t \leqslant t_0 + L$ such that $\sum_{t=t_0}^{t_0+L} Q_\Phi(w')(t) \geq 0$ then $\mathfrak{B}$ is $\Phi-$dissipative.*

Proof. Since for all L long samples of w it holds that $\sum_{t=t_0}^{t_0+L} Q_\Phi(w)(t) \geq 0$ then $\sum_{-\infty}^{\infty} Q_\Phi(w)(t) \geq 0$, hence, $\mathfrak{B}$ is $\Phi-$dissipative. □

Remark 6.14. A possible future research direction is to investigate if $\Phi - L-$dissipativity can prove $\Phi-$dissipativity and to come up with an effective way of verifying conditions of Theorem 6.10. We believe a possible starting point would be to consider how to infer the storage function from data. □

6.4 Summary

We have shown a necessary and sufficient condition to determine whether a system is dissipative with respect to a given supply rate using data. We have also introduced the notion of $L-$dissipative for finite observed trajectory.

Chapter 7

Flatness and parametrizations

7.1 Introduction

The notion of *differential flatness* for continuous-time system was first introduced in [14] in 1995. Since then it has been studied by several authors see [12, 15, 34, 35, 66], and applied to some control problems [29, 49, 66, 97]. Differential flatness or simply *flat/flatness*, is a property of systems where by the system variables such as the state, inputs and outputs variables are parametrized by a set of free variables and a finite number of derivatives of these free variables. These free variables are called *flat outputs* or *linearly flat outputs* in the case of linear systems. Discrete-time flatness, i.e. *difference flatness*, was first studied in [66] then in [6, 16, 28]. In the discrete setting, system variables are parametrized by flat outputs and a finite number of their backward and/or *forward* shifts.

In the behavioural framework flatness is first mentioned in [66] and covered in more details in [72] for a class of linear time-invariant systems. The author of [72] defines linear flat systems as systems whose behaviour admits an observable image representation with the flat outputs as the latent variables and this is proven in Theorem 5.1 p. 54 of [72]. In the theorem, it is also shown that flatness is preserved under unimodular transformation of the image representation.

In this chapter, in a behavioural setting, we shall propose a different definition of linear flatness. The definition is based on trajectories, consquently, we accomodate cases when there is a partition of manefist variables such that some of the manifest variables are the flat variables. We also present a characterization of the image and kernel representations of flat systems, see [36] for characterization for flat outputs. In [36], flat outputs are characterized by polynomial matrices which the authors calls *defining matrices*.

We show how given an observable image representation, one can find flat variables by finding a square sub-matrix of the image representation with a constant determinant. We then study how our results on characterization of the image and kernel representations of flat system fit with the notion of control as interconnection. We give a characterization of the kernel representation of a controller interconnected with a flat system via full interconnection. The case of partial interconnection is also studied. We prove necessary and sufficient conditions for the control variable to be the flat variables. Finally, we give a characterization of all controllers that interconnect with flat systems via the control variables.

We end the chapter by showing a parametrization of all controllable controlled behaviours and all controllable controllers

7.2 Flat systems

Consider a system

$$x(t+1) = f(x(t), u(t)) \tag{7.1}$$

where $t \in \mathbb{Z}$, $x(t) \in \mathbb{R}^{\mathtt{n}}$ is the state, $u(t) \in \mathbb{R}^{\mathtt{m}}$ is the input and $f : \mathbb{R}^{\mathtt{n}} \times \mathbb{R}^{\mathtt{m}} \to \mathbb{R}^{\mathtt{n}}$.

Definition 7.1. The system in (7.1) is called *difference flat* (or simply *flat)* if there exist non-negative integers $\mathtt{p}$, $\mathtt{q}$ and functions $h : \mathbb{R}^{\mathtt{n}} \times \mathbb{R}^{(\mathtt{p}+1)\mathtt{m}} \to \mathbb{R}^{\mathtt{m}}$, $\alpha : \mathbb{R}^{(\mathtt{q}+1)\mathtt{m}} \to \mathbb{R}^{\mathtt{n}}$, $\beta : \mathbb{R}^{(\mathtt{q}+1)\mathtt{m}} \to \mathbb{R}^{\mathtt{m}}$ such that (x, u) satisfies (7.1) if and only if there exists $y(t) \in \mathbb{R}^{\mathtt{m}}$ such that

$$\begin{aligned} x(t) &= \alpha(y(t), \sigma y, \sigma^2 y, \ldots, \sigma^{\mathtt{q}} y) \\ u(t) &= \beta(y(t), \sigma y, \sigma^2 y, \ldots, \sigma^{\mathtt{q}} y), \end{aligned}$$

and

$$y(t) = h(x(t), u(t), \sigma u, \sigma^2 u, \ldots, \sigma^{\mathtt{p}} u).$$

In other words, (7.1) is flat if there exists a set of independent variables y called *flat outputs* such that x and u are functions of these flat outputs and a finite number of their backward shifts.

We turn our attention to the concepts of flatness covered in [72]. We state the discrete version of definition of linear flat behaviours of [72] as follows.

Definition 7.2. Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$. $\mathfrak{B}$ is called *linearly flat* if there exists a non-negative integer $\mathtt{l}$, matrices $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ and $L \in \mathbb{R}^{\mathtt{l}\times\mathtt{w}}[\xi]$ such that

i. $\mathfrak{B} = \{w \mid \text{there exist } \ell \in (\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}} \text{ such that } w = M(\sigma)\ell\}$,

ii. $\ell \in (\mathbb{R}^{\mathtt{l}})^{\mathbb{Z}}$ such that i. holds implies $\ell = L(\sigma)w$.

From Definition 7.2 linearly flat system behaviours are those that admit an observable image representation as already mentioned. Furthermore, if i)-ii) in Definition 7.2 holds then ℓ is called the *linearly flat latent variable* for $\mathfrak{B}$.

The following is worth noting about Definition 7.2:

- The definition does not use basic fundamentals of behaviours, i.e. systems are defined as trajectories not using representation;
- As a consequence of using representations in the definition only latent variables can be considered as flat variables;
- Controllability is implicitly assumed, since only controllable systems admit an image representation.

Consider the following example, which we shall also use later in the discussion of our proposed definition of flatness.

Example 7.1. Consider the example of motorized stage Example 6.2 pp. 168-170 of from [34]. See Figure 6.6, p.168 of [34] for the positioning system and the parameters; M moving mass, M_B base mass, k stiffness coefficient and r damping coefficient. The system has two manifest variable $\mathrm{col}(x, x_B)$ and the manifest behaviour of the system is induced

$$R := \begin{bmatrix} M\xi^2 & M_B\xi^2 + r\xi + k \end{bmatrix} \tag{7.2}$$

In order to define flat variable on p.170 the authors introduce a latent variable y such that now the system admits an observable image representation induced by

$$M := \begin{bmatrix} \frac{1}{k}(M_B\xi^2 + r\xi + k) \\ -\frac{M}{k}\xi^2 \end{bmatrix} \tag{7.3}$$

with the left inverse of M

$$L := \begin{bmatrix} 1 - \frac{r}{k}\xi & \frac{1}{M}(M_B - r^2) - \frac{M_B r}{MK}\xi \end{bmatrix} \tag{7.4}$$

In this case the system whose behaviour is induce by (7.3) is flat since $\mathrm{col}(x, x_B)$ is a function of y and y is free. Moreover, y can be defined as a function of $\mathrm{col}(x, x_B)$ using L. □

We now state our definition.

Definition 7.3. Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$. Then $\mathfrak{B}$ is called *linearly flat* if there exists a partition of the manifest variables (w_1, w_2) such that:

a. w_1 is free,

b. w_2 is observable from w_1.

From Definition 7.3 it follow that if (a) and (b) holds then w_1 is called *linearly flat variable.*

The following is worth noting the following about Definition 7.3.

i. It is representations free and the flat variables are not restricted to be the system latent variables in contrast to Definition 7.2.

ii. It can be shown that Definition 7.2 is a special case of Definition 7.3 where the system admits a minimal hybrid representation

$$R(\sigma)w = M(\sigma)\ell$$

with $R = I_{\mathtt{w}}$ and $M(\lambda)$ full column rank for all $\lambda \in \mathbb{C}$. Clearly, in such cases w is observable from ℓ and ℓ is free.

iii. The Definition is not such that one need to introduce some latent variable in order to define flat output as in Example 7.1. Furthermore, in the case where one has to introduce latent variables like in Example 7.1 then the definition still holds.

iv. Like in [72], the number of flat variables is not required to be equal to the number of inputs of the system. We shall elaborate on this point later in this section.

v. In the case $\mathfrak{B} = (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ then $\mathfrak{B}$ is flat if the w_2 variables are non void, otherwise, $\mathfrak{B}$ is not flat.

vi. $\mathfrak{B}$ is not required to be controllable. However, one can readily see that indeed the definition implies controllability as we sure in the result.

Proposition 7.4. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$. If $\mathfrak{B}$ is flat then $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$.*

Proof. From Definition 7.3, then there a exists a partition of the variables (w_1, w_2) of $\mathfrak{B}$ such that w_1 is free and w_2 is observable from w_1. w_2 is observable from w_1 implies that there exists $M \in \mathbb{R}^{\mathtt{w}_2 \times \mathtt{w}_1}[\xi]$ such that $w_2 = M(\sigma)w_1$. Now since w_1 is free then it follows from Theorem 2.16 that $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$. □

We now present a characterization of the image and kernel representation of flat systems.

Theorem 7.5. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and assume that $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{1}}[\xi]$ full column rank for all $\lambda \in \mathbb{C}$ and $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ full row rank are such that $\mathfrak{B} = im(M(\sigma)) = \ker(R(\sigma))$. The following statements are equivalent:*

1. *$\mathfrak{B}$ is flat,*
2. *there exists a unimodular sub-matrix $U \in \mathbb{R}^{\mathtt{1}\times\mathtt{1}}[\xi]$ of M and a matrix $P \in \mathbb{R}^{(\mathtt{w}-\mathtt{1})\times\mathtt{1}}[\xi]$ such that*
$$MU^{-1} = \begin{bmatrix} I_1 \\ P \end{bmatrix},$$
3. *there exist $R_1 \in \mathbb{R}^{(\mathtt{w}-\mathtt{1})\times\mathtt{1}}[\xi]$ such that $\begin{bmatrix} R_1 & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix}$ induces a minimal kernel representation of $\mathfrak{B}$*
4. *there exists a partition of $R = \begin{bmatrix} R' & R'' \end{bmatrix}$, with $R' \in \mathbb{R}^{\mathtt{g}\times\mathtt{1}}[\xi]$ and $R'' \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ such that R'' is unimodular.*

Proof. We start by proving the equivalence between (3) and (4). (3) $\Longrightarrow$ (4) Notice that $\begin{bmatrix} R_1 & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix}$ is full row rank, hence, it induces a minimal kernel representation. Define $\mathtt{g} := (\mathtt{w}-\mathtt{1})$, $R' := R_1$ and $R'' := -I_{\mathtt{g}}$. Since $R'' = -I_{\mathtt{g}}$ then R'' is square and has a constant determinant, hence, R'' is unimodular.

The proof of (4) $\Longrightarrow$ (3) is as follows. Since R'' is unimodular then there exists $U' \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ such that $U'R'' = -I_{\mathtt{g}}$. Define $\mathtt{g} =: (\mathtt{w}-\mathtt{1})$ and $R_1 := U'R'$ then

$$U' \begin{bmatrix} R' & R'' \end{bmatrix} = \begin{bmatrix} R_1 & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix},$$

moreover, $\begin{bmatrix} R_1 & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix}$ is full row rank, hence, it induces a minimal representation of $\mathfrak{B}$.

Now we prove (1) $\Longleftrightarrow$ (3). We start with (3) $\Longrightarrow$ (1), let $w \in \mathfrak{B}$ and partition $w = \mathrm{col}(w_1, w_2)$ with respect to the partition of $\begin{bmatrix} R_1 & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix}$. Then

$$\begin{bmatrix} R_1(\sigma) & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0,$$

hence, $w_2 = R_1(\sigma)w_1$ which implies that w_2 is observable from w_1. Moreover, $w_2 = R_1(\sigma)w_1$ does not impose any restrictions on w_1, therefore, w_1 is free. From Definition 7.3 then $\mathfrak{B}$ is flat.

