# GLOBAL TOTAL LEAST SQUARES

# Global Total Least Squares

## A method for the construction of open approximate models from vector time series

(Globale Totale Kleinste Kwadraten
Een methode voor het construeren van open benaderende modellen op grond van vectortijdreeksen)

Proefschrift

Ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr. P.W.C. Akkermans M.A.
en volgens besluit van het college voor promoties.

De openbare verdediging zal plaats vinden op
vrijdag 20 Januari 1995 om 13.30 uur
door

Berend Roorda
geboren te Assen

Promotiecommissie

| | |
|---|---|
| Promotores: | Prof. dr. T. Kloek |
| | Prof. dr. ir. J.C. Willems |
| Co-promotor: | Dr. C. Heij |
| Overige leden: | Prof. dr. H. Bart |
| | Prof. dr. H.K. van Dijk |
| | Dr. B. Hanzon |

Aan Hannie

The book is no. 88 of the Tinbergen Institute Research Series. This series is established through cooperation between Thesis Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

# Preface

PhD research is a largely open proces in which a stimulating environment plays a crucial role. I am indebted to several people and institutions that shaped such an environment for me during the years since April 1990.
First and for all, I would like to thank my supervisor, Christiaan Heij. He set me on the right track with the PhD project 'Dynamical aspects of time series models; a system theoretic approach': it is indeed what this thesis is about. His continuous support, his careful reading and correcting of my writings and his clear thoughts on the subject were of invaluable interest. The sometimes long discussions formed an important ingredient of the pleasant side of doing research.
I thank my promotors, Teun Kloek and Jan Willems, for their accurate reading of the manuscript and the numerous suggestions for improvement, each from their own viewpoint. Jan's courses on systems theory 'from the behavioural point of view' once made me choose for a specialization in that direction as a graduate student, and I have never regretted it. I also thank the members of the scientific comittee, Bernard Hanzon, Harm Bart and Herman van Dijk, for their interesting comments.
The courses at the Dutch Systems and Control Theory Network, the stimulating conferences of the European SCIENCE project on System Identification, and their financial support have been of great value for my research. Thanks also go to the NWO and the Royal Dutch Shell for their financial contribution for visiting conferences.
My colleagues at the Econometric Institute and the Tinbergen Institute have contributed substantially through discussions and a friendly atmosphere. I thank them for their positive influence on the daily work.
Special acknowledgements are given to my brother Dirk and to my colleague André Lucas for solving all kinds of LaTeXproblems. I thank Anne-Marie Haas for her linguistic hints on parts of the manuscript.
Finally, I would like to thank my family and friends for all those things that are so hard to put into words. They kept me going. This holds in particular for Hannie. I dedicate this book to her, with love.

Berend Roorda
Rotterdam, November 1994

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Dynamics deals with matter in motion or, more generally, with the evolution of phenomena over time. Virtually everything can be studied from this perspective, albeit that for instance the dynamical aspect of a phenomenon at rest is not quite fascinating. The same is true for the opposite of standstill, in which the past has no implications for the future. The interesting case is in between, if there is both change and memory, if history contains information about a changing future.

Dynamical models describe relations between past and future. They form the central objects in systems theory, incorporating system identification and control as main branches, and in time series analysis. System identification and time series analysis are partly overlapping fields of research, both concerning the construction of dynamical models on the basis of observed behaviour. Dynamic modelling is also the subject of this monograph.

The literal meaning of identification is to determine the identity of a phenomenon, to reconstruct what it is. In system identification this applies to the situation in which the qualitative aspects of a phenomenon are known, and the numerical value of some physical parameters is estimated from measurements. The qualitative aspects determine a model class, in which each model corresponds to specific values of the parameters. By analysing the quantitative aspects of observed behaviour of the phenomenon, the true values are estimated. A typical example is the estimation of the value of electrical elements in a circuit whose structure is known. As the relation between parameters and data is often non-linear, even for linear dynamic models, this is in general a highly non-trivial task, which gives parameter estimation and algorithmic aspects a well-deserved prominent place in identification.

In many situations however, the objective of identification is better described as formulating approximate statements on the behaviour of a phenomenon, rather than as aiming at an exact description of the internal structure. This involves a shift in three dimensions.

1. *exact* $\rightarrow$ *approximate*

Even if the ultimate goal is exact modelling, this is often far beyond the ability of the modeller, and then it is preferable to formulate more modest aims that

are realistic. In particular if there is lack of knowledge on the physical structure of the phenomenon, or if its structure is too complicated to be exploited by the modeller, it is in general not possible to discover the truth. For such relatively abstract objects as economic phenomena, the involved processes are in general extremely complex, and existing background theory is often not beyond dispute. Moreover, the limited amount of available data may induce severe limitations on the achievable model complexity. In such situations one cannot expect much more from models than to describe approximate features of the phenomenon of interest.
Approximation is not only due to the limitations of the modeller, an exact model is in many cases not even desirable as it would be too complex to be useful. This is most apparent in model reduction, which concerns reduction of the complexity of a model of which the internal structure might be precisely known but is too involved to be tractable. This amounts to a conscious exchange of exactness for simplicity. For real phenomena a similar type of trade-off is often essential.

2. *internal* $\rightarrow$ *external*

Approximate models often do not reflect the internal structure of phenomena. If the internal mechanism is unknown, complex or abstract, this is often not feasible, not desirable or not well-defined. Consequently, one has to concentrate on the external behaviour of a phenomenon, so this shift is more or less inherent in approximate modelling. For example, most macro-economic models consist of approximate statements on economic quantities, without reflecting the actual structure of the underlying economic reality.

3. *closed* $\rightarrow$ *open*

In general it is impossible or undesirable to isolate a phenomenon completely from its environment. By definition, the environment falls outside the scope of the model, but may affect its behaviour. Then it is reasonable to work with open models, which take into account external effects that are in principle left unexplained.
To summarize, we view a model as a description of phenomena that is

1. *approximate*, as it leaves details unexplained;
2. *external*, as it leaves the internal structure unexplained;
3. *open*, as it leaves the effect of the environment unexplained.

In this monograph we introduce a novel approach to dynamic modelling that takes full account of this view on models, both in methodology and in terminology. We consider the proposed method, called Global Total Least Squares (GTLS), as a straightforward and practical elaboration of these principles in the context of linear, time-invariant systems. It is a deterministic approach for modelling on the basis of multivariable time series by approximate dynamic

relationships. This method is described in detail in the next chapters, and it is compared with existing methods in identification and time series analysis.
We take account of the approximate character of models by allowing for a gap between model behaviour and observed time series, also called the *misfit* of a model. This is in contrast with the idea that the mere existence of such a gap is a reason to reject a model. Then one would have to increase the model complexity until this gap has become invisible, or absorb it in the model by incorporating stochastic disturbances. In approximate modelling we measure the misfit of models, and consider it as one of the factors that determine model quality.
We take account of the external character by *defining* models in terms of their external behaviour. For the GTLS method this means that we define models as the solution set of difference equations, while we view the equations themselves merely as a tool to describe such sets. In fact we will make use of different representations of such sets, called *isometric state representations*, which form the corner stone of the algorithmic aspects of the GTLS approach.
We take account of the open character of models by allowing for *indeterminism*, i.e. some arbitrariness in the behaviour that is left completely unspecified by the model. So the effect of the environment of a phenomenon is reflected by a certain degree of freedom, corresponding to that part of the model behaviour that is not governed by the phenomenon itself but determined by external effects. In the GTLS method the environment is represented by 'unknown inputs', variables whose value is left completely unspecified. This amounts to partial models with less equations than the number of components in the time series.
Summarizing, our view on models as approximate, open descriptions of the external behaviour is reflected in our methodology as follows:

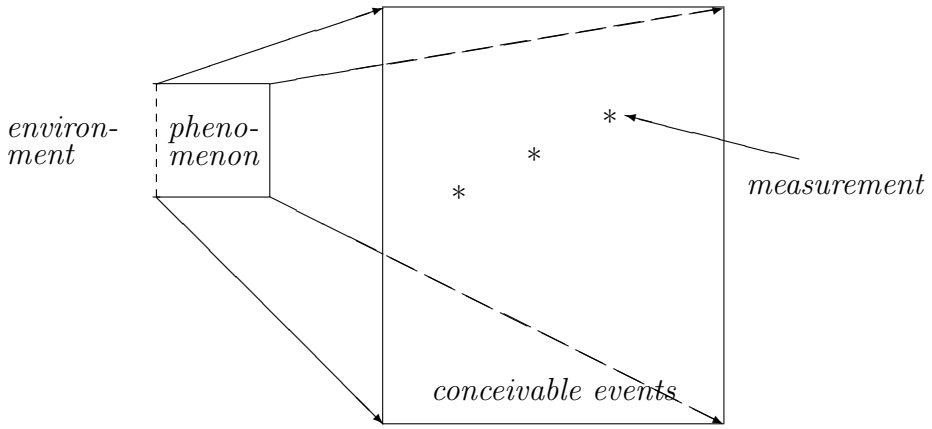
1. *approximate* ⇒ allow for a *misfit* of models.
2. *external* ⇒ define models in terms of their *external behaviour*.
3. *open* ⇒ allow for *indeterminism*.

In our opinion, in time series analysis and identification there is a certain tension between theory and practice concerning these points, even in the rich theory on linear, time-invariant (stationary) modelling. Despite the fact that in many applications these aspects are obviously relevant, and probably recognized by the modeller, they are not fully reflected in the terminology and methods that are used, both in system identification and econometrics. In the sequel of this chapter we motivate this further by discussing these aspects at a general and intuitive level. We first introduce some general notions concerning our starting point of modelling, and then indicate the distinctive contrasts with the mainstream approach in time series analysis and system identification.

## 1.1 Starting Point for Modelling

The starting point for our discussion on modelling is visualized in Fig. 1.1.

**Figure 1.1**: Starting point for modelling.




The concepts occurring in this diagram have the following meaning. A *phenomenon* is a part of the reality that is more or less conceived as an entity. One may think of physical phenomena, or more abstract ones as economic quantities like inflation or interest on loans.

The *environment* of a phenomenon is defined as that part of reality that may be of influence on the behaviour of the phenomenon, but falls *outside* the scope of the model. By definition this excludes detailed statements on the environment in models, as this would amount to incorporating it in the model. But it does not exclude to account for the *effect* of the environment on a phenomenon, for instance by considering it as a source of indeterminism.

*Measurements* are the result of our perception of a phenomenon. They represent the observed behaviour of a phenomenon, and for simplicity we consider them as known facts.

The set of *conceivable events* denotes what theoretically might be thinkable. It corresponds to the imaginable outcomes of measuring, while the measurements point out some events that actually have taken place. A measurement has several conceivable outcomes, otherwise it is a completely trivial one that need not to be carried out. So it is inherent in measuring to *imagine* unobserved events, to think of behaviour that does not take place. As the notion of conceivable events obviously precedes the measurements, we can conclude that it must be inherent in our perception of a phenomenon.

For clarity we emphasize at this point that by conceivable events we mean events that we can in principle imagine to happen, without necessarily considering them as realistic, as serious candidates for real events at some time or under certain circumstances. For example, we think of a planet taking some place at some time, and only then we can ask ourselves what kind of orbit it actually follows. The set of conceivable events contains in this case for instance a square orbit, while only in a later stage of modelling the phenomenon of planetary motion, after taking measurements, this may be ruled out as an unrealistic one that never is to occur anywhere in the universe.

## 1.2 External Modelling

We have sketched the following situation in which modelling takes place. We want to model a phenomenon, a part of reality. Somehow we have become aware of the phenomenon, and have an idea about its type of behaviour, the type of events that take place. Further we did observe some of those conceivable events that actually have taken place, represented by the available measurements. This has been depicted in Fig. 1.1.
Roughly speaking, the task of the modeller is to explain the measurements. The ultimate answer would be a description of the internal structure of the phenomenon from which it is completely clear why it behaved like we observed, and how it will behave under different circumstances. As we have argued in the introduction, this is often not achievable, and one cannot expect much more than an approximate description of the external behaviour of the phenomenon. So from this perspective, models represent theory on conceivable events. An obvious question is, what kind of statements a model makes about such events. In some situations it may be the aim of the modeller just to give a clear representation of the observed facts, without making any statement on unobserved behaviour. Think for instance of the design of a town plan. However, often a model does not only concern the measurements, but also the general features of a phenomenon, containing information about what could come, or has been, or would be under different circumstances or manipulations. Then models are not only a *simplified* representation of reality, they are also intended to tell us more than we actually observed, to *extend* our insight in reality. In our opinion this is a strong driving force behind modelling, to gain knowledge on the unobserved, the idea that from what we observed we can learn something more general about the phenomenon. Having registered what actually happened, the modeller tries to tell something about what will happen, or could have happened under different circumstances. Therefore we propose the following viewpoint.

*A model indicates to what extent* conceivable *behaviour is* realistic.

Stated otherwise, a model indicates to what extent imaginable events are in correspondence with how a phenomenon *really might* behave. Viewing models as representing theory on the external behaviour of a phenomenon, it is natural also to concentrate on the external behaviour of models, to judge them by their implications at the level of concrete events. This is indeed the central idea of our view on models. We judge models by their logical content, their concrete implications for reality. This means that we do not take into account how such implications may be described by *underlying* mechanisms or reasons. In terms of figure Fig. 1.1, such notions refer to what is *behind* the plane of conceivable events, to the interior of the phenomenon, while in the 'behavioural' approach we will not look behind the plane of conceivable events. We will concentrate on *what* a phenomenon might do, not on *how* or *why*.
Consider for example the modelling of the phenomenon 'earthy gravity' as "the gravitational force is 10 $m/s^2$". This is not a direct statement about events.

Its logical content contains statements as "objects fall with an acceleration of 10 $m/s^2$ to the earth". We consider 'force' as an auxiliary term, an abstract notion used to formulate statements on events.