(1) $\Longrightarrow$ (3) Let $\mathfrak{B}$ be flat with a flat partition $\mathrm{col}(w_1, w_2)$ where w_1 is the flat variable. From Definition 7.3 since w_2 is observable from w_1 then there exists $F \in \mathbb{R}^{(\mathtt{w}-\mathtt{1})\times\mathtt{1}}$ such that $w_2 = F(\sigma)w_1$. Therefore,

$$\begin{bmatrix} F & -I_{(\mathtt{w}-\mathtt{1})} \end{bmatrix}(\sigma) \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0$$

induces a minimal kernel representation of $\mathfrak{B}$. To conclude the proof define $R_1 := F$.
We proceed to show the equivalence between (2) and (3).
(3) $\implies$ (2) Let $\mathfrak{B} = \ker\left(\begin{bmatrix} R_1(\sigma) & -I_{\mathtt{w}-1}\end{bmatrix}\right)$ and notice that

$$\begin{bmatrix} R_1 & -I_{\mathtt{w}-1}\end{bmatrix}\begin{bmatrix} I_1 \\ R_1\end{bmatrix} = 0.$$

It follows from Proposition 2.18 that $\begin{bmatrix} I_1 \\ R_1\end{bmatrix}$ induces a minimal image representation of $\mathfrak{B}$. Define $P := R_1$ then

$$\mathfrak{B} = \operatorname{im}\left(\begin{bmatrix} I_1 \\ P(\sigma)\end{bmatrix}\right) = \operatorname{im}(M(\sigma)).$$

Partition

$$M = \begin{bmatrix} M_1 \\ M_2\end{bmatrix}$$

appropriately with the partitions of $\begin{bmatrix} I_1 \\ P\end{bmatrix}$. By the assumption that M is full column rank for all $\lambda \in \mathbb{C}$ and the fact that two matrices $\begin{bmatrix} I_1 \\ P\end{bmatrix}$ and $\begin{bmatrix} M_1 \\ M_2\end{bmatrix}$ induce an observable image representation of the same behaviour if and only if there exists a unimodular matrix $W \in \mathbb{R}^{\mathtt{1}\times\mathtt{1}}[\xi]$ such that

$$\begin{bmatrix} I_1 \\ P\end{bmatrix} = \begin{bmatrix} M_1 \\ M_2\end{bmatrix} W,$$

see Proposition 2.23, then $I_1 = M_1 W$. Now since W is unimodular then $M_1 = W^{-1}$. To complete the proof define $U := W^{-1}$.
(2) $\implies$ (3) Let P be such that

$$\begin{bmatrix} w_1 \\ w_2\end{bmatrix} = \begin{bmatrix} I_1 \\ P(\sigma)\end{bmatrix}\ell,$$

with $\ell \in (\mathbb{R}^{\mathtt{1}})^{\mathbb{Z}}$. Then $w_1 = \ell$, and this implies that $w_2 = P(\sigma)w_1$. Therefore,

$$\begin{bmatrix} P(\sigma) & -I_{\mathtt{w}-1}\end{bmatrix}\begin{bmatrix} w_1 \\ w_2\end{bmatrix} = 0,$$

define $R_1 := P$ then $\mathfrak{B} = \ker\left(\begin{bmatrix} R_1(\sigma) & -I_{\mathtt{w}-1}\end{bmatrix}\right)$. □

Remark 7.6. Given an image representation $\mathfrak{B} = \operatorname{im}(M(\sigma))$, finding flat manifest variables involves finding an $\mathtt{1}\times\mathtt{1}$ sub-matrix of M with constant determinant. □

Let $w \in \mathfrak{B}$, recall that a partition of the variables of $\mathfrak{B}$ defined by $w =: \operatorname{col}(w_1, w_2)$ defines an input/output partition if w_1 is maximally free. Furthermore, let $\mathfrak{B} = \ker(R(\sigma))$

where $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ is full row rank, from Theorem 2.30 an input/output partition is equivalent to a partition of $R = \begin{bmatrix} -Q & P \end{bmatrix}$ with $Q \in \mathbb{R}^{\mathtt{g}\times\mathtt{m}}[\xi]$, $P \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ and $\det(P) \neq 0$.

Corollary 7.7. *Let $\mathfrak{B} \in \mathscr{L}^{\mathtt{w}}$ be flat with a flat partition of variables* $\mathrm{col}(w_1, w_2)$ *and w_1 the flat variable. Then w_1 is maximally free.*

Proof. Since $\mathfrak{B}$ is flat it follows from Theorem 7.5 that $\begin{bmatrix} R_1 & -I_{(\mathtt{w}-1)} \end{bmatrix}$ induces a minimal kernel representation of $\mathfrak{B}$. Define $Q := R_1$ and $P := -I_{(\mathtt{w}-1)}$ then $\det(P) \neq 0$. Therefore, $\begin{bmatrix} Q & P \end{bmatrix}$ is an input/output partition, hence, w_1 is maximally free. □

It follows from Corollary 7.7 that flat variable are maximally free, consequently, the number of flat variables is equal to the number of inputs.

Example 7.2. Let $\mathfrak{B}$ with a kernel representation

$$\underbrace{\begin{bmatrix} \sigma^2+1 & \sigma & -1 & 0 \\ \sigma+1 & \sigma^2+\sigma & -\sigma-1 & 1 \end{bmatrix}}_{R} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0. \tag{7.5}$$

Since $\mathrm{rank}(R(\lambda)) = \mathrm{rank}(R) = 2$ then $\mathfrak{B}$ is controllable, hence we can find an observable image representation of by finding unimodular matrices

$$U := \begin{bmatrix} -\xi-1 & 1 \\ -1 & 0 \end{bmatrix}$$

and

$$V := \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\xi \\ 0 & 1 & \xi & 1 \\ 1 & 0 & 0 & \xi^3+\xi^2 \end{bmatrix}$$

such that

$$URV = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Define

$$M := V \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -\xi \\ \xi & 1 \\ 0 & \xi^3+\xi^2 \end{bmatrix}$$

and

$$L := \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} V^{-1} = \begin{bmatrix} \xi & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Then $\mathfrak{B}$ admits an observable image representation

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = M(\sigma) \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix},$$

moreover,

$$L(\sigma) \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix},$$

Compute 2×2 minors of M and the submatrix of M with constant minor is

$$U_1 = \begin{bmatrix} 0 & 1 \\ 1 & -\xi \end{bmatrix}.$$

Define

$$M' := MU_1^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \xi^2 + 1 & \xi \\ \xi^3 + \xi^2 & 0 \end{bmatrix}.$$

Now $\text{rank}(M'(\lambda)) = 2$, hence, $\mathfrak{B}$ also admits an observable image representation

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \sigma^2 + 1 & \sigma \\ \sigma^3 + \sigma^2 & 0 \end{bmatrix} \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix}$$

with $\text{col}(w_1, w_2)$ as the flat variables. Notice that $\mathfrak{B}$ also admits a kernel representation

$$\begin{bmatrix} \sigma^2 + 1 & \sigma & -1 & 0 \\ \sigma^3 + \sigma^2 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0$$

Finally, notice that $\mathtt{m}(\mathfrak{B}) = 2$ and the flat variables are w_1, w_2, hence, the number of flat variables is equal to the input cardinality. $\square$

7.3 Flatness and control as interconnection

In this section we study the control of flat systems using the interconnection paradigm introduced in Chapter 3. We show a characterization of the kernel representation of controllers interconnected with flat systems via both full and partial interconnections. In the partial interconnection we also study full information control.

7.3.1 Flatness and full interconnection

We start with a characterization of kernel representations of controllers that implements a given controlled behaviour via full interconnection with a flat system. We shall use $\mathcal{P}$, $\mathcal{K}$ and $\mathcal{C}$ to denote the to-be-controlled system behaviour, controlled behaviour and controller behaviour, respectively.

Theorem 7.8. *Let $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}_{contr}$ be a flat to-be-controlled system with a flat partition (w_1, w_2) and w_1 the flat variable. Assume that $\mathcal{C}$ regularly implements $\mathcal{K} \subseteq \mathcal{P}$. Then there exists*

$$C := \begin{bmatrix} C_1 & 0 \end{bmatrix} \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi], \tag{7.6}$$

such that $\mathcal{C} = \ker(C(\sigma))$.

Proof. Since $\mathcal{P}$ is flat with w_1 as the flat variable then it follows from Theorem 7.5 that $\mathcal{P}$ admit a minimal kernel representation

$$\begin{bmatrix} R(\sigma) & -I \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0.$$

Now, $\mathcal{K} \subseteq \mathcal{P}$ implies that a controller $\mathcal{C}$ exists implementing $\mathcal{K}$, see Theorem 3.2. Let $\mathcal{C} = \ker(D(\sigma))$ where D induces a minimal representation and partition D appropriately with respect to the partitions of $\begin{bmatrix} R & -I \end{bmatrix}$, i.e.

$$\mathcal{K} = \ker \underbrace{\left(\begin{bmatrix} R(\sigma) & -I \\ D_1(\sigma) & D_2(\sigma) \end{bmatrix} \right)}_{=: K(\sigma)}$$

Since $\begin{bmatrix} R & -I \end{bmatrix}$ and D are full row rank and by the assumption that $\mathcal{C}$ regularly implements $\mathcal{K}$ then K induces a minimal kernel representation of $\mathcal{K}$, see Proposition 3.7. Define the unimodular matrix

$$U := \begin{bmatrix} I & 0 \\ D_2 & I \end{bmatrix}.$$

Pre-multiply K by U

$$UK = \begin{bmatrix} R & -I \\ C_1 & 0 \end{bmatrix}$$

where $C_1 := D_2 R + D_1$. Now since U is unimodular and pre-multiplication by a unimodular matrix does not change the behaviour then

$$\mathcal{K} = \ker(K(\sigma)) = \ker(U(\sigma)K(\sigma)).$$

Moreover, since $\mathcal{C}$ regularly implements $\mathcal{K}$ then $\mathcal{C} = \ker\left(\begin{bmatrix} C_1(\sigma) & 0 \end{bmatrix}\right)$. □

Let $\mathcal{C} \in \mathscr{L}^{\mathtt{w}}$ and define a componentwise partition of $w \in \mathcal{C}$ by $w =: \mathrm{col}(w_1, w_2)$. Assume that $C \in \mathbb{R} \in \mathbb{R}^{\mathtt{t}\times\mathtt{w}}[\xi]$ induces a minimal kernel representation of $\mathcal{C}$. Define a partition of the columns of $C =: \begin{bmatrix} C_1 & C_2 \end{bmatrix}$ appropriately with respect to the componentwise partition of w.

Definition 7.9. $\mathcal{C}$ is said to *act only* on w_2 if $C_1 = 0$.

From the above definition, we say $\mathcal{C}$ acts only the w_2 if $\mathcal{C}$ admits a kernel representation of the form $\ker\left(\begin{bmatrix} 0_{\bullet\times\mathtt{w}} & C_2(\sigma) \end{bmatrix}\right)$.

From the literature, it has been shown that flatness is useful in solving problems like reference tracking and motion planning, see p.73 of [36]. Therefore, once the flat output are obtained the control of the system can be done by restricting only the flat outputs. Consequently, in the case of flat system covered in this work, if one get a representation of a controller that interconnects with a flat system such that the controller is not only acting on the flat output it is then necessary to find a representation such that the controller acts on the flat output. Moreover, such a representation always exists as shown in Proposition 7.8. Another case where they may be the need to find a controller acting on some of the variable is in Example 5.3.3 on Chapter 5. Notice that in the example, in Figure 5.2 the desired behaviour, the controller is acting only on the variables $\mathrm{col}(i_s, v)$. From the results of Algorithm 6 the representation of the controller is such the controller is acting on all the variables, see (5.9). Clearly, in this case for easy synthesis of the controller it is necessary to find a representation of the controller such that it is acting on $\mathrm{col}(i_s, v)$.

In the following results we show how to use a kernel representation of $\mathcal{P}$ to find a kernel representation of $\mathcal{C}$, $\begin{bmatrix} 0_{\bullet\times\mathtt{w}} & C_2(\sigma) \end{bmatrix}$, given any kernel representation of $\mathcal{C}$.

Proposition 7.10. *Let $\mathcal{P}$ be the to-be-controlled system and assume that $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ is full row rank and is such that $\mathcal{P} = \ker(R(\sigma))$. Define a partition of the columns of $R =: \begin{bmatrix} R_1 & R_2 \end{bmatrix}$ appropriately with respect to the partition of $w = \mathrm{col}(w_1, w_2)$ above. Assume that $\mathcal{K} \subseteq \mathcal{P}$ and that $\mathcal{C}$ regularly implements $\mathcal{K}$. Furthermore, assume that $\mathcal{C}$ can act only on the w_2 variables. Then for any minimal kernel representation of $\mathcal{C}$, $C' = \begin{bmatrix} C'_1 & C'_2 \end{bmatrix} \in \mathbb{R}^{\mathtt{t}\times\mathtt{w}}[\xi]$ partitioned with respect to the partition of $w = \mathrm{col}(w_1, w_2)$ with $C'_1 \neq 0$, there exists a unimodular matrix $U \in \mathbb{R}^{(\mathtt{g}+\mathtt{t})\times(\mathtt{g}+\mathtt{t})}[\xi]$ defined by*

$$U := \begin{bmatrix} I_{\mathtt{g}} & 0_{\mathtt{g}\times\mathtt{t}} \\ U' & -I_{\mathtt{t}} \end{bmatrix}$$

such that

$$\begin{bmatrix} R_1 & R_2 \\ 0 & C_2 \end{bmatrix} = U \begin{bmatrix} R_1 & R_2 \\ C'_1 & C'_2 \end{bmatrix}.$$

Furthermore, let $G \in \mathbb{R}^{\mathtt{t}\times(\mathtt{g}+\mathtt{t})}[\xi]$ *be such that its rows are a basis of the left syzygy of* $\begin{bmatrix} R_1 \\ C_1' \end{bmatrix}$ *then*

$$\begin{bmatrix} 0 & C_2 \end{bmatrix} = G \begin{bmatrix} R_1 & R_2 \\ C_1' & C_2' \end{bmatrix}. \tag{7.7}$$

Proof. Since both $\begin{bmatrix} R_1 & R_2 \end{bmatrix}$ and $\begin{bmatrix} C_1' & C_2' \end{bmatrix}$ induce minimal representation of their respective behaviours and $\mathcal{C}$ regularly implements $\mathcal{K}$ then it follows from Proposition 3.7 that $\begin{bmatrix} R_1 & R_2 \\ C_1' & C_2' \end{bmatrix}$ induces a minimal kernel representation of $\mathcal{K}$. Now by the assumption that $\mathcal{C}$ can act only on the w_2 variables and the fact that pre-multiplication of a kernel representation by a unimodular matrix does not change the behaviour induced by that kernel representation then U exists such that

$$\begin{bmatrix} R_1 & R_2 \\ 0 & C_2 \end{bmatrix} = U \begin{bmatrix} R_1 & R_2 \\ C_1' & C_2' \end{bmatrix}.$$

Now to show the structure of U partition U appropriately with the partition of $\begin{bmatrix} R_1 & R_2 \\ C_1' & C_2' \end{bmatrix}$, i.e.

$$\begin{bmatrix} R_1 & R_2 \\ 0 & C_2 \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} R_1 & R_2 \\ C_1' & C_2' \end{bmatrix}$$

where $U_{11} \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$, $U_{12} \in \mathbb{R}^{\mathtt{g}\times\mathtt{t}}[\xi]$, $U_{21} \in \mathbb{R}^{\mathtt{t}\times\mathtt{g}}[\xi]$ and $U_{22} \in \mathbb{R}^{\mathtt{t}\times\mathtt{t}}[\xi]$. Then

$$\begin{aligned} R_1 &= U_{11}R_1 + U_{12}C_1 \\ R_2 &= U_{11}R_2 + U_{12}C_2 \\ 0 &= U_{12}R_1 + U_{22}C_1. \end{aligned}$$

From first two equations above and the fact that $\mathrm{col}(R, C')$ is full row rank, this yields $U_{11} = I_{\mathtt{g}}$ and $U_{12} = 0_{\mathtt{g}\times\mathtt{t}}$, hence, $U = \begin{bmatrix} I_{\mathtt{g}} & 0_{\mathtt{g}\times\mathtt{t}} \\ U_{21} & U_{22} \end{bmatrix}$. From $0 = U_{21}R_1 + U_{22}C_1'$ and assumption that $\mathcal{C}$ act only on the w_2 variables then $U_{22} = -I_{\mathtt{t}}$, therefore, $U = \begin{bmatrix} I_{\mathtt{g}} & 0_{\mathtt{g}\times\mathtt{t}} \\ U_{21} & -I_{\mathtt{t}} \end{bmatrix}$. Define $U' := U_{21}$, furthermore, define $C_2 := U'R_2 - C_2'$.

Now to show (7.7), notice that

$$0_{\mathtt{t}\times\mathtt{m}} = \begin{bmatrix} U' & -I_{\mathtt{t}} \end{bmatrix} \begin{bmatrix} R_1 \\ C_1' \end{bmatrix}.$$

Let $\mathcal{M}$ be the $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times(\mathtt{g}+\mathtt{t})}[\xi]$ generated by the rows of $\begin{bmatrix} U' & -I_{\mathtt{t}} \end{bmatrix}$. Since all $\mathbb{R}[\xi]$-submodules of $\mathbb{R}^{1\times\bullet}[\xi]$ are finitely generated and have a basis, then $\mathcal{M}$ has basis generators $\alpha_1, \ldots, \alpha_{\mathtt{t}}$. Define $G := \mathrm{col}(\alpha_1, \ldots, \alpha_{\mathtt{t}})$, and partition $G = \begin{bmatrix} G_1 & G_2 \end{bmatrix}$ with respect to the rows of R_1 and C_1'. Compute,

$$\begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} R_1 & R_2 \\ C_1' & C_2' \end{bmatrix} = \begin{bmatrix} 0_{\mathtt{t}\times\mathtt{m}} & G_1 R_2 + G_2 C_2' \end{bmatrix},$$

define $C_2 := G_1 R_2 + G_2 C_2'$. Hence, $\mathcal{C} = \ker\left(\begin{bmatrix} 0_{\mathtt{t}\times\mathtt{m}} & C_2(\sigma) \end{bmatrix}\right)$. □

It follows from Proposition 7.10 that to find $C = \begin{bmatrix} C_1 & 0 \end{bmatrix}$ in Theorem 7.8 it is sufficient to find a matrix G whose rows are a basis of the left syzygy of $\mathrm{col}(-I, D_2)$, then $C = GK$. We illustrate the above procedure with an example.

Example 7.3. Let the to-be controlled system with a kernel representation

$$\begin{bmatrix} \sigma + \frac{1}{2} & 1 & -\sigma - \frac{1}{4} & 0 \\ 0 & \sigma + \frac{1}{3} & 1 & -\sigma - \frac{1}{3} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_2 \\ w_3 \end{bmatrix} = 0 \tag{7.8}$$

be interconnected with a controller

$$\begin{bmatrix} -\frac{603}{5000} - \frac{2411\sigma}{10000} & -\frac{2417}{10000} - \frac{17\sigma}{10000} & -\frac{693}{10000} - \frac{2629\sigma}{5000} & -\frac{3833}{5000} + \frac{17\sigma}{10000} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0.$$

Hence, the controlled behaviour admits a kernel representation

$$\underbrace{\begin{bmatrix} \sigma + \frac{1}{2} & 1 & -\sigma - \frac{1}{4} & 0 \\ 0 & \sigma + \frac{1}{3} & 1 & -\sigma - \frac{1}{3} \\ -\frac{603}{5000} - \frac{2411\sigma}{10000} & -\frac{2417}{10000} - \frac{17\sigma}{10000} & -\frac{693}{10000} - \frac{2629\sigma}{5000} & -\frac{3833}{5000} + \frac{17\sigma}{10000} \end{bmatrix}}_{:=D} \begin{bmatrix} w_1 \\ w_2 \\ w_2 \\ w_3 \end{bmatrix} = 0.$$

We aim to find a kernel representation of the controller acting only on the variables w_3 and w_4. Define the matrix K as the first two columns of the matrix D, i.e.

$$K := \begin{bmatrix} \sigma + \frac{1}{2} & 1 \\ 0 & \sigma + \frac{1}{3} \\ -\frac{603}{5000} - \frac{2411\sigma}{10000} & -\frac{2417}{10000} - \frac{17\sigma}{10000} \end{bmatrix}.$$

Using `Singular` we compute the left syzygy of K as

$$G = \begin{bmatrix} \frac{2411\sigma}{10000} + \frac{1391}{4865} & \frac{17\sigma}{10000} + \frac{3}{5000} & \sigma + \frac{1}{3} \end{bmatrix}.$$

Consequently, by left multiplying D by G we get a kernel representation of the controller acting only on w_3 and w_4 as

$$\begin{bmatrix} 0 & 0 & -\frac{329\sigma^2}{429} - \frac{767\sigma}{2000} - \frac{123}{2888} & -\frac{959\sigma}{1250} - \frac{1200}{4693} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = 0.$$

□

7.3.2 Flatness and partial interconnection

Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$, in the following result we give necessary and sufficient conditions for (w, c) to be a flat partition with c as the flat variable.

Proposition 7.11. *Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$. c is a flat variable if and only if w is observable from c and c is free.*

Proof. The proof follows from Definition 7.3. □

From Proposition 7.11, clearly not all $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$ are flat as we discuss. Let $\mathcal{P} = \Pi_w(\mathcal{P}_{full})$ and $\mathcal{N} \in \mathscr{L}^{\mathtt{w}}$ defined by (3.2) be such that $\mathcal{N} \subseteq \mathcal{P}$. If we assume that $\mathcal{N}$ is not trivial then there exists $(w, 0) \in \mathcal{P}_{full}$ with non-zero w. In this case w is not observable from c. Therefore, flatness for $\mathcal{P}_{full}$ with flat partition (w, c) and c as the flat variable is only possible in the case $\mathcal{N} = \{0\}$, henceforth we focus on full information control. Recall that under full information control the to-be-controlled variables w are observable from the control variables c. Therefore, $\mathcal{P}_{full}$ flat with c as the flat variable is a necessary condition for full information control. Now we state under which condition does full information imply flatness.

Theorem 7.12. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ and $M \in \mathbb{R}^{\mathtt{g}\times\mathtt{c}}[\xi]$ be such that $R(\sigma)w = M(\sigma)c$ induces a minimal hybrid representation of $\mathcal{P}_{full}$. If $\mathtt{w} = \mathtt{g}$ and w is observable from c then $\mathcal{P}_{full}$ is flat.*

Proof. Let $\mathtt{w} = \mathtt{g}$ then R is square, furthermore, since w is observable from c then R admits an inverse. Consequently, $\mathcal{P}_{full}$ admits a minimal kernel representation

$$\begin{bmatrix} F(\sigma) & -I \end{bmatrix} \begin{bmatrix} c \\ w \end{bmatrix} = 0,$$

where $F := R^{-1}M$. Therefore, it follow from Theorem 7.5 that $\mathcal{P}_{full}$ is flat and c is the flat variable. □

Under the conditions of Theorem 7.12 we now show a characterization of kernel representations of $\mathcal{C}$, a controller that implements some controlled behaviour via partial interconnection with c.

Theorem 7.13. *Let $\mathcal{P}_{full} \in \mathscr{L}^{\mathtt{w}+\mathtt{c}}$ be flat with c as the flat variable then there exists $R' \in \mathbb{R}^{\mathtt{g}\times\mathtt{c}}[\xi]$ such that*

$$\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix} \begin{bmatrix} c \\ w \end{bmatrix} = 0$$

induces a minimal representation of $\mathcal{P}_{full}$. Furthermore, let $R_1 \in \mathbb{R}^{\mathtt{g}\times\mathtt{c}}[\xi]$ and $R_2 \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ with R_2 unimodular be such that

$$\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R_1(\sigma) & R_2(\sigma) \end{bmatrix}\right)$$

induces a minimal representation as well. Let a full row rank matrix $D \in \mathbb{R}^{\bullet\times\mathtt{w}}[\xi]$ be such that $\mathcal{K} = \ker(D(\sigma))$. Assume that $\mathcal{K}$ is implementable by partial interconnection through c with respect to $\mathcal{P}_{full}$. Then the following statements are equivalent:

1. *$C \in \mathbb{R}^{\bullet\times\mathtt{c}}[\xi]$, is such that $\mathcal{C} = \ker(C(\sigma))$ implements $\mathcal{K}$,*
2. *$< C > = < DR' >$,*
3. *$< C > = < DR_2^{-1}R_1 >$.*

Proof. The existence of R', R_1 and R_2 such that

$$\mathfrak{B}_{full} = \ker\left(\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix}\right) = \ker\left(\begin{bmatrix} R_1(\sigma) & R_2(\sigma) \end{bmatrix}\right)$$

follows from Theorem 7.5. Now recall that if $\mathcal{K}$ is implementable via partial interconnection of $\mathcal{P}_{full}$ with $\mathcal{C}$ then $\mathcal{K}$ is defined by (3.1). We start with the equivalence between (1) and (2). (2) $\Longrightarrow$ (1) is as follows. Assume that $< C > = < DR' >$. Let $(w, c) \in \mathcal{P}_{full}$ such that $c \in \ker(DR'(\sigma))$ then $DR'(\sigma)c = 0$. Since $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix}\right)$ then

$$\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix} \begin{bmatrix} c \\ w \end{bmatrix} = 0,$$

hence, $w = R'(\sigma)c$. Substitute $R'(\sigma)c = w$ into $DR'(\sigma)c = 0$ then $D(\sigma)w = 0$. Now since $\mathcal{K} = \ker(D(\sigma))$ then $w \in \mathcal{K}$. Consequently, $(w, c) \in \mathcal{P}_{full}$ is such that $c \in \ker(C(\sigma))$ and $w \in \mathcal{K}$, hence, $\mathcal{C}$ implements $\mathcal{K}$.
(1) $\Longrightarrow$ (2) Assume that $\mathcal{C}$ implements $\mathcal{K}$, we need show that $\mathcal{C} = \ker(DR'(\sigma))$ then it follows directly that $< C > = < DR' >$. We start with the inclusion $\mathcal{C} \subseteq \ker(DR'(\sigma))$. Let $(w, c) \in \mathcal{P}_{full}$ such that $c \in \mathcal{C}$ and $w \in \mathcal{K}$. Now since $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix}\right)$ then

$$\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix} \begin{bmatrix} c \\ w \end{bmatrix} = 0,$$

hence, $w = R'(\sigma)c$. Moreover, since $\mathcal{K} = \ker(D)(\sigma)$ then $D(\sigma)w = 0$. Substitute $w = R'(\sigma)c$ into $D(\sigma)w = 0$ then $DR'(\sigma)c = 0$. Therefore, $c \in \ker(DR')(\sigma)$ which implies that $\mathcal{C} \subseteq \ker(DR'(\sigma))$. Now to show the inclusion $\mathcal{C} \supseteq \ker(DR'(\sigma))$, let $c \in \ker(DR'(\sigma))$ and $w \in \ker(D(\sigma))$ then $DR'(\sigma)c = 0$ and $D(\sigma)w = 0$. Therefore, $DR'(\sigma)c = D(\sigma)w$ which implies that $R'(\sigma)c = w$. Hence,

$$\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix} \begin{bmatrix} c \\ w \end{bmatrix} = 0.$$