Two types of models play an important role in systems theory and time series analysis, namely deterministic and stochastic models. Deterministic models are defined in terms of equations and inequalities, in terms of restrictions that may be fulfilled or not, while stochastic models make less strict distinctions by using the concept of probability. From the external viewpoint, i.e., abstracting from model mechanisms, deterministic models correspond to a bipartition of events into those which are in accordance with the model (the model behaviour), and the rest. Similarly, from the external perspective stochastic models amount to a probability distribution for the phenomenon's behaviour.
According to the foregoing, we will *define* deterministic models in terms of their behaviour. This is the central idea in the so-called behavioural approach to systems theory, as introduced by Willems in a series of papers [44, 45, 46]. There this idea is worked out consequently, incorporating theory on the formulation of model quality from this perspective.
The behavioural approach is in contrast with the common use in identification and time series analysis in which models are defined as processes, input/output mechanisms, transfer functions, or as sets of equations. Model quality is then usually defined in terms of model parameters, corresponding to coefficients in these descriptions. We view such definitions as representations of models, as a way of describing their external behaviour, as indispensable tools in estimation procedures, but not as models themselves. Consequently, the quality of models should be expressed without reference to numerical representations, directly in terms of their external behaviour.
This is not only a matter of terminology, it is also of practical importance. It helps to make a clear distinction between parametrization problems, involving the relation between models and their numerical representation, and real identification problems, concerning the relation between models and observations. If identification is viewed as a parameter estimation problem, the uniqueness of the parametrization is essential, but often hard to achieve, especially in multivariable analysis. By concentrating on the behaviour in model estimation, this non-uniqueness of parameters is less important as the parameters are deprived of any intrinsic meaning. In fact, it turns out that multivariable models do not cause substantial additional difficulties in our estimation procedures.

## 1.3 Approximate Modelling

We have argued that in describing the external behaviour of phenomena it is natural to define a model as a set of statements on conceivable events. The model statements can have various interpretations. In their strict interpretation they are supposed to be in exact correspondence with reality. Then the behaviour of a deterministic model is claimed to consist precisely of all possible events, and the distribution of a stochastic model is supposed to describe the

true likelihood of the occurrence of events. More precisely, this should reflect the probability of observations as a sample from a population. We will refer to this as *exact* modelling. This is relevant for instance in a simulation context, where the data is generated by known models.
In practice models most often have the less strict interpretation of being approximative descriptions instead of making statements on the behaviour of a phenomenon that are exactly in correspondence with reality. Approximate stochastic models indicate the likelihood of events to some accuracy level, or express the degree of belief of the modeller which are by definition not claimed to reflect a real stochastic mechanism. Approximate deterministic models describe which events play a dominant role, are relatively large or frequent components in the behaviour of the phenomenon. Another form of approximate modelling is worst-case identification, where the main purpose is to describe a set of imaginary events that surely contains all realistic ones, rather than excluding all unrealistic behaviour.
At this point we like to motivate our choice in this monograph concerning

$$\textit{deterministic} \leftrightarrow \textit{stochastic}$$

modelling. One may state that the stochastic approach is by far the most popular one in identification and especially in time series analysis. For an overview on stochastic time series analysis we refer to [7, 17, 24]. Nonetheless, we will use the deterministic approach, for the following three reasons.
Firstly, the use of stochastics is not compelling in approximate modelling, even in the presence of random, irregular properties of the phenomenon of interest. Namely, as we allow for a gap between models and reality, we may leave some details unexplained and concentrate on features of more regular nature. Further, a part of the irregularities may be assigned to the influence of the environment. As the environment is left unexplained, this induces some indeterminism in the model. For clarity we remark that according to the standard terminology a model can be both deterministic and indeterministic, as it may contain degrees of freedom without using a stochastic framework. So our first argument is that stochastics is not the only way to handle irregularities, disturbances and randomness, as open approximate deterministic models are also fairly flexible.
Secondly, the use of a stochastic framework is in many cases not very convincing. The gap between reality and our relatively simple models has not always the stylized nature of a random variable. From a methodological point of view it is then preferable to *exclude* uncomprehended aspects from modelling, instead of *incorporating* them as a stochastic disturbance.
This brings us to a more pragmatic motive for using the deterministic approach: approximate stochastic modelling is much more complex. The question to what extent a given observation deviates from a certain stochastic distribution is much more involved than the question to what extent it deviates from a certain set of events, because there is a less clear borderline between stochastic disturbances *in* the model, and the gap *between* the phenomenon and the proposed stochastic properties. The main problem is not

so much that this borderline is soft, i.e., that the disturbances in the model mix with deviations from the model, but that this borderline has a much more complex nature than for deterministic models. For example, if we state that a phenomenon behaves like a Gaussian white noise process, this implicitly proposes a certain distribution for the standard deviation of the fourth moment of the second order sample autocorrelation, and an observation that is in good correspondence with most other implications of the model might point out a serious deviation from this aspect. This also illustrates that the likelihood of observations is by no means sufficient to formalize to what extent a stochastic model is approximate, as such a model has many facets that can be tested in innumerably many ways.

## 1.4 Open Modelling

As stated before, it is often impossible or undesirable to isolate phenomena from their environment. This raises the question in how far the environment should be taken into account in modelling. One possibility is to incorporate the environment in the model, which means in fact that one changes his mind and start to model an extended phenomenon, namely the original one plus its environment. This might work if the scope of the model has been chosen inappropriately, but more often than not it is not feasible to model the environment in a satisfactory way. Moreover, the same problem returns, now concerning the environment of the environment and so on. In order to avoid ambiguity in terminology we repeat our definition of the environment of a phenomenon as that part of reality that may be of influence on the behaviour of the phenomenon, but falls *outside* the scope of the model.
This means that the environment causes indeterminism, as some aspects of the external behaviour of the phenomenon are not determined by its intrinsic properties but by external unmodelled effects. As we argued before, it is then natural to allow for indeterminism in the model as well, by incorporating some degrees of freedom in the model behaviour. A straightforward way to formalize this is to represent the effect of the environment as unspecified inputs, which are not part of the data (as exogenous variables), but have to be constructed during the identification process. The value of these inputs is left completely free by the model, without a preference for small values over large ones. This corresponds to partial models, in which some aspects of the phenomenon's behaviour are assigned to unmodelled external effects.
This deviates from the common deterministic approaches in which the environment often consists of known inputs, i.e., variables whose value is known to the modeller and forms part of the measurements in the modelling process. This refers to situations in which the modeller has a precise idea about which variables influence the phenomenon of interest, and moreover the known inputs often represent control variables that can be manipulated. In fact this means that the inputs fall within the scope of the model, so the phenomenon may be considered as an isolated input/output system.

We remark that our approach also results in models that can be given the interpretation of input/output systems, but in a different way. The inputs represent the effect of the environment and hence they are not part of the data. As such they do not correspond to the usual notion of inputs, and in fact they are better called *auxiliary inputs*, which reflects the fact that it is just an auxiliary term, a way of talking about indeterminism in models. The comparison between our approach and input/output modelling is in fact even more subtle, but this will be discussed later on. We will also show how to modify the GTLS approach for modelling on the basis of input/output data without allowing for approximation errors for the input variables.
In stochastic modelling the environment is often represented by certain exogenous variables or by random variables with a specific distribution, most often white noise. The first corresponds to 'known inputs', which is comparable to the input/output setting as described above. The latter is a more loose way of modelling the environment, but nevertheless much stronger than considering completely unspecified inputs. Modelling the environment as a specific random variable in general induces a preference for small effects, and in many cases models are even estimated on the basis of minimizing the size of the shocks from the environment. However, incorporating pure indeterminism causes problems in the stochastic approach, as free variables with completely unspecified value do not fall within the definition of a random variable.
Summarizing, we account for the effect of the environment by indeterminism in the model behaviour, which is in contrast with most approaches in which one presumes certain stochastic properties of the environment and/or represent it by known inputs.

## 1.5 Overview

This monograph may be divided in three parts. In the first part we describe *what* the GTLS method is, and *why* we consider it as a useful addition to time series analysis. In the second part we describe *how* to construct GTLS models from a given time series. The last part consist of applications and extensions. An overview is given in the following scheme.
We have chosen for a somewhat abstract description of the GTLS method in which we concentrate on the essential features, without any reference to model representation or parametrization issues. The static case serves as a simple illustration of the method at the same abstract level. In order to make things more concrete, we give alternative formulations in terms of difference equations and input/output systems. The method is put in contrast with other least squares methods, which further explains the term *global* and *total* in the name GTLS.
The part on the solution method is of a somewhat technical nature. The basic ingredients are state space representations, Riccati equations, the singular value decomposition and iterative Gauss-Newton algorithms. Although we have tried to make this part self-contained, some familiarity with these issues may

**Table 1.1**: Overview of monograph.

| | |
|---|---|
| **What & Why?** | Chapter 2 |
| What is GTLS? | Section 2.1 |
| Why GTLS? | Section 2.2 |
| Further explanation of GTLS | |
| - the static case | Section 2.3 |
| - description by difference equations | Section 2.4.2 |
| - description by input/output systems | Section 2.4.3,4 |
| - comparison with other least squares methods | Section 2.5 |
| | |
| **How?** | Chapter 3 – 5 |
| Tool: the Isometric State Representation | Chapter 3 |
| Evaluation of GTLS criterion | Chapter 4 |
| Construction of GTLS models | Chapter 5 |
| | |
| **Applications and Extensions** | Chapter 6 |
| in econometrics | Section 6.1 – 6.3 |
| in systems theory | Section 6.4 – 6.5 |

be helpful. The method is illustrated by a simple leading example. Appendix B contains an implementation of the algorithms in Matlab.
The applications in the last chapter concern time series on economic quantities (interest rates in the United States and seasonal consumption and income in the German Federal Republic), a simulation experiment with multiple outputs (also called a simultaneous equation model), results on optimal model reduction, and the construction of models on the basis of input/output data.
A brief explanation of some basic notions and the explanation of the used notation are listed in the back.

The main part of this monograph has been published in [33, 36, 34, 32]. The first paper concerns a heuristic method for the construction of models from data that is used in Section 5.6 to determine initial models for the iterative algorithms. In [34] the GTLS problem is analysed for infinite square summable time series, while in [32] we concentrate on finite data. In [36] we give a description of the GTLS problem in terms of difference equations. We remark that [12] contains an analysis of the GTLS problem for single-input single output systems in terms of Hankel matrices.
The applications on economic data and the various extensions of the GTLS method described in Chapter 6 have not been published elsewhere.

# Chapter 2

# Global Total Least Squares

Global Total Least Squares (GTLS) is a method for modelling a dynamical phenomenon on the basis of time series reflecting several quantitative aspects of the phenomenon. The goal is to determine the main characteristics of the joint evolution of these aspects. We take a deterministic approach and restrict ourselves to linear, time-invariant models. So we aim at an approximate description of the behaviour of the phenomenon in terms of linear relations that do not vary over time.

One of the main themes of this monograph is to give an interpretation of the complete modelling process at the level of external behaviour, without reference to internal model mechanisms. This is based on the consideration that models often do not reflect the internal structure of the phenomenon, but only concern their external behaviour. Then it is natural to concentrate on the external behaviour of models as well. In this chapter we describe the GTLS method according to this view point. Models are defined in terms of their behaviour, namely as sets of time series, and model criteria are formulated in geometrical terms, namely the dimension of models and their distance to data. This results in a purely set-theoretic description of the GTLS method, to determine models with minimal distance to the data under size constraints.

This description of GTLS is simple in the sense that it only involves basic mathematical notions as set, dimension and distance. Moreover, in this way we restrict ourselves entirely to the relation between models and data, without having to pay attention to technical topics as the relation between models and parameters in numerical representations, as we need not make any reference to representations at all.

On the other hand we feel that such a description might be considered as somewhat unusual or perhaps as too abstract. Therefore, as an illustration we discuss the special case of static modelling which will evoke a clear geometrical picture of what GTLS is about. Further we give alternative, more concrete formulations in terms of difference equations and transfer functions. The method is compared with others in order to illustrate some distinctive aspects. We remark that the next chapters, concerning the construction of GTLS models, are quite concrete in terms of numerical representations of systems.

We conclude this chapter by giving a motivation of the specific choices under-

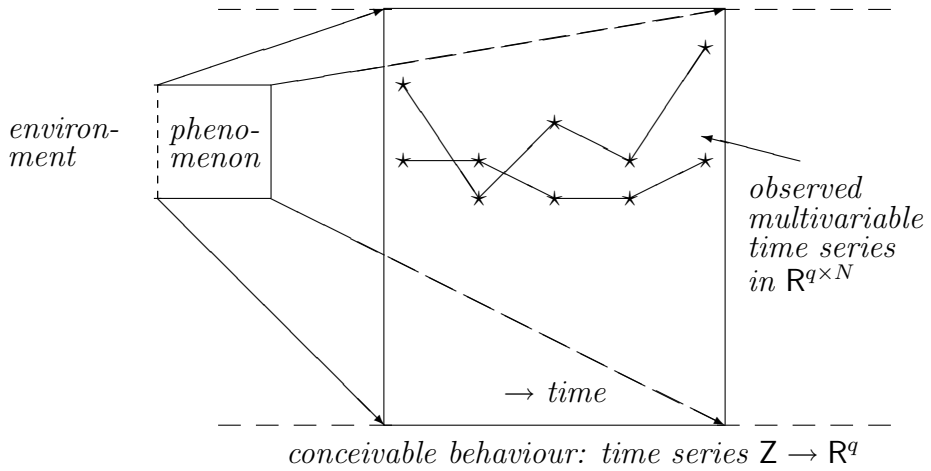
lying the method, and indicate some limitations.
We remark that we treat several extensions of the standard GTLS problem in Chapter 6. This concerns e.g. the scaling of variables, and periodic models.

## 2.1 Behavioural Description

### 2.1.1 Starting Point

The starting point for GTLS modelling is visualized in Fig. 2.1.

**Figure 2.1**: Starting point for GTLS.




The aim is to determine the realistic content of conceivable behaviour (represented by the set of time series over $\mathsf{Z}$) from knowledge on what actually has taken place (represented by an observed time series). In the picture the data consists of two components ($q = 2$) over five time instants ($N = 5$).