Now since $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix}\right)$ then $(w, c) \in \mathcal{P}_{full}$. Since $c \in \ker(DR'(\sigma))$ then $\ker(DR'(\sigma))$ implements $\mathcal{K}$. Now by the assumption that $\mathcal{C}$ implements $\mathcal{K}$ then there exists $c' \in \mathcal{C}$ such that $(w, c') \in \mathcal{P}_{full}$. From $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R'(\sigma) & -I_{\mathtt{w}} \end{bmatrix}\right)$ then $w = R'(\sigma)c'$ and $w = R'(\sigma)c$. Hence $R'(\sigma)c = R'(\sigma)c'$, therefore, $c = c'$ which implies that $c \in \mathcal{C}$. Consequently, $\mathcal{C} \supseteq \ker(DR'(\sigma))$.
(1) $\iff$ (3) Let R_2^{-1} be such that $R_2^{-1}R_2 = I_{\mathtt{w}}$ then

$$R_2^{-1}\begin{bmatrix} R_1(\sigma) & R_2(\sigma) \end{bmatrix} = \begin{bmatrix} R_2^{-1}R_1(\sigma) & I_{\mathtt{w}} \end{bmatrix}.$$

Since $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R_1(\sigma) & R_2(\sigma) \end{bmatrix}\right)$ induces a minimal representation and R_2^{-1} is unimodular, then $\mathcal{P}_{full} = \ker\left(\begin{bmatrix} R_2^{-1}R_1(\sigma) & I_{\mathtt{w}} \end{bmatrix}\right)$. The rest of the proof follows from the same argument used to prove (1) $\iff$ (2). □

It follows from Theorem 7.13, that given a desired behaviour that is to be implemented via partial interconnection with a flat system with the control variable flat, finding a controller amounts to finding a basis of a module generated DR'.

7.4 Parametrizations

In this section, we show a parametrization of the image representation of all implementable $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}_{contr}$. In the literature, parametrizations has been studied in: [33], where a parametrization of all stabilizing controllers has been presented for the case of full interconnection; and [51] where the problem of parametrization of all controllers and stabilizing controllers has been solved for both full and partial interconnection.

Let $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $\mathcal{K} \subseteq \mathcal{P}$, we consider a case where the manifest controlled behavior $\mathcal{K}$ is specified by restricting the latent variables, i.e.

$$\begin{aligned} w &= M(\sigma)\ell, \\ C(\sigma)\ell &= 0 \end{aligned}$$

where $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ induces an observable image representation of $\mathcal{P}$ and $C \in \mathbb{R}^{\mathtt{k}\times\mathtt{l}}[\xi]$ induces a minimal kernel representation of the controller acting on the latent variables of $\mathcal{P}$.

Theorem 7.14. *Let $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ induce an observable image representation of $\mathcal{P}$, $\mathtt{l} > \mathtt{k}$ and $C \in \mathbb{R}^{\mathtt{k}\times\mathtt{l}}[\xi]$ be full row rank for all $\lambda \in \mathbb{C}$. Assume that there exists a unimodular matrix $G \in \mathbb{R}^{\mathtt{l}\times\mathtt{l}}[\xi]$ such that $CG = \begin{bmatrix} I_{\mathtt{k}} & 0_{\mathtt{l}-\mathtt{k}} \end{bmatrix}$. Define*

$$MG =: \begin{bmatrix} M_1 & M_2 \end{bmatrix}$$

with $M_2 \in \mathbb{R}^{\mathtt{w}\times(\mathtt{l}-\mathtt{k})}[\xi]$ and assume that $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}_{contr}$ is implementable. Then the following statements are equivalent:

1. *$\mathcal{K}$ is defined by*

$$\mathcal{K} := \{w \mid \exists\ \ell \text{ such that } w = M(\sigma)\ell,\ C(\sigma)\ell = 0\},$$

2. *M_2 induces an image representation of $\mathcal{K}$,*
3. *there exist $K \in \mathbb{R}^{\mathtt{l}\times(\mathtt{l}-\mathtt{k})}[\xi]$, such that $K(\lambda)$ is full column rank for all $\lambda \in \mathbb{C}$, such that $im(K(\sigma)) = \ker(C(\sigma))$ and $\mathcal{K} = im((MK)(\sigma))$*

Proof. We start with the equivalence between (1) and (2). (1) $\Longrightarrow$ (2) Define $\ell' := G^{-1}(\sigma)\ell$ then

$$\mathcal{K} = \left\{w \mid w = MG(\sigma)\ell',\ \begin{bmatrix} I_{\mathtt{k}} & 0_{\mathtt{l}-\mathtt{k}} \end{bmatrix}(\sigma)\ell' = 0\right\}.$$

Now partition $\ell' = \mathrm{col}(\ell_2', \ell_1')$ with respect to the partition of $\begin{bmatrix} I_{\mathtt{k}} & 0_{\mathtt{l}-\mathtt{k}} \end{bmatrix}$ then

$$\begin{bmatrix} I_{\mathtt{k}} & 0_{\mathtt{l}-\mathtt{k}} \end{bmatrix}(\sigma)\begin{bmatrix} \ell_1' \\ \ell_2' \end{bmatrix} = 0,$$

which implies that $\ell_1' = 0$. Consequently,

$$\begin{aligned} w &= MG(\sigma)\begin{bmatrix} 0 \\ \ell_2' \end{bmatrix} \\ &= \begin{bmatrix} M_1 & M_2 \end{bmatrix}(\sigma)\begin{bmatrix} 0 \\ \ell_2' \end{bmatrix}, \end{aligned}$$

therefore, $\mathcal{K} = \mathrm{im}(M_2(\sigma))$.

(2) $\Longrightarrow$ (1) Since $\mathcal{P} = \mathrm{im}(M(\sigma))$ then $w = M(\sigma)\ell$ for some latent variable ℓ. Now since $MG = \begin{bmatrix} M_1 & M_2 \end{bmatrix}$ then

$$w = \begin{bmatrix} M_1 & M_2 \end{bmatrix} G^{-1}(\sigma)\ell.$$

Define $G^{-1}(\sigma)\ell =: \mathrm{col}(\ell_1'', \ell_2'')$ with respect to the partitions of $\begin{bmatrix} M_1 & M_2 \end{bmatrix}$. Since $w = M_2(\sigma)\ell'$ for some latent variable ℓ' then for all $w \in \mathcal{K}$

$$\begin{aligned} \begin{bmatrix} M_1 & M_2 \end{bmatrix}(\sigma)\begin{bmatrix} \ell_1'' \\ \ell_2'' \end{bmatrix} &= M_2(\sigma)\ell' \\ M_1(\sigma)\ell_1'' + M_2(\sigma)\begin{bmatrix} \ell_2'' - \ell' \end{bmatrix} &= 0 \\ \begin{bmatrix} M_1 & M_2 \end{bmatrix}(\sigma)\begin{bmatrix} \ell_1'' \\ \ell_2'' - \ell' \end{bmatrix} &= 0 \end{aligned}$$

Since M induces an observable image representation then $\begin{bmatrix} M_1 & M_2 \end{bmatrix}$ admits a left inverse which implies that

$$\begin{bmatrix} \ell_1'' \\ \ell_2'' - \ell' \end{bmatrix} = 0.$$

Hence, $\ell_1'' = 0$ and $\ell_2'' = \ell'$. Therefore,

$$G^{-1}(\sigma)\ell = \begin{bmatrix} 0 \\ \ell' \end{bmatrix}.$$

Partition the rows of G^{-1} appropriately with respect to the partition of $\mathrm{col}(0, \ell')$, i.e $G^{-1} = \mathrm{col}(G_1^{-1}, G_2^{-1})$. then

$$\begin{bmatrix} G_1^{-1} \\ G_2^{-1} \end{bmatrix}(\sigma)\ell = \begin{bmatrix} 0 \\ \ell' \end{bmatrix}.$$

Define $C := G_1^{-1}$ then

$$\mathcal{K} := \{w \mid w = M(\sigma)\ell,\ C(\sigma)\ell = 0\}.$$

(1) $\iff$ (3) Since $CG = \begin{bmatrix} I_{\mathtt{k}} & 0_{\mathtt{1-k}} \end{bmatrix}$ partition $G = \begin{bmatrix} G_1 & G_2 \end{bmatrix}$ with $G_1 \in \mathbb{R}^{\mathtt{1}\times\mathtt{k}}[\xi]$ and $G_2 \in \mathbb{R}^{\mathtt{1}\times(\mathtt{1-k})}[\xi]$ then $CG_2 = 0$. Define $K := G_2$ then it follows from Proposition 2.18 that $\mathrm{im}(K(\sigma)) = \ker(C(\sigma))$. The proof of (1) $\Longrightarrow$ (3) is as follows. Let the manifest behaviour of $\mathcal{K}$ be

$$\begin{aligned} w &= M(\sigma)\ell, \\ C(\sigma)\ell &= 0. \end{aligned}$$

Since $\mathrm{im}(K(\sigma)) = \ker(C(\sigma))$ then $\ell = K(\sigma)\ell'$, therefore, $w = MK(\sigma)\ell'$.
(3) $\Longrightarrow$ (1) Let $\ell = K(\sigma)\ell'$. Now since $\mathcal{K} = \mathrm{im}(MK(\sigma))$ then $w = MK(\sigma)\ell'$ implies that $w = M(\sigma)\ell$. Since $\mathrm{im}(K(\sigma)) = \ker(C(\sigma))$ then $C(\sigma)\ell = 0$ then the manifest behavior of

$\mathcal{K}$ is given by

$$\begin{aligned} w &= M(\sigma)\ell, \\ C(\sigma)\ell &= 0. \end{aligned}$$

□

Assume that $\mathcal{P}$ is flat with (w_1, w_2) as the flat partition of the variables $\mathcal{P}$ and w_1 the flat variable. We show a characterization of a controller $\mathcal{C}$ that implements $\mathcal{K}$ in Theorem 7.14.

Theorem 7.15. *Under the assumptions of Theorem 7.14, let $w = \text{col}(w_1, w_2)$ be a flat partition with w_1 flat. Then statements (1)-(3) in Theorem 7.14 are equivalent with,*

4. let $C' \in \mathbb{R}^{\mathtt{k}\times\mathtt{l}}[\xi]$. Then there exist a unimodular matrix $U \in \mathbb{R}^{\mathtt{l}\times\mathtt{l}}[\xi]$ such that $C = C'U$ and

$$\begin{bmatrix} C'(\sigma) & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0$$

implements $\mathcal{K}$.

Proof. Since $\mathcal{P} = \text{im}(M(\sigma))$ where M induces an observable representation then

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = M(\sigma)\ell = \begin{bmatrix} M_1(\sigma) \\ M_2(\sigma) \end{bmatrix} \ell.$$

Now since w_1 is flat it follows from Theorem 7.5 that M_1 is unimodular, which implies $\ell = M_1^{-1}(\sigma)w_1$.

Now we prove (1) $\iff$ (4). Since $C(\sigma)\ell = 0$ then $(CM_1^{-1})(\sigma)w_1 = 0$. Now Define $C' := CM_1^{-1}$ then

$$C'(\sigma)w_1 = 0$$

which implies

$$\begin{bmatrix} C'(\sigma) & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0$$

implements $\mathcal{K}$. Now define $M_1 =: U$ which completes our proof. □

In Theorem 7.17 we show another parametrization of the image representation of controllable controlled behaviour. First consider the following standard result about left prime matrices.

Lemma 7.16. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ be left prime. Then there exists a matrix $R' \in \mathbb{R}^{(\mathtt{w}-\mathtt{g})\times\mathtt{w}}[\xi]$ such that col(R, R') is unimodular.*

Proof. Follows from Lemma 8.1.20 p.290 of [77]. □

See Algorithm 2.1 p.2183 of [71] on how to compute R' in Lemma 7.16.

Theorem 7.17. *Let full row rank $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ and $C \in \mathbb{R}^{\mathtt{c}\times\mathtt{w}}[\xi]$ be such that $\mathcal{P} = \ker(R(\sigma))$, $\mathcal{C} = \ker(C(\sigma))$. Assume that $\mathcal{P} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and that $\mathcal{C}$ regularly implements $\mathcal{K} \in \mathscr{L}^{\mathtt{w}}_{contr}$ via full interconnection with $\mathcal{P}$. Then, there exists a matrix $S \in \mathbb{R}^{(\mathtt{w}-\mathtt{t})\times\mathtt{w}}[\xi]$, with $\mathtt{t} = \mathtt{c} + \mathtt{g}$, such the matrix $T \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\xi]$ defined by $T := \mathrm{col}(R, C, S)$ is unimodular. Let $K \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\xi]$ such that $TK = I_{\mathtt{w}}$ and partition $K = row(N, Q, M)$ such that $N \in \mathbb{R}^{\mathtt{w}\times\mathtt{g}}[\xi], Q \in \mathbb{R}^{\mathtt{w}\times(\mathtt{t}-\mathtt{g})}$ and $M \in \mathbb{R}^{\mathtt{w}\times(\mathtt{w}-\mathtt{t})}[\xi]$ are all full column rank for all $\lambda \in \mathbb{C}$. Hence,*

$$\begin{bmatrix} R \\ C \\ S \end{bmatrix} \begin{bmatrix} N & Q & M \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}. \tag{7.9}$$

Then $\mathcal{P} = im([Q(\sigma)\ M(\sigma)])$ and $\mathcal{K} = im(M(\sigma))$.

Proof. Since $\mathcal{C}$ regularly implements $\mathcal{K}$ and both R, C are full row rank then $\mathrm{col}(R, C)$ is full row rank, see Proposition 3.7. Moreover, since $\mathcal{K}$ is controllable then $\mathrm{col}(R, C)$ is left prime. Therefore, the existence of S such that T is unimodular follows from Lemma 7.16. To show that $\mathcal{P} = \mathrm{im}([Q(\sigma)\ M(\sigma)])$ from (7.9) since $R(\begin{bmatrix} Q & M \end{bmatrix}) = 0$ and R induces a minimal kernel representation of $\mathcal{P}$ then it follows from Proposition 2.18 that $\mathrm{row}(Q, M)$ induces an image representation of $\mathcal{P}$. Now since

$$\begin{bmatrix} R \\ C \end{bmatrix} M = 0,$$

and $\begin{bmatrix} R \\ C \end{bmatrix}$ induces a minimal kernel representation of $\mathcal{K}$ then it follows from Proposition 2.18 that M induces a minimal image representation of $\mathcal{K}$. □

We end this section by showing a parametrization of all controllable controller that implements controllable controlled behaviours. Recall that in Theorem 5.8 both $\mathcal{C}$ and $\mathcal{D}$ were not required to be elements of $\mathscr{L}^{\mathtt{c}}_{contr}$ and $\mathscr{L}^{\mathtt{w}}_{contr}$, respectively. In the next result we consider when both $\mathcal{C} \in \mathscr{L}^{\mathtt{c}}_{contr}$ and $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$.

Theorem 7.18. *Let $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and assume that $M_1 \in \mathbb{R}^{\mathtt{w}\times\mathtt{l}}[\xi]$ induce an image representation of $\mathcal{D}$. Under the assumptions of Theorem 5.8, then the following statements are equivalent:*

1. *$\mathcal{C} \in \mathscr{L}^{\mathtt{c}}_{contr}$ implements $\mathcal{D}$,*
2. *$\mathcal{C} = im((NRM_1)(\sigma))$.*

Proof. Let $\mathcal{P}_{full}$, Y, N, D, G and R as in Theorem 5.8.
(1) $\Longrightarrow$ (2) To show $\mathcal{C} \subseteq \text{im}((NRM_1)(\sigma))$ let $(w,c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$ and $c \in \mathcal{C}$. Since c is observable from w then

$$\begin{aligned} c &= Y(\sigma)w \\ &= (NR + GD)(\sigma)w \end{aligned}$$

Now since $\mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$, then $w = M_1(\sigma)\ell$ for some latent variable ℓ, consequently,

$$c = (NRM_1 + GDM_1)(\sigma)\ell.$$

Now since

$$\mathcal{D} = \ker(D(\sigma)) = \text{im}(M_1(\sigma))$$

then it follows from Proposition 2.18 that $DM_1 = 0$, hence, $c = (NRM_1)(\sigma)\ell$. To show the converse inclusion, let $(w,c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$ and $c \in \text{im}((NRM_1)(\sigma))$. Since c corresponds to $w \in \mathcal{D}$ and $\mathcal{C}$ implements $\mathcal{D}$ then it follows from Proposition 4.10 $c \in \mathcal{C}$. Hence, $\mathcal{C} \supseteq \text{im}((NRM_1)(\sigma))$.
(2) $\Longrightarrow$ (1) Let $c = (NRM_1)(\sigma)\ell$. Since $\mathcal{D} = \text{im}(M_1(\sigma))$ then there exists $w \in \mathcal{D}$ such that

$$\begin{aligned} c &= (NR)(\sigma)w \\ &= (NR)(\sigma)w + 0 \\ &= (NR)(\sigma)w + (GD)(\sigma)w, \end{aligned}$$

such that $(w,c) \in \mathcal{P}_{full}$. From Proposition 4.10 $(w,c) \in \mathcal{P}_{full}$ such that $w \in \mathcal{D}$ then $c \in \mathcal{C}$. Therefore, $\mathcal{C}$ implements $\mathcal{D}$. □

7.5 Summary

We have studied flat systems, presented a trajectory based definition of linear flat systems. Using this new definition we characterized image and kernel representations of flat system and shown how to find flat outputs given an observable image representation of a flat system. We then characterized a kernel representation of a controller interconnected with a flat system via full interconnection. We have shown how to find the kernel representation such that the controller only acts on the flat variables. In the case of partial interconnection we showed under which conditions are the control variables flat. Then under such conditions we gave a parametrization of all controllers that acts of on the control variable. Finally, we gave a parametrization of of the image representations of all controllable controlled behaviours.

Chapter 8

Conclusions and future work

We have developed new results on data-driven control and flat systems. Our data-driven results are such that the data is used to find a representation for a controller that implements the desired behaviour. We have also shown a parametrization of all controllable controllers and image representations of controlled systems. Finally, we proposed a trajectory-based definition of flat systems; based this definition we showed a characterization of the kernel and image representation of flat systems. We also studied the control of flat systems using interconnection.

In Chapters 2 and 3, we focused mainly on setting up the tools, i.e. definitions and mathematical results necessary for the results developed in Chapters 4,5,6 and 7. Therefore, the material covered in Chapters 2 and 3 is taken from various sources in the literature, all of which have been referenced within the chapters. In Chapter 2, we covered concepts of the behavioural framework for linear time-invariant difference systems. Chapter 3 is dedicated to the theory of control as interconnection. We discussed the two types of interconnections namely full and partial interconnections. We also studied the notion of implementability, and regular interconnections.

Chapter 4: Using partial interconnection, we developed a data-driven solution to the problem of finding control variables such that the to-be-controlled variables follows a prescribed path for a given finite time interval. The solution is characterized as follows.

- In Theorem 4.3, we showed how to compute a trajectory of the desired behaviour using a trajectory of the to-be-controlled system. Then in Theorem 4.6, we proved sufficient condition under which such a trajectory computed in Theorem 4.3 is sufficiently informative about the desired behaviour.
- Under the assumptions of Theorem 4.6, in Theorem 4.9 we showed how to find a trajectory of the desired behaviour that is equal to the given prescribed path for the given time interval. Then, under the assumption that the control variables

are observable from the to-be-controlled ones, we compute the control variable trajectory that corresponds to it. Finally, in Lemma 4.11 we prove that if a controller implements the desired behaviour then the control variable trajectory belongs to such a controller.

- Our solution to the prescribed path data-driven control problem is summarised in Algorithm 3 on p. 46.

Chapter 5: We covered two completely data-driven control problems. In both problems, we showed how to find a controller when given trajectories from the to-be-controlled system and the desired behaviour. For the first solution we use partial interconnection and is summarised as follow.

- In Theorem 5.8, under the assumption that the control variables are observable from the to-be-controlled variables we showed a parametrization of an observability map such that we can construct control variable trajectories from the to-be-controlled variable trajectories. Then in Lemma 5.9 we used the Hankel matrices associated with observed control and to-be-controlled variable trajectories of the to-be-controlled system to find a real matrix O such that we can compute the control variable trajectory corresponding to the given trajectory of the desired behaviour.

- In Proposition 5.11, we prove that if a controller implements the desired behaviour then control variable trajectory corresponding to the given trajectory of the desired behaviour computed using O belongs to such a controller. Then in Theorem 5.12, we prove sufficient conditions under which such control variable trajectory is sufficiently informative about the controller behaviour.

- Finally, under the assumption Theorem 5.12 the data-driven partial interconnection solution is summarised in Algorithm 4 on p. 60.

The full interconnection solution is summarised as follows.

- In Theorem 5.20, we proved necessary and sufficient conditions suitable for finding a controller from data under full interconnection.

- In Algorithm 6 on p. 66, under the assumption of Theorem 5.20 and Lemma 5.19 we showed how to find a representation of the controller .

- The correctness of Algorithm 6 is proved in Proposition 5.21.

Chapter 6: We presented some results on data-driven dissipativity.

- In Theorem 6.10, we proved necessary and sufficient conditions under which one can determine whether a system is half-line dissipative with respect to a given supply rate using observed trajectories.
- We also introduce the notion of L-dissipativity for finite length observed trajectories.

Chapter 7: We studied the notion of flat systems and parametrizations in behavioural setting. We proposed an alternative definition of flatness then characterized flat system based on the definition. we also studied control of flat system via control as interconnections. Finally we parametrized image representations of all controllable controlled system and controllable controllers.

- We presented a trajectory based definition of flat systems, see Definition 7.3, and a characterization of flat systems in Theorem 7.5. We showed that flat systems admits an observable image and minimal kernel representations with a unimodular submatrix. Hence, finding flat outputs amounts to finding the unimodular submatrix of the kernel or image representation and the flat outputs are the manifest variables, corresponding to the unimodular submatrix.
- We then studied control of flat systems. In Theorem 7.8 we showed a characterization of a kernel representation of controllers that interconnect with flat system via full interconnection. We showed that using the results of Proposition 7.10 one can always find a kernel presentation of the controller such that the controller is only acting on the flat variables.
- Theorem 7.11, under partial interconnection, we proved necessary and sufficient conditions for control variables to be flat variables. We then showed that having control variables as flat outputs, is a necessary condition for full information control.
- In Theorem 7.12, we stated sufficient condition for full information control to imply flatness.
- In Theorem 7.13, under the conditions of Theorem 7.12, we showed a characterization of kernel representations of controllers that implements some controlled behaviour via partial interconnection with a flat system.
- In Theorem 7.14, we showed a parametrization of the image representation of the controlled behaviour in the case where the manifest behaviour of the controlled behaviour is specified by restricting latent variables of the to-be-controlled system behaviour.
- In Theorem 7.18, under the assumptions of Theorem 5.8, we presented a parametrization of the image representation of all controllable controllers that implements controllable desired behaviours.

8.1 Future work

In this section we outline possible future research direction.

Consistency. Throughout this thesis, our data-driven control results are based on the assumption that the given trajectories are of infinite length, but in practice trajectories are of finite length. As already mentioned in Remark 4.2 the issue of consistency still remains an important open question for future research.

Let $t, N \in \mathbb{Z}_+$ and a finite length N trajectory w. Recall that, a procedure is called *consistent* if the system identified using $w_{|[t,t+N]}$, denoted by $\mathfrak{B}^*_{|[t,t+N]}$, converge to the "true system", $\mathfrak{B}^*$, as $N \longrightarrow \infty$. Therefore, we need to study and develop sufficient conditions for the procedures developed in Algorithms 4 and 6 to be consistent.

The issue of convergence to a true system in a behavioural setting, i.e. without restriction to a particular represenation, has been addressed in [20]. Using the *gap matrix*, $d(\mathfrak{B}_1, \mathfrak{B}_2)$, which measure distance between behaviours, in Theorem 8. p. 1000 of [20] the authors shows that the Global Total Least Square is consistent as almost surely for $N \longrightarrow \infty$ then it holds that $d(\mathfrak{B}^*_{|[t,t+N]}, \mathfrak{B}^*) \longrightarrow 0$. Note that the gap between behaviours has close connection between angle between behaviours, covered in Chapter 5 section 5.2.1, see p. 200 of [59]. Therefore, a possible starting point would be show the consistence of Algorithms 4 and 6 using angle between behaviours.