Compared to the general starting point as depicted in Fig. 1.1, we now assume that the phenomenon is of such a type that its behaviour can be represented by a multivariable time series. This means that it should consist of several quantitative aspects whose value can be measured at regularly spaced time instants. Alternatively, it may concern several separate phenomena. This is merely a matter of terminology, and we will view the object of modelling as one phenomenon with several aspects. According to the terminology used in the introduction, we say that the conceivable behaviour of the phenomenon consists of time series with a certain number of components.
Further we assume that measurements are available consisting of a multivariable time series, representing behaviour that actually has taken place. The time interval corresponding to the measurements is called the observation interval, which we assume to be finite, and we suppose that there are no missing data points.
Conceivable events are defined as time series over the infinite time axis $\mathsf{Z}$, the set of all integers $\{\ldots, -1, 0, 1, \ldots\}$, irrespective of the actual observation interval, for the following reasons. Models often do not only concern the observation interval, but represent more general statements concerning also the

future evolution of phenomena, or their unobserved past. Then it is appropriate to define conceivable behaviour according to such a purpose. As this may depend on the application, we choose a time axis that contains all finite intervals. Afterwards one can restrict the attention to those model statements concerning relevant time instants, and ignore the rest. Further, we will consider time-invariant models, which almost automatically apply to the whole time axis $\mathsf{Z}$. Another reason for taking the infinite time axis $\mathsf{Z}$ is that it streamlines the representation of models in our approach.
The starting point as sketched above can be summarized in mathematical notation as follows. Let $q$ denote the number of aspects, or equivalently, the number of components of the time series, and let $N$ be the length of the observation interval, defined as the number of time instants. For simplicity we start numbering the time instants at $t = 1$, so we consider observation intervals $\{1, \ldots, N\}$, or $[1, N]$ for short.

> *measurements:*
> time series with $q$ components $w : T \to \mathsf{R}^q$ on the observation interval $T = [1, N] \subset \mathsf{Z}$ or, equivalently, matrices $w \in \mathsf{R}^{q \times N}$.
> *conceivable events:*
> the set of all infinite time series with $q$ components $\{w' : \mathsf{Z} \to \mathsf{R}^q\}$, or equivalently, the set $(\mathsf{R}^q)^{\mathsf{Z}}$.

**Example.** We consider measurements consisting of the (multivariable) time series

$$\{w(1), \ldots, w(5)\} = \begin{bmatrix} 3 & 0 & 2 & 1 & 4 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix},$$

cf. Fig. 2.1. So $q = 2$, $N = 5$, and the observation interval $T$ equals $\{1, 2, 3, 4, 5\}$. The time series $w$ is identified with a $(2 \times 5)$-matrix in which the $i$-th column represents a 'fact' (concerning two aspects of a phenomenon, or concerning two phenomena) at time $i$. The set of conceivable events consists of all time series on $\mathsf{Z}$ with two components, representing arbitrary (imaginary) behaviour of the two aspects at every time, including non-observed past and future. $\diamond$

We consider the measurements as observed facts about the phenomenon. The conceivable events represent which behaviour is theoretically imaginable, they indicate the type of behaviour. According to our discussion in the introduction, we view modelling as making statements about the realistic content of conceivable behaviour. So on the basis of our knowledge of what actually has taken place, we try to determine to what extent time series over $\mathsf{Z}$ are in correspondence with how the phenomenon really might behave, e.g. in the future, in the past, or under different circumstances.

### 2.1.2 Model Class

In the GTLS method we use deterministic models, representing restrictions on time series that may be satisfied or not. We define models in terms of their

external behaviour, the set of time series that are in accordance with the model restrictions.

**Definition 2.1.1 (Models)** *Models are subsets of time series, that is, with behaviour* $\mathcal{B} \subset (\mathsf{R}^q)^{\mathsf{Z}}$.

The symbol $\mathcal{B}$ stands for behaviour, and in our approach we identify models with their behaviour.
We will not consider all types of models, but make a rigorous choice for some structural properties. The first condition is that models should reflect time-invariant relationships, i.e., properties of the phenomenon that do not depend on time. This is expressed at the level of behaviours by the condition of *shift-invariance*. Let $\sigma$ denote the shift operator, defined by $(\sigma f)(t) = f(t+1)$. For sets of time series it is defined as the shift of all elements.

**Definition 2.1.2 (Shift-Invariance)** *A model* $\mathcal{B} \in (\mathsf{R}^q)^{\mathsf{Z}}$ *is called shift-invariant if* $\sigma\mathcal{B} = \mathcal{B}$.

This means that if a time series belongs to a model, all its forward and backward shifts should also be contained in it.
Secondly, we will restrict our attention to linear models, i.e. models that are linear spaces. Let $\lambda\hat{w}$ denote scalar multiplication of all components of the time series $\hat{w}$ by $\lambda \in \mathsf{R}$, and let $\lambda\mathcal{B}$ denote the set $\{\lambda\hat{w};\ \hat{w} \in \mathcal{B}\}$.

**Definition 2.1.3 (Linearity)** *A model* $\mathcal{B} \in (\mathsf{R}^q)^{\mathsf{Z}}$ *is called linear if* $\lambda\mathcal{B} + \mu\mathcal{B} = \mathcal{B}$ *for all* $\lambda, \mu \in \mathsf{R}$.

These two conditions characterize the model class in the GTLS method. So we define

**Definition 2.1.4 (Model Class)** *In GTLS the model class consists of linear, shift-invariant subspaces of time series over* $\mathsf{Z}$ *with* $q$ *components. It is denoted as* $\mathsf{B}^q$*, or if we leave* $q$ *unspecified, by* $\mathsf{B}$.

Such models reflect linear relationships that are independent of the time instant. As they are closely related to dynamical systems, we will also refer to them as *systems* if we discuss their properties rather than their relation to data.
**Example.** As an illustration, we consider the system

$$\mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^2;\ \hat{w}_2(t) = \alpha\hat{w}_2(t-1) + \beta\hat{w}_1(t)\}, \tag{2.1}$$

where $\alpha, \beta$ denote fixed real numbers. This system consists of a linear, shift-invariant set of time series with two components, so it belongs to $\mathsf{B}^2$. Clearly every (set of) difference equation(s) define(s) a linear, shift-invariant system. Conversely, under a mild extra condition all linear, shift-invariant systems admit a description in terms of linear difference equations with constant coefficients, which is discussed in Section 2.4. ◇

We remark that this model class does not form the main innovative aspect of the GTLS approach. Although its definition without reference to model parameters might be somewhat unusual, it is determined by the quite conventional structural properties of linearity and time-invariance. The distinctive feature of GTLS concerns the way in which the accuracy of these models are measured, as described in the next section.

### 2.1.3 Model Criteria

We have defined models as subsets of time series that are linear and shift-invariant. The central question is now which of them are good models for an observed time series. Stated in our terminology, which subsets of time series form a good description of the phenomenons behaviour? In this section we indicate the two criteria on which the GTLS approach is based.
Firstly, a model should be in correspondence with the data. In exact modelling this would mean that the model should contain the observation in its behaviour, or more precisely, in the restriction of its behaviour to the observation interval. In mathematical notation, exact modelling requires $w \in \mathcal{B}_T$, with $w$ denoting the observation and $\mathcal{B}_T$ the restriction of $\mathcal{B}$ to the observation interval $T$.
As motivated in the introduction, it is often not realistic to stick to this severe requirement, and it is then more appropriate to impose it only in an approximate sense. This means that a model should be accurate, i.e., that it should contain a time series that is close to the data. So the first criterion in our approach is

*accuracy:* a model should have a low misfit

The formalization of this requires a kind of distance measure between data (one time series) and a model (a set of time series). We consider the Euclidean distance between the data and the closest element in the model behaviour. In principle the Euclidean norm is only defined for vectors, but by identifying a matrix $\tilde{w} \in \mathsf{R}^{q\times N}$ with a point in $\mathsf{R}^{qN}$ we can extend its definition to matrices in the obvious way. So we define for $\tilde{w} \in \mathsf{R}^{q\times N}$,

$$\|\tilde{w}\|^2 := \Sigma_{t\in T}\Sigma_{i=1}^{q}\tilde{w}_i(t)^2 \tag{2.2}$$

where $\tilde{w}_i$ denotes the $i$-th component of $\tilde{w}$, which means that we take into account the 'sum of squares', both over the time instants and the components. This is also called the Frobenius norm of the matrix $\tilde{w} \in \mathsf{R}^{q\times N}$. For the Euclidian norm of a vector $a$ we use the notation $|a|$, so that $\|\tilde{w}\|^2 = \Sigma_{t\in T}|\tilde{w}(t)|^2$. The misfit of a model is defined as follows.