In Algorithms 4, the observed data is used to a compute control variable trajectory, using an empirical observability map, which is then used to find a controller. Hence, it will be worth studying under which condition does control variable trajectory computed using an empirical observability map implies that as $N \longrightarrow \infty$ the controller will converge to a "true controller." We believe a possible solution would be to prove the following.

Conjecture 8.1. *Let N and t as before. Assume that $\mathcal{P}^*$ and $\mathcal{D}^*$ are the true to-be-controlled and desired behaviours, respectively. Futhermore, assume that $\mathcal{P}^*_{|[t,t+N]}$ and $\mathcal{D}^*_{|[t,t+N]}$ are the identified to-be-controlled and desired behaviours from finite length $\tilde{w}$ and $\tilde{w}_d$, respectively. Algorithms 4 and 6 are consistent if it hold that as $N \longrightarrow \infty$ then $\mathcal{P}^*_{|[t,t+N]}$ and $\mathcal{D}^*_{|[t,t+N]}$ converges to $\mathcal{P}^*$ and $\mathcal{D}^*$, respectively.*

Extension to 2D systems. 2D systems have been studied in a behavioural setting, for a class of *linear shift-invariant systems*, see pp.11-13 of [56]. Ibid. a 2D system is defined by a triple $\Sigma := (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, where $\mathbb{T} \in \mathbb{Z}^2$ is the *index set*, $\mathbb{W}$ the signal space and $\mathfrak{B} \subseteq \mathbb{W}^{\mathbb{T}}$ the behaviour. Necessary and sufficient conditions for linearity, shift invariance and completeness of Σ are given in Proposition 5. p. 13 of [56]. The class of linear shift-invariant complete 2D systems with $\mathtt{w}$ variables is denoted by $\mathbb{L}^{\mathtt{w}}$.

Let the *backward shift* operator

$$\sigma_1 : \mathbb{W}^{\mathbb{T}} \to \mathbb{W}^{\mathbb{T}}$$

be defined by

$$\sigma_1 w(t_1, t_2) := w(t_1 + 1, t_2)$$

and the *downward shift* operator

$$\sigma_2 : \mathbb{W}^{\mathbb{T}} \to \mathbb{W}^{\mathbb{T}}$$

be defined by

$$\sigma_2 w(t_1, t_2) := w(t_1, t_2 + 1)$$

It can be shown that 2D systems admits a kernel representation of the form

$$R(\sigma_1, \sigma_2)w = 0$$

where $R \in \mathbb{R}^{\mathtt{g} \times \mathtt{w}}[\xi_1, \xi_2]$ and $\mathbb{R}^{\mathtt{g} \times \mathtt{w}}[\xi_1, \xi_2]$ is the ring of polynomial in the indeterminate ξ_1 and ξ_2.

The author of [56] also studies the notions of controllability and observability for 2D systems. Control as interconnection has been studied in [57], where the notion of regular full interconnection for 2D systems has been addressed, see Lamma 3.3 p.115. In Lemma 3.5 p.117 of [56], condition of implementability by regular interconnection are given. The partial interconnection has been studied in [73].

Based on the conditions of In Lemma 3.5 p.117 of [56]. Let $\mathcal{P}, \mathcal{D} \in \mathbb{L}^{\mathtt{w}}$ be the to-be-controlled system and desired behaviours. Consider the following 2D data-driven control full interconnection problem.

Problem 8.2. 2D full interconnection case. *Assume that*

1. *$\mathcal{P}$ is controllable in the 2D sense*

Given

i. $\tilde{w} \in \mathcal{P}$,

ii. $\tilde{w}_d \in \mathcal{D}$.

Find, if it exists, a controller $\mathcal{C}$ that implements $\mathcal{D}$ via regular full interconnection with $\mathcal{P}$.

We believe in order to find a solution to the above 2D full interconnection problem, one needs to address the following:

1. Define the notion of sufficiently informative trajectories in the case of 2D systems.
2. Then provide a characterization of sufficiently informative.

A possible starting point would be to consider $\tilde{w} = \tilde{w}(0,0), \tilde{w}(1,0), \tilde{w}(0,1), \tilde{w}(1,1), \tilde{w}(2,0), \tilde{w}(2,1), \tilde{w}(1,2), \tilde{w}(2,2), \tilde{w}(3,0), \ldots$ and define

$$\mathcal{H}(\tilde{w}) := \begin{bmatrix} \tilde{w}(0,0) & \tilde{w}(1,0) & \tilde{w}(0,1) & \tilde{w}(1,1) & \tilde{w}(2,0) & \tilde{w}(2,1) & \ldots \\ \tilde{w}(1,0) & \tilde{w}(2,0) & \tilde{w}(1,1) & \tilde{w}(2,1) & \tilde{w}(3,0) & \tilde{w}(3,1) & \ldots \\ \tilde{w}(0,1) & \tilde{w}(1,1) & \tilde{w}(0,2) & \tilde{w}(1,2) & \tilde{w}(2,1) & \tilde{w}(2,2) & \ldots \\ \tilde{w}(1,1) & \tilde{w}(2,1) & \tilde{w}(1,2) & \tilde{w}(2,2) & \tilde{w}(3,0) & \tilde{w}(3,2) & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots \end{bmatrix}. \tag{8.1}$$

The matrix $\mathcal{H}(\tilde{w})$ consists of the backwards shifts and downwards shifts of $\tilde{w}$. Therefore, to address (1) and (2) above one needs to find if possible, conditions under which col span$(\mathcal{H}(\tilde{w})) = \mathcal{P}$ and conditions we can recover the laws of $\mathcal{P}$ using $\mathcal{H}(\tilde{w})$. That is, find a set of annihilators, $\mathfrak{N}_{\mathcal{P}}$.

Some of the challenges to this problem includes the fact that even though $\mathfrak{N}_{\mathcal{P}}$ would be a finitely generated module, it is not free, see p.18 of [56]. Furthermore, one needs to formally define some of the concepts of 1D systems like the lag of $\mathfrak{B}$ and lag structure for 2D systems.

Approximate controller. In all the data-driven algorithms presented we test for implementability, i.e. $\mathcal{D} \subseteq \mathcal{P}$, and in the case when $\mathcal{D} \nsubseteq \mathcal{P}$ there is no solution to a given problem. This leads us to question is it possible to find an "approximate" solutions? That is, assume that $\mathcal{D} \nsubseteq \mathcal{P}$, is it possible to find $\mathcal{D}' \subseteq \mathcal{P}$ such that we can find a controller that "approximately implements" $\mathcal{D}$. The term approximately implement needs to be formally defined. We also need to prove under which conditions controller exists approximately implementing $\mathcal{D}$. Using angles between behaviours one needs to show that if $\mathcal{P}, \mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and $\Theta(\mathcal{P}, \mathcal{D}) \neq \frac{\pi}{2}$ then a controller possibly exists approximately implementing $\mathcal{D}$.

Assume that $\mathcal{P}, \mathcal{D} \in \mathscr{L}^{\mathtt{w}}_{contr}$ and that $\mathcal{D} \nsubseteq \mathcal{P}$. Let $w_d \in \mathcal{D}$, then a possible way of finding and approximate solution would be to solve the following minimization problem

$$\underset{w'_d \in \mathcal{P}}{\text{minimize}} \quad ||w_d - w'_d||^2_{l_2^{\mathtt{w}}} \tag{8.2}$$

A solution to the minimization problem in (8.2) has been presented in [42] and [58]. Now let $\mathcal{D}'$ be such that $w'_d \in \mathcal{D}'$ and $\mathcal{D}' \subseteq \mathcal{P}$, then we need to develop and prove conditions on w_d and w such that w'_d computed from observed trajectories is sufficiently informative about $\mathcal{D}'$.

Applications. As shown with the example in Chapter 5 section 5.3.3 a possible application of Algorithm 6 is power factor correction. Moreover, as already mentioned in the Introduction (Chapter 1), data-driven control methods are gaining popularity in application, therefore, one of the long-term research objectives is to find a suitable application for the algorithms presented.

Other developments. Throughout this thesis the data-driven control solutions developed are based on the assumption that the desired behaviour is controllable, but in practical application it is not always the case that the desired behaviour is controllable. For example in a stabilization problem then the desired is autonomous. Therefore, an open question for research is to develop conditions and algorithm such that given data from autonomous desired behaviour one can find a controller.

Appendix A

Background material

A.1 Polynomial matrices

Definition A.1. Let $l, L \in \mathbb{Z}$. A *Laurent polynomial* in the indeterminate ξ over the field of rational numbers is defined by

$$R(\xi^{-1}, \xi) := R_l \xi^l + R_{l-1}\xi^{l-1} + \cdots + R_{L+1} + R_L \xi^L.$$

Definition A.2. A polynomial matrix $U \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ (or Laurent polynomial matrix $W \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi^{-1}, \xi]$) is called *unimodular* if there exists $V \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ ($X \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi^{-1}, \xi]$) such that $VU = I_{\mathtt{g}}$ ($XW = I_{\mathtt{g}}$).

Unimodular matrices are characterized as follows.

Lemma A.3. *A polynomial matrix* (*or Laurent polynomial matrix*) $U \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ ($W \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi^{-1}, \xi]$) *is unimodular if and only if* $U^{-1} \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ ($W^{-1} \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi^{-1}, \xi]$).

Proof. See Lemma 6.3-1 p. 375 of [27] and pp. 575-576 of [79]. □

It follows from the above Lemma that if U is unimodular then $\det(U)$ is a nonzero constant, and for Laurent polynomial $\det(W) = \alpha\xi^{\mathtt{d}}$ with $\alpha \neq 0$ and $\mathtt{d} \in \mathbb{Z}$.

Without loss of generality, pre-multiplying a Laurent polynomial matrix by a unimodular matrix one gets a polynomial matrix. Henceforth, in this thesis we work with polynomial matrices and we define the rank of a polynomial matrix as follows.

Definition A.4. Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$.

(i) The *row* (*column*) *rank* of R over the field of rational functions is the maximal number of independent rows (columns).

(ii) R is of full row (column) rank if its rank is equals to the number of rows (columns).

We shall denote by rank(R) the rank of R. rank(R) is different from rank($R(\lambda)$) which is the rank of the constant matrix $R(\lambda)$ for some specific $\lambda \in \mathbb{C}$. If R is not full row rank, then it can be shown that there exists a unimodular matrix $U \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ such that $R' := UR$ has last $\mathtt{g}-\mathtt{q}$ rows of zeros, where $\mathtt{q}$ is the number of rows such that R is full row rank; this is shown in Theorem 2.5.23 p. 58 of [50].

Definition A.5. Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$. R is called *left prime* if there exists a matrix $G \in \mathbb{R}^{\mathtt{w}\times\bullet}[\xi]$ such that $RG = I_{\mathtt{g}}$.

G is called the *right inverse* of R. Left prime matrices are characterized as follows.

Proposition A.6. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ with rank$(R) = \mathtt{g}$. Then R is left prime if and only if rank$(R(\lambda)) = \mathtt{g}$ for all $\lambda \in \mathbb{C}$.*

Proof. Follows from the proof of Proposition B.8 pp. 176-177 of [52]. □

Definition A.7. Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$. R is called *right prime* if there exists a matrix $G \in \mathbb{R}^{\bullet\times\mathtt{g}}[\xi]$ such that $GR = I_{\bullet}$.

In this case G is called *left inverse* of R, moreover, the following result holds true.

Proposition A.8. *Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ with rank$(R) = \mathtt{w}$. Then R is right prime if and only if rank$(R(\lambda)) = \mathtt{w}$ for all $\lambda \in \mathbb{C}$.*

Proof. See Proposition B.8 pp. 176-177 of [52]. □

Definition A.9. Let $r = r_0 + r_1\xi + \cdots + r_L\xi^L \in \mathbb{R}^{1\times\mathtt{w}}[\xi]$. The *coefficient vector* of r is defined by

$$\tilde{r} := \begin{bmatrix} r_0 & r_1 & \dots & r_L & 0 & \dots & \dots \end{bmatrix}.$$

$\tilde{r}$ has infinite number of entries, which are all zero everywhere except for a finite number of entries. We continue to define the coefficient matrix and show how to recover polynomials matrix from coefficient matrix.

Definition A.10. Let $R = R_0 + R_1\xi + \cdots + R_L\xi^L \in \mathbb{R}^{\mathtt{p}\times\mathtt{q}}[\xi]$. The *coefficient matrix* of R is defined by

$$\tilde{R} := \begin{bmatrix} R_0 & R_1 & \dots & R_L & 0_{\mathtt{p}\times\mathtt{q}} & \dots & \dots \end{bmatrix}.$$

$\tilde{R}$ has infinite number of columns, which are all zero everywhere except for a finite number of elements. To recover R from $\tilde{R}$,

$$R := \begin{bmatrix} R_0 & R_1 & \dots & R_L & 0_{\mathtt{p}\times\mathtt{q}} & \dots & \dots \end{bmatrix} \begin{bmatrix} I_{\mathtt{q}} \\ I_{\mathtt{q}}\xi \\ \vdots \\ I_{\mathtt{q}}\xi^L \\ 0_{\mathtt{q}} \\ \vdots \end{bmatrix}.$$

We define the right shift of $\tilde{R}$ by

$$\sigma_R \tilde{R} := \begin{bmatrix} 0_{\mathtt{p}\times\mathtt{q}} & R_0 & R_1 & \dots & R_L & 0_{\mathtt{p}\times\mathtt{q}} & \dots \end{bmatrix}.$$

Let $\mathtt{k} \in \mathbb{Z}_+$ then $\sigma_R^{\mathtt{k}}\tilde{R}$ denotes $\mathtt{k}$ right shifts of $\tilde{R}$.

Definition A.11. Let $r \in \mathbb{R}^{1\times\mathtt{w}}[\xi]$. Then the *degree* of r is the highest degree of all the entries of r.

We denote by $\deg(r)$ the degree r. The degree of $R \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ is defined as follows.

Definition A.12. Let R be defined by $R := R_0 + R_1\xi + \dots + R_L\xi^L$ with $R_L \neq 0$, then L is the *degree* of R.

We denote the degree of R by $\deg(R)$. Let $R = \mathrm{col}(r_1, r_2, \dots, r_{\mathtt{g}})$ with $r_i \in \mathbb{R}^{1\times\mathtt{w}}[\xi]$ for $i = 1, \dots, \mathtt{g}$. Now let $\deg(r_i) = d_i$ and $r_i := r^i_{d_i}\xi^{d_i} + r^i_{d_i-1}\xi^{d_i-1} + \dots + r^i_0$ and define $\tilde{L} := \mathrm{col}(r^1_{d_1}, \dots, r^{\mathtt{g}}_{d_{\mathtt{g}}})$. $\tilde{L}$ is called the *leading row coefficient matrix* of R. Similarly in the case R is written in terms of its column, we can define *leading column coefficient matrix* of R.

Definition A.13. Let $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ be full row (column) rank. Then R is called *row* (*column*) *proper* if the leading row (column) coefficient matrix of R is full row (column) rank.

The following result holds about row (column) properness.

Theorem A.14. *Any nonsingular polynomial matrix $R \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ is row* (*column*) *equivalent to a row* (*column*) *proper matrix. That is, there exists a unimodular matrix $U \in \mathbb{R}^{\mathtt{g}\times\mathtt{g}}[\xi]$ ($U \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\xi]$) such that UR (RU) is row* (*column*) *proper.*

Proof. See Theorem 2.5.7 p.27 of [91]. □

The notions of proper and strictly proper rational matrices are defined as follows.

Definition A.15. Let $H \in \mathbb{R}^{\mathtt{g}\times\mathtt{m}}(\xi)$. H is *proper* if

$$\lim_{\xi\to\infty} H(\xi) < \infty.$$

Moreover, H is *strictly proper* if

$$\lim_{\xi\to\infty} H(\xi) = 0.$$

Properness and strictly proper rational are characterized as follows.

Lemma A.16. *Let $N \in \mathbb{R}^{\mathtt{g}\times\mathtt{m}}[\xi]$ and $D \in \mathbb{R}^{\mathtt{m}\times\mathtt{m}}[\xi]$. Define*

$$H(\xi) := ND^{-1}.$$

If H is strictly proper (proper) then every column of N has degree strictly less than (less than or equal to) that of the corresponding column of D.

Proof. See Lemma 6.3-10, p. 383 of [27]. □

It turns out that the converse of Lemma A.16 is not always true, see an example on p. 383 of [27]. Hence the need for conditions on D for H to be proper, such conditions are stated in the following result.

Lemma A.17. *Let H, D and N as in Lemma A.16. If D is column proper then H is strictly proper (proper) if and only if each column of N has degree less than (less than or equal) the degree of the corresponding columns of D.*

Proof. See p. 385 Lemma 6.3-11 of [27]. □

A.2 Modules and Syzygies

Definition A.18. A *module* over a ring $\mathcal{R}$ or $\mathcal{R}$-module $\mathcal{M}$, is an abelian group $(\mathcal{M}, +)$ such that for all $\alpha_1, \alpha_2 \in \mathcal{M}$ and $r \in \mathcal{R}$ then $\alpha_1 + \alpha_2 \in \mathcal{M}$ and $r\alpha_i \in \mathcal{M}$ for $i = 1, 2$.

Module over a ring is a generalization of the notion of vector space, wherein scalars belong to a ring instead of being restricted to belong to a field, see pp. 413-414 of [1].

Definition A.19. Let $\mathcal{R}$-module $\mathcal{M}$ and $\mathtt{g} \in \mathbb{Z}_+$. An ordered set, $\{\alpha_1, \alpha_2, \ldots, \alpha_{\mathtt{g}}\}$, of the elements of $\mathcal{M}$ is said to *generate* $\mathcal{M}$ or to *span* $\mathcal{M}$ if every element $\alpha \in \mathcal{M}$ is a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_{\mathtt{g}}$, i.e.

$$\alpha = r_1\alpha_1 + r_2\alpha_2 + \cdots + r_{\mathtt{g}}\alpha_{\mathtt{g}},$$

for $r_1, r_2, \ldots, r_{\mathtt{g}} \in \mathcal{R}$.

The elements $\alpha_1, \alpha_2, \ldots, \alpha_{\mathtt{g}} \in \mathcal{M}$ in Definition A.19 are called *generators*.

Definition A.20. An $\mathcal{R}$-module $\mathcal{M}$ is *finitely generated* if there exists a finite number of generators.

Definition A.21. Let $\mathcal{R}$-module $\mathcal{M}$ and $\mathtt{g} \in \mathbb{Z}_+$. $\mathcal{M}$ is said to be *free* if there exists a set of linearly independent generators, $\alpha_1, \alpha_2, \ldots, \alpha_{\mathtt{g}} \in \mathcal{M}$. That is,

$$r_1\alpha_1 + r_2\alpha_2 + \cdots + r_{\mathtt{g}}\alpha_{\mathtt{g}} = 0 \Rightarrow r_1 = r_2 = \cdots = r_{\mathtt{g}} = 0.$$

A set of linearly independent generators of $\mathcal{M}$ is called *bases generators*. The *dimension* of $\mathcal{M}$ is defined as the number of basis generators and is denoted by $\dim(\mathcal{M})$.

Definition A.22. A *submodule* $\mathcal{N}$ of an $\mathcal{R}$-module $\mathcal{M}$ is a non-empty set $\mathcal{N} \subset \mathcal{M}$ such that $n_1, n_2 \in \mathcal{N}$ and $r \in \mathcal{R}$ then $n_1 + n_2 \in \mathcal{N}$ and $rn_i \in \mathcal{N}$ for $i = 1, 2$.

$\mathcal{N}$ is called an $\mathcal{R}$-submodule of $\mathcal{M}$. Now let $\mathcal{R} = \mathbb{R}[\xi]$ and $\mathtt{w} \in \mathbb{Z}_+$. Then all $\mathbb{R}[\xi]$-submodules of $\mathbb{R}^{1\times\mathtt{w}}[\xi]$ are finitely generated and free, see p. 84 of [85].

Definition A.23. Let $\mathcal{M}$ be a $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times\mathtt{w}}[\xi]$. The *syzygy* of $\mathcal{M}$ is defined by

$$\mathcal{M}^{\perp} := \left\{ n \in \mathbb{R}^{1\times\mathtt{w}}[\xi] \mid n(\xi)m(\xi)^{\top} = 0 \text{ for all } m \in \mathcal{M} \right\}.$$

$\mathcal{M}^{\perp}$ is also an $\mathbb{R}[\xi]$-submodule of $\mathbb{R}^{1\times\mathtt{w}}[\xi]$, moreover, $\mathcal{M}^{\perp} \cap \mathcal{M} = \{0\}$ and $\dim(\mathcal{M}^{\perp}) + \dim(\mathcal{M}) = \mathtt{w}$, see p. 84 of [85].

The term syzygy is also associated with $G \in \mathbb{R}^{\mathtt{g}\times\mathtt{w}}[\xi]$ as follows

$$\mathcal{A}_1 = \left\{ n_1 \in \mathbb{R}^{1\times\mathtt{g}}[\xi] \mid n_1(\xi)G(\xi) = 0 \right\},$$
$$\mathcal{A}_2 = \left\{ n_2 \in \mathbb{R}^{1\times\mathtt{w}}[\xi] \mid G(\xi)n_2(\xi)^{\top} = 0 \right\}.$$

Both $\mathcal{A}_1$ and $\mathcal{A}_2$ are $\mathbb{R}[\xi]$-modules, and are called *left syzygy* and *right syzygy* of G, respectively.

A.3 Principal angles

Let $\mathcal{G}$ and $\mathcal{T}$ be two subspaces in $\mathbb{R}^{\mathtt{m}}$, and assume

$$p = \dim(\mathcal{G}) \geq \dim(\mathcal{T}) = q \geq 1.$$

for $p, q \in \mathbb{Z}_+$. Let $u_1, \ldots, u_q$ be a basis of $\mathcal{T}$ and $v_1, \ldots, v_q$ complemented with $p - q$ vectors be a basis of $\mathcal{G}$.

Definition A.24. The *principal angles* $\theta_1, \theta_2, \ldots, \theta_q \in [0, \pi/2]$ are defined recursively by

$$\cos(\theta_k) = \max_{u \in \mathcal{G}} \max_{v \in \mathcal{T}} u^\top v = u_k^\top v_k$$

subject to

$$\|u\| = \|v\| = 1, \qquad u^\top u_i = 0, \qquad v^\top v_i = 0,$$

for $k = 1, \ldots, q$ and $i = 1, \ldots, k-1$.

The vectors $u_1, \ldots, u_q$ and $v_1, \ldots, v_q$ are called *principal vectors*, see. p. 579 of [5] and p. 2009 of [32]. The angles can be computed using singular value decomposition (SVD), see Algorithm 3.1 p. 2015 of [32]. If the largest principal between $\mathcal{G}$ and $\mathcal{T}$ is zero then $\mathcal{T} \subseteq \mathcal{G}$ see Theorem 12.4.2 p. 604 of [18].

References

[1] M. Artin. *Algebra*. Pearson Prentice Hall, 2011.

[2] M. N. Belur. *Control in a Behavioral Context*. Doctoral dissertation, University of Groningen, 2003.

[3] M. N. Belur and H. L. Trentelman. Stabilization, pole placement, and regular implementability. *IEEE Transactions on Automatic Control*, 47(5):735–744, 2002.

[4] M. N. Belur, H. L. Trentelman, and J. C. Willems. Regular implementability and its application to stabilization of system behaviors. In *Proceedings of the 40th IEEE Conference on Decision and Control.*, volume 4, pages 3479–3484, 2001.

[5] Å. Björck and G. H. Golub. Numerical Methods for Computing Angles Between Linear Subspaces. *Mathematics of Computation*, 27(123):579–594, 1973.

[6] T. Boukhobza and G. Millérioux. Graphical characterization of the set of all flat outputs for structured linear discrete-time systems. In *6th Symposium on System Structure and Control, SSSC*, 2016.

[7] J.-T. H. Chan. Data-based synthesis of a multivariable linear-quadratic regulator. *Automatica*, 32(3):403–407, 1996.

[8] N. L. C. Chui and J. M. Maciejowski. Subspace identification- a Markov parameter approach. *International Journal of Control*, 78(17):1412–1436, 2005.

[9] T. Cotroneo. *Algorithms in Behavioral Systems Theory*. Doctoral dissertation, University of Groningen, 2003.

[10] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum. *Feedback control theory*. Courier Corporation, 2013.

[11] G. E. Dullerud and F. Paganini. *A course in robust control theory: a convex approach*. Springer Science & Business Media, 2013.

[12] A. Fabianska and A. Quadrat. Flat multidimensional linear systems with constant coefficients are equivalent to controllable 1-d linear systems. In *Proceedings of International Symposium on Mathematical Theory of Networks and Systems, MTNS*, pages 24–28, 2006.

[13] S. Fiaz, K. Takaba, and H. L. Trentelman. Tracking and regulation in the behavioral framework. *Automatica*, 47(11):2343–2354, 2011.

[14] M. Fliess, J. Lèvine, P. Martin, and P. Rouchon. Flatness and defect of non-linear systems: introductory theory and examples. *International Journal of Control*, 61(6):1327–1361, 1995.

[15] M. Fliess, J. Lèvine, P. Martin, and P. Rouchon. A Lie-Bäcklund approach to equivalence and flatness of nonlinear systems. *IEEE Transactions on Automatic Control*, 44(5):922–937, 1999.

[16] M. Fliess and R. Marquez. Towards a module theoretic approach to discrete time linear predictive control. In *14th International Symposium on Mathematical Theory of Networks and Systems, MTNS*, pages 2–7, 2004.