**Definition 2.1.5 (Misfit)** *The misfit of a model $\mathcal{B} \in \mathsf{B}^q$ with respect to a time series $w : T \to \mathsf{R}^q$ is defined as*

$$\mathrm{d}(w, \mathcal{B}) := \min_{\hat{w} \in \mathcal{B}_T} \|w - \hat{w}\|. \tag{2.3}$$

This will also be called the 'total least squares' criterion, where the term total refers to the fact that we allow for deviations in all components. The squared misfit is comparable to the so-called 'sum of squares residuals (RSS)' in (orthogonal) regression models, but the definition of what residuals are in GTLS is quite different from other methods, as explained in Section 2.5.

The second criterion underlying the GTLS approach is that a model should be powerful, i.e., it should exclude many conceivable events as unrealistic or hardly present in reality. This is an obvious criterion: the smaller a behaviour, the stronger the statements it represents on the behaviour of a phenomenon. In the context of deterministic behavioural modelling this means a preference for small models. So we also aim at

*power:* a model should be small.

As the models we consider are defined as linear subspaces of time series, an obvious way to formalize this is in terms of their dimension. The most straightforward way to do this would be to define the size of models simply as their dimension as a subspace of $(\mathsf{R}^q)^{\mathsf{Z}}$. However, this is infinite in general as $\mathsf{Z}$ is infinite. Therefore we consider their dimension on finite time intervals. As models are time-invariant, this only depends on the length of the interval. In our definition of the size of a model we also exploit the fact that the increase of the dimension of $\mathcal{B}_{[1,t]}$ with the length of the interval is constant (except for small $t$).

**Definition 2.1.6 (Size, Rank, Degree)** *The size of a system $\mathcal{B} \in \mathsf{B}$ is defined as the pair of integers $(m, n)$ such that $\dim(\mathcal{B}_{[1,t]}) = mt + n$ for $t \geq n$. Then $m$ is called the rank of the system, and $n$ its degree. Sizes are partially ordered by $(m', n') \leq (m, n)$ if $m' \leq m$ and $n' \leq n$.*

*Proof of correctness.* For a proof of the existence and unicity of $(m, n)$ we refer to Appendix A.2.

The rank of a system denotes its degree of freedom at each time instant. This reflects indeterminism in the model that can be considered as due to external effects that continuously affect the phenomenon. The degree of a system corresponds to the freedom due to initial conditions that are left unspecified by the model. Namely, the dimension of $\mathcal{B}_{[1,t]}$ is $n$ larger than the dimension on the same interval if the past would have been specified. We will give a further interpretation in terms of representations of models in Section 2.4.

**Example.** Again we consider the system (2.1),

$$\mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^2;\ \hat{w}_2(t) = \alpha \hat{w}_2(t-1) + \beta \hat{w}_1(t)\}.$$

In addition we suppose that $\alpha$ and $\beta$ are not both zero. Then this restriction is not effective on intervals of length 1, so $\mathcal{B}_{[1,1]}$ is the set of all data points in $R^2$, which has dimension 2. The set $\mathcal{B}_{[1,2]}$ is a three-dimensional space, as $\hat{w}_1(1)$,

$\hat{w}_2(1)$ and $\hat{w}_1(2)$ can be chosen arbitrary, and determine the value of $\hat{w}_2(2)$. Similarly, the dimension of $\dim\mathcal{B}_{[1,t]}$ equals $1+t$, from which it easily follows that $\text{rank}(\mathcal{B}) = 1$. This corresponds to the fact that at each time instant one variable can be chosen arbitrarily. In addition, one initial condition can be chosen, which corresponds to the fact that $\text{degree}(\mathcal{B}) = 1$. For $\alpha = 0$ and $\beta = 0$ the degree is zero. Later on we will see that, roughly speaking, the rank corresponds to the number of components minus the number of equations, and the degree to the sum of the lags of the equations. $\diamond$

### 2.1.4 Global Total Least Squares

Two criteria for model quality have been formulated, accuracy and power, which are formalized in the definition of respectively the misfit and size of models, see Definitions 2.1.5 and 2.1.6. We want to keep both small, but they are competing objectives and somehow we have to make a trade-off between them. As an illustration of this, we consider the following two extreme cases.

| *extremely small:* | *extremely large:* |
|---|---|
| $\text{rank}(\mathcal{B}) = 0$, $\text{degree}(\mathcal{B}) = 0$ | $\text{rank}(\mathcal{B}) = q$, (then $\text{degree}(\mathcal{B}) = 0$) |

In the first case $\mathcal{B}$ consists only of the zero time series, so it has misfit $\mathrm{d}(w, \mathcal{B}) = \|w\|$ with respect to every observation $w$. This corresponds to a trivial isolated (autonomous) system that declares all events as unexplained details. In the second case, all time series on the observation interval are compatible with the system, in particular the observation itself, so $\mathrm{d}(w, \mathcal{B}) = 0$. This corresponds to an overly open system that declares everything as due to external unmodelled effects.
This shows that the optimization of both criteria separately leads to trivial models that do not assign any aspect of the data to intrinsic properties of the phenomenon. The interesting cases are somewhere in between both extremes. We take a simple approach to the involved trade-off by considering minimizing the misfit of a model for a specified constraint on its size.

> *A GTLS model is a linear, shift-invariant space of time series of restricted dimension that has minimal total least squares distance to a given multivariable time series.*

A more formal description is as follows. Let $\mathsf{B}^{q,m,n}$ denote the systems in $\mathsf{B}^q$ with size at most $(m, n)$, i.e., linear, shift-invariant systems with rank at most $m$ and degree at most $n$.

**Definition 2.1.7 (The Global Total Least Squares Problem)**
*Given:*

- *an observation* $w : T \to \mathsf{R}^q$, *with* $T = [1, N] \subset \mathsf{Z}$
- *tolerated size* $(m, n)$,

*determine:*

- *a system* $\mathcal{B}^* \in \mathsf{B}^{q,m,n}$ *with minimal misfit* $\mathrm{d}(w, \mathcal{B})$.

$\mathcal{B}^*$ *is called a GTLS model for* $w$ *of size* $(m, n)$.

The name Global Total Least Squares is based on the used misfit criterion of Definition 2.1.5. So 'least squares' refers to summation of squares of deviations at each time, 'total' to the fact that we allow for deviations in all components. The term 'global' is used to reflect the fact that we approximate the data by a time series that belongs as a whole to a linear, shift-invariant model. This is put in contrast with other methods in Section 2.5.

The GTLS problem involves a double minimization. The inner minimization, the evaluation of the misfit $\mathrm{d}(w, \mathcal{B})$, amounts to optimization over a linear space. Secondly, we have to determine a system for which this misfit is as small as possible. This is the hard part in the GTLS approach, as it consists of a non-quadratic optimization problem over a non-convex set.
For further explanation we give an alternative formulation in which both minimizations are combined into one. Let $\mathcal{B}^{q,m,n}$ denote all time series that belong to a system in $\mathsf{B}^{q,m,n}$, i.e.,

$$\mathcal{B}^{q,m,n} := \{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ \text{there exists a } \mathcal{B} \in \mathsf{B}^{q,m,n} \text{ such that } \hat{w} \in \mathcal{B}\}. \quad (2.4)$$

Elements in this set are called regular, provided that $m < q$, as they satisfy linear, time-invariant relations. The GTLS problem can be interpreted as an optimal approximation of the data by regular time series, as follows.

**Definition 2.1.8 (GTLS as Decomposition)**
*Given:*

- *an observation* $w : T \to \mathsf{R}^q$, *with* $T = [1, N] \subset \mathsf{Z}$
- *tolerated size* $(m, n)$,

*determine:*

- *a decomposition* $w = \hat{w}^* + \tilde{w}^*$ *with* $\hat{w}^* \in \mathcal{B}_T^{q,m,n}$ *and* $\|\tilde{w}^*\|$ *minimal.*

## 2.2 Motivation

In the introduction we have argued that it is often realistic to view models as *approximate* descriptions of the *external* behaviour of a phenomenon that moreover should have an *open* character as the effect of the environment should be taken into account. This viewpoint is the main theme of the introduction, where it is discussed at a general level. The GTLS approach fits perfectly in this perspective.

1. GTLS models are *approximate*, as a misfit between models and data is allowed, and hence they leave details unexplained (corresponding to $\tilde{w}^*$ in Definition 2.1.8).

2. GTLS models are *externally* defined in terms of their behaviour, so they leave the internal structure of phenomena unexplained.

3. GTLS models are *open*, as they contain a degree of freedom at each time instant, denoted by their rank. So they leave a part of the behaviour unexplained that may be considered as due to external effects.

The GTLS method means a choice for a deterministic approach, as opposed to using stochastic models. This has been motivated in Section 1.3 and 1.4, by pointing at the fact that the formalization of our modelling principles is relatively straightforward in a deterministic setting.
In the description of the GTLS method we took our starting point in the definition of models as subsets of time series, cf. Definition 2.1.1. This can be considered as a consequence of defining deterministic models in terms of their external behaviour, in the context of time series analysis. Once this definition has been adopted, it is natural to define the other modelling concepts in the same spirit. This forms the main motivation for the way in which the GTLS approach is described.
The preceding considerations give a general motivation for the GTLS method and the used terminology. In the sequel we discuss additional choices behind the approach, concerning the model class and the used criteria for model quality.

### 2.2.1 Motivation of Model Class

The model class in GTLS is characterized by two properties, linearity and time-invariance. These are quite conventional assumptions in identification for which there has already been developed a rich theory, but nevertheless even in this classical context our modelling principles lead to a different method. The GTLS method may serve as an illustration of these principles in a well-known context. Notice however that it concerns multivariable analysis, for which the theory is less extensive. Moreover, in Chapter 6 we treat various extensions in which these assumptions are weakened, e.g. by replacing time-invariance by periodicity. We further remark that the algorithm for determining an optimal approximation, within a *given* system, of the data, applies to linear time-varying models as well.
A more intrinsic argument to impose linearity and time-invariance is that many phenomena expose a corresponding type of behaviour, albeit often only approximately. This is reflected by the fact that many physical laws express linear, time-invariant relations, and also in economic theory they play an important role, as in the simple law that in equilibrium supply equals demand. Moreover, non-linear systems can be approximated by linearizations around a certain working point, and the variation of properties of phenomena over time

may be slow, so that they are approximately constant on short intervals. In these situations several linear, time-invariant models could be estimated from parts of the data.
We hasten to admit that in many situations the reason to impose these properties is of much more pragmatic nature. Generally speaking, in identification the model class cannot be based only on knowledge about the type of phenomenon that is modelled, but one also has to take into account practical limitations. Firstly, the construction of models should be feasible. In particular this requires numerical representation of models, and for linear, time-invariant systems this is well-developed. In the estimation of GTLS models from data we will make use of theory on model representation, more specifically of state representations, albeit that we develop a new variant. Secondly, the resulting models should be transparent enough to give insight in the phenomenon's behaviour. The restriction to linear time-invariant systems guarantees that once a GTLS model has been constructed, there is an extensive theory available for analysing its properties.
We want to remark, however, that adopting linearity and time-invariance cannot always be assigned exclusively to pragmatism. Simplification is not only due to practical limitations, it can also be viewed as inherent in modelling. Only if the object of modelling is of such a simple nature that we can overlook it completely, 'true' models come into the picture, as in simulation studies with a relatively simple data generating process.
Therefore we adopt the viewpoint that in principle the model class may be chosen as one likes, and we then judge it afterwards on the basis of the results. This is perhaps an overly defensive standpoint, but in our opinion it is in accordance with a significant part of the modelling practice.
The question may arise how far the restriction to linear, time-invariant models limits the scope of the GTLS approach. This question is hard to answer in general, but we want to emphasize that the answer is not that it only applies to phenomena that exposes 'linear', 'time-invariant' or 'stationary' behaviour (here we use these terms in an informal way). Notice that GTLS models contain indeterminism at each time instant, which is reflected by their rank. In this way we can incorporate 'non-stationary' behaviour in models, which is then viewed as due to varying conditions in the environment and hence left unexplained. Further, if non-linearities have a relatively small effect on the external behaviour, in approximate modelling they may be declared to be the unexplained details that form the gap between models and data.

**Remark on linearity and time-invariance.**
We also want to point at a general methodological consideration concerning the choice of a model class, concerning the *coherence* of model restrictions. We discuss this topic in an informal way, as in fact it goes somewhat beyond the scope of this monograph.
By definition a GTLS model represents linear, time-invariant relations, each of which has concrete implications at each time instant (due to shift-invariance) and for each quantity (due to linearity). For example, consider the relation $y(t) = u(t-1)$. This induces an infinite number of concrete restrictions, among which 'if $u(1) = 3$ then

$y(2) = 3$', 'if $y(2) = 3.27$ then $u(1) = 3.27$' and 'if $u(2) = 3$ then $y(3) = 3$'. We view this as a coherent set of restrictions, emanating from one and the same relatively abstract model law denoted as '$y(t) = u(t-1)$'. In our discussion below we also make this distinction between concrete *restrictions* involving the occurrence of one event and more abstract *relations* reflecting a general property.

The coherence of concrete restrictions is a quite essential feature of models. Namely, this makes it possible that, at least in principle, *one* relation can be confirmed by the data *several* times, as it induces several concrete restrictions. Then, depending on the strength of the relation, on the accuracy of confirmation of its concrete implications by the data, and on how often they are confirmed, this gives us (some) confidence in adopting the relation as a general property of the phenomenon. This means that we also accept its implications that have not been confirmed by the data, e.g. involving the future (prediction) or different circumstances (simulation). Clearly every statement involving hypothetical, unobserved behaviour can only be motivated by the data indirectly, namely by the confirmation of coherent restrictions concerning the data.

Imposing linearity and time invariance is one way to guarantee such coherence, and of course both can be modified in many ways. One could replace linearity by more sophisticated functional forms, or time-invariance by the weaker requirement that regularities may change in a certain time-invariant way. Our main point is that in time series analysis these conditions cannot simply be dropped but have to be replaced by other structural requirements that somehow guarantee coherence of concrete restrictions implied by models, both in time and quantity.

In order to illustrate this we consider the consequences of simply releasing either time-invariance or linearity. If we drop time-invariance, we may freely impose completely different linear restrictions at each time. For instance, this does not exclude the model

$$\mathcal{B} = \{\lambda(w^- \wedge w \wedge w^+);\ \lambda \in \mathsf{R}\} \tag{2.5}$$

with $w^-$ and $w^+$ fixed arbitrary backward- and forward extensions of the data, and $\wedge$ denoting concatenation. Notice that this model is linear, exact, as it has zero misfit, and extremely small, as it consists of a one dimensional space. So with respect to the criteria of accuracy and size it is an excellent model, but clearly it makes no sense. The concrete restrictions concerning the observation interval are only trivially supported by the data, the others even not at all, and due to the complete lack of coherence in time they do not reflect a general property that in principle could be confirmed repeatedly by the data.

Similarly, if we drop the condition of linearity, we may freely impose completely different time-invariant relations for different quantities. So we may consider all functional forms that are independent of time, and without imposing additional model structure this also leads to nonsense models. For example, for data $w : T \to \mathsf{R}^q$ we may come up with the model

$$\mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ \hat{w}(t) = f(\hat{w}(t-1))\} \tag{2.6}$$

with $f$ a time-invariant function on $\mathsf{R}^q$ that satisfies $f(w(t-1)) = w(t)$ (if there are data points $w(t_1)$ and $w(t_2)$ that coincide exactly, then $f$ is not a function but a relation). This model is time-invariant, exact and extremely small, but in general the restrictions for different quantities are poorly motivated, and again we would not have much confidence in its predictions or any statement concerning hypothetical situations.

So we may summarize that in time series modelling one always has to guarantee coherence of model restrictions, both in time and quantity, and imposing linearity and time-invariance is a rigorous way to do so. ◇

### 2.2.2 Motivation of Model Criteria

The model criteria underlying the GTLS method, accuracy and power, are of quite general nature. Clearly both are desirable properties of models which do not need further motivation. Here we motivate how they are formalized in the GTLS approach, and make some remarks on whether they are sufficient to characterize model quality.

The criterion of accuracy has been formalized in the definition of misfit, cf. Definition 2.1.5, which in fact expresses the absence of accuracy. In this definition we take into account the distance between the data and the closest element in the model. This is an obvious idea, while the choice for measuring the distance by the Euclidean- or total least squares norm is mainly a matter of convenience. Least squares estimation plays a dominant role both in time series analysis and system identification, not in the last place because it has some convenient mathematical properties that play a role in the estimation of models. It also concerns stochastic modelling, as it is equivalent with maximum likelihood estimation for normally distributed data generating processes. The idea is that a deviation of 2 is considered as 4 times as worse as a unit deviation. Of course this involves some arbitrariness, and a disadvantage is that exceptional high deviations are given a huge weight, which is not always desired. On the other hand, some choice has to be made, and 'least squares' has proven its value in practice as a reasonable option. Moreover, it has convenient mathematical properties that play a role in the estimation of models.

To allow for deviations in all components is somewhat less conventional. In most approaches one makes a prior distinction into input- and output components, or into exogenous and endogenous variables and does not approximate the first group. In our approach the choice for a *total* least squares criterion is natural, as we do not make such a distinction but treat all variables on the same footing. As compared with the use of exogenous variables, we represent the environment by 'unknown' exogenous variables, as discussed in Section 1.4.

Another aspect of the definition of misfit is its global character. As already mentioned, this refers to the fact that the data is approximated by a time series that belongs as a whole to a linear, time-invariant system. This is an obvious choice in our exposition, as a local version would amount to approximating the data part by part, which seems artificial. Nevertheless, the majority of the identification methods is of this nature. Namely, models are often defined in terms of difference equations, and deviations are then defined as an *equation error*, in terms of the local implications of these equations. In particular, many methods are based on minimizing the first step ahead prediction error of the equations. We further discuss this issue in Section 2.5, where we compare GTLS with several local methods.

Concerning the definition of size, once models have been defined as linear subspaces it is an obvious choice to consider their dimension. The rank and degree as defined in Definition 2.1.6 determine this dimension completely on intervals with length at least equal to the degree. In practice the tolerated degree is often chosen much smaller than the observation interval, so that the precise dimension on shorter intervals may be considered as a detail. A more pragmatic argument not to consider the dimension on shorter intervals is that we cannot handle such a refinement in our procedures.
Summarizing, accuracy and power are convincing model criteria, and their formalization in terms of misfit and size is straightforward.

The question may arise whether these criteria are sufficient to characterize good models. First we want to emphasize that in a general context the answer is no. If we simply drop the condition of linearity or time-invariance, these criteria certainly do not exclude completely trivial models, as has been discussed in Section 2.2.1. So in the following discussion the restriction to linear, time-invariant models is essential. Further, the quality of models may depend on specific requirements dictated by the type of application, which of course we cannot take into account at this general level.
In the GTLS approach the above question amounts to whether a linear, shift-invariant model with low misfit and small size is a good model. Generally speaking, we may say that such a model indeed has its value, in the following respect. If the size of a system is small, this means that the time series belonging to it satisfy non-trivial linear relations at each time-instant exactly. The lower the rank and degree of a system, the lower its dimension, so the stronger the statements for its behaviour. Now if the misfit is also small, these model statements apply to the data as well, modulo some small unexplained details. So a GTLS model with small size and low misfit with respect to the data represents non-trivial general properties of the behaviour of the underlying phenomenon that are repetitively confirmed by the data up to some small deviations. In many situations this is considered as a good characterization of model quality.

### 2.2.3 Model Objectives

The last aspect of the GTLS approach we want to discuss is the way in which the two criteria of accuracy and power are worked into one final objective. We consider the minimization of the misfit of models with respect to data under a constraint on their size, cf. Definition 2.1.7. Solving this problem for different restrictions on the rank and degree gives an impression of the involved trade-off between the tolerated size and the resulting approximation error. The search for a reasonable size constraint is facilitated by the fact that the error decreases for increasing tolerated rank or degree. Moreover, by definition the rank of a system with $q$ components is between zero and $q$, and as the data is finite, only a finite number of degrees is relevant. So in principle all combinations can be tried out, while in practice it often suffices to consider only a few of them.

We could also have considered the converse approach, namely minimizing the size under a misfit constraint. This is a valid approach as well, and the advantage is that then the bound consists of one number instead of two. A less attractive aspect is that then this bound is a real number, so that in principle there are infinitely many relevant bounds.
Notice that by considering the minimization of the misfit for all relevant size constraints we would also obtain the complete answer to this reverse optimization problem. So from a logical point of view both approaches are in fact equivalent.
A more general approach to make a trade-off between accuracy and power would be to develop a criterion function that relates