[17] V. Ghanbari, P. Wu, and P. J. Antsaklis. Large-scale dissipative and passive control systems and the role of star and cyclic symmetries. *IEEE Transactions on Automatic Control*, 61(11):3676 – 3680, 2016.

[18] G. H. Golub and C. F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[19] G-M. Greuel and G. Pfister. *A Singular introduction to commutative algebra*. Springer Science & Business Media, 2012.

[20] C. Heij and W. Scherrer. Consistency of system identification by global total least squares. *Automatica*, 35(6):993–1008, 1999.

[21] D. J. Hill and P. J. Moylan. Dissipative dynamical systems: Basic input-output and state properties. *Journal of the Franklin Institute*, 309(5):327 – 357, 1980.

[22] Z.-S. Hou and Z. Wang. From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences*, 235:3–35, 2013.

[23] M. Ikeda, Y. Fujisaki, and N. Hayashi. A model-less algorithm for tracking control based on input-output data. *Nonlinear Analysis: Theory, Methods & Applications*, 47(3):1953–1960, 2001.

[24] T. Jain, J. J. Yamé, and D. Sauter. Trajectory-based real-time control of an electrical circuit against unknown faults. *Journal of the Franklin Institute*, 351(2):986–1000, 2014.

[25] D. Jeltsema. Time-varying phasors and their application to power analysis. In *Mathematical Control Theory I*, pages 51–72. Springer, 2015.

[26] A. A. Julius, J. C. Willems, M. N. Belur, and H. L. Trentelman. The canonical controllers and regular interconnection. *Systems & Control Letters*, 54(8):787–797, 2005.

[27] T. Kailath. *Linear systems.* Prentice-Hall, 1980.

[28] A. Kaldmäe and Ü. Kotta. On flatness of discrete-time nonlinear systems. *IFAC Proceedings Volumes*, 46(23):588–593, 2013.

[29] C. Kandler, S. X. Ding, T. Koenings, N. Weinhold, and M. Schultalbers. A differential flatness based model predictive control approach. In *2012 IEEE International Conference on Control Applications*, pages 1411–1416, 2012.

[30] O. Kaneko and T. Fujii. Discrete-time average positivity and spectral factorization in a behavioral framework. *Systems & Control Letters*, 39(1):31 – 44, 2000.

[31] O. Kaneko, K. Takaba, and P. Rapisarda. Totally dissipative systems. *Systems and Control Letters*, 54(7):705 –711, 2005.

[32] A. V. Knyazev and M. E. Argentati. Principal angles between subspaces in an a-based scalar product: Algorithms and perturbation estimates. *SIAM Journal on Scientific Computing*, 23(6):2008–2040, 2002.

[33] M. Kuijper. Why do stabilizing controllers stabilize? *Automatica*, 31(4):621–625, 1995.

[34] J. Lévine. *Analysis and control of nonlinear systems: A flatness-based approach.* Springer Science & Business Media, 2009.

[35] J. Lévine. On necessary and sufficient conditions for differential flatness. *Applicable Algebra in Engineering, Communication and Computing*, 22(1):47–90, 2011.

[36] J. Lèvine and D.V. Nguyen. Flat output characterization for linear systems using polynomial matrices. *Systems & Control Letters*, 48(1):69–75, 2003.

[37] L. Ljung. *System Identification: Theory for the User.* Prentice Hall; 2 edition, 1999.

[38] I. Markovsky. A missing data approach to data-driven filtering and control. *IEEE Transaction on Automatic Control*, to appear, 2017.

[39] I. Markovsky and P. Rapisarda. On the linear quadratic data-driven control. In *Proceedings of European Control Conference*, pages 5313–5318, 2007.

[40] I. Markovsky and P. Rapisarda. Data-driven simulation and control. *International Journal of Control*, 81(12):1946–1959, 2008.

[41] I. Markovsky, J. C. Willems, P. Rapisarda, and B. L. M. De Moor. Algorithms for deterministic balanced subspace identification. *Automatica*, 41(5):755–766, 2005.

[42] I. Markovsky, J. C. Willems, S. Van Huffel, B. De Moor, and R. Pintelon. Application of structured total least squares for system identification. In *Proc. of the 43rd Conf. on Decision and Control*, pages 3382–3387, 2004.

[43] I. Markovsky, J. C. Willems, S. Van Huffel, and B. L. M. De Moor. *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach.* Number 11 in Monographs on Mathematical Modeling and Computation. SIAM, 2006.

[44] T. M. Maupong, J. C. Mayo-Maldonado, and P. Rapisarda. On linear flat systems. in preparation, 2017.

[45] T.M. Maupong, J. C. Mayo-Maldonado, and P. Rapisarda. On Lyapunov functions and data-driven dissipativity. In *IFAC Proceedings volumes*, volume Accepted, 2017.

[46] T.M. Maupong and P. Rapisarda. Data-driven control: the full interconnection case. In *Proceedings of the 22nd International Symposium on Mathematical Theory of Networks and Systems*, pages 395–398, 2016.

[47] T.M. Maupong and P. Rapisarda. Data-driven control: A behavioral approach. *Systems & Control Letters*, 101:37–43, 2017.

[48] J. Mayo-Maldonado. *Switched linear differential systems.* Doctoral dissertation, Unviersity of Southampton, 2015.

[49] J. Oldenburg and W. Marquardt. Flatness and higher order differential model representations in dynamic optimization. *Computers and Chemical Engineering*, 26(3):385 – 400, 2002.

[50] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach (Texts in Applied Mathematics).* Springer, 1997.

[51] C. Praagman, H. L. Trentelman, and R. Zavala Yoe. On the parametrization of all regularly implementing and stabilizing controllers. *SIAM Journal on Control and Optimization*, 45(6):2035–2053, 2007.

[52] S. Rao. *Energy methods for lossless systems using quadratic differential forms.* Doctoral dissertation, Unviersity of Southampton, 2008.

[53] P. Rapisarda. *Linear differential systems.* Doctoral dissertation, University of Groningen, 1998.

[54] P. Rapisarda and I. Markovsky. Why "state" feedback? *IFAC Proceedings volumes*, 41(2):12285–12290, 2008.

[55] P. Rapisarda and H. L. Trentelman. Identification and data-driven model reduction of state-space representations of lossless and dissipative systems from noise-free data. *Automatica*, 47(8):1721 – 1728, 2011.

[56] P. Rocha. *Structure and representation of 2-D systems.* Doctoral dissertation, University of Groningen, 1990.

[57] P. Rocha and J. Wood. Trajectory control and interconnection of 1d and nd systems. *SIAM Journal on Control and Optimization*, 40(1):107–134, 2001.

[58] B. Roorda and C. Heij. Global total least squares modeling of multivariable time series. *IEEE Transactions on Automatic Control*, 40(1):50–63, 1995.

[59] B. Roorda and S. Weiland. Optimal angle reduction−a behavioral approach to linear system approximation. *Linear Algebra and its Applications*, 337(1):189 – 235, 2001.

[60] M. G. Safonov. Focusing on the knowable: Controller invalidation and learning. *Control Using Logic-Based Switching*, pages 224–233, 1997.

[61] M. G. Safonov and F. B. Cabral. Fitting controllers to data. *Systems & Control Letters*, 43(4):299 – 308, 2001.

[62] M. G. Safonov and T.-C. Tsao. The unfalsified control concept and learning. *IEEE Transactions on Automatic Control*, 42(6):843–847, 1997.

[63] C. Scherer, P. Gahinet, and M. Chilali. Multiobjective output-feedback control via lmi optimization. *IEEE Transactions on Automatic Control*, 42(7):896–911, 1997.

[64] Y. Shen, L. Hao, and S. X. Ding. Real-time implementation of fault-tolerant control systems with performance optimization. *IEEE Transactions on Industrial Electronics*, 61(5):2402–2411, 2014.

[65] G. Shi and R. E. Skelton. Markov Data-Based LQG Control. *Journal of Dynamic Systems, Measurement, and Control*, 122(3):551–559, 2000.

[66] H. Sira-Ramirez and S. K. Agrawal. *Differentially flat systems.* CRC Press, 2004.

[67] T. Söderström and P. Stoica. *System identification.* Prentice-Hall, Inc., 1988.

[68] H. L. Trentelman. On behavioral equivalence of rational representations. In *Perspectives in Mathematical System Theory, Control, and Signal Processing*, pages 239–249. Springer, 2010.

[69] H. L. Trentelman and J. C. Willems. H_∞ control in a behavioral context: the full information case. *IEEE Transactions on Automatic Control*, 44(3):521–536, 1999.

[70] H. L. Trentelman and J. C. Willems. Synthesis of dissipative systems using quadratic differential forms: Part II. *IEEE Transactions on Automatic Control*, 47(1):70–86, 2002.

[71] H. L. Trentelman, R. Z. Yoe, and C. Praagman. Polynomial embedding algorithms for controllers in a behavioral framework. *IEEE Transactions on Automatic Control*, 52(11):2182–2188, 2007.

[72] H.L. Trentelman. On flat systems behaviors and observable image representations. *Systems & Control Letters*, 51(1):51–55, 2004.

[73] H.L. Trentelman and D. Napp Avelli. On the regular implementability of nd systems. *Systems & Control Letters*, 56(4):265 – 271, 2007.

[74] M. E. Valcher and J. C. Willems. Observer synthesis in the behavioral approach. *IEEE Transactions on Automatic Control*, 44(12):2297–2307, 1999.

[75] A. J. Van der Schaft. Achievable behavior of general systems. *Systems & Control Letters*, 49(2):141–149, 2003.

[76] P. van Overschee and B. L. M. De Moor. *Subspace Identification for Linear Systems: Theory - Implementation - Applications*. Springer, 2011.

[77] M. Vidyasagar. Control System Synthesis: a factorization approach. *Synthesis lectures on control and mechatronics*, 2(1):1–367, 1985.

[78] J. C. Willems. Dissipative dynamical systems part I: General theory. *Archive for Rational Mechanics and Analysis*, 45(5):321–351, 1972.

[79] J. C. Willems. From time series to linear system-Part I. Finite dimensional linear time invariant systems. *Automatica*, 22(5):561–580, 1986.

[80] J. C. Willems. From time series to linear system-Part II. Exact modelling. *Automatica*, 22(6):675–694, 1986.

[81] J. C. Willems. Models for dynamics. *Dynamics reported*, 2:171–269, 1989.

[82] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Transactions on Automatic Control*, 36(3):259–294, 1991.

[83] J. C. Willems. LQ-control: a behavioral approach. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, volume 4, pages 3664–3668, 1993.

[84] J. C. Willems. On interconnections, control, and feedback. *IEEE Transactions on Automatic Control*, 42(3):326–339, 1997.

[85] J. C. Willems. The behavioral approach to open and interconnected systems. *IEEE Control Systems*, 27(6):46–99, 2007.

[86] J. C. Willems. Dissipative dynamical systems. *European Journal of Control*, 13(2):134–151, 2007.

[87] J. C. Willems. Representations of linear systems. In *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on*, pages 168–171. IEEE, 2008.

[88] J. C. Willems, A. A. Julius, M. N. Belur, and H. L. Trentelman. The canonical controller and its regularity. In *Proceedings of the 42th IEEE Conference on Decision and Control.*, volume 4, pages 1639–1644. IEEE, 2003.

[89] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. M. De Moor. A note on persistency of excitation. *Systems & Control Letters*, 54(4):325–329, 2005.

[90] J. C. Willems and H. L. Trentelman. Synthesis of dissipative systems using quadratic differential forms: Part I. *IEEE Transactions on Automatic Control*, 47(1):53–69, 2002.

[91] W. A. Wolovich. *Linear Multivariable Systems*, volume 11. Springer Science & Business Media, 1974.

[92] T. Wonghong and S. Engell. Automatic controller tuning via unfalsified control. *Journal of Process Control*, 22(10):2008 – 2025, 2012.

[93] B. R. Woodley. *Model free subspace based H_∞ control.* PhD thesis, Stanford University, 2001.

[94] B. R. Woodley, J. P. How, and R. L. Kosut. Direct unfalsified controller design-solution via convex optimization. In *Proceedings of the 1999 American Control Conference*, volume 5, pages 3302–3306. IEEE, 1999.

[95] B. R. Woodley, J. P. How, and R. L. Kosut. Model free subspace based H_∞ control. In *Proceedings of the 2001 American Control Conference*, volume 4, pages 2712–2717, 2001.

[96] D. Youla, H. Jabr, and J. Bongiorno. Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case. *IEEE Transactions on Automatic Control*, 21(3):319–338, 1976.

[97] M. Zandi, A. Payman, J. P. Martin, S. Pierfederici, B. Davat, and F. Meibody-Tabar. Energy management of a fuel cell/supercapacitor/battery power source for electric vehicular applications. *IEEE Transactions on Vehicular Technology*, 60(2):433–443, 2011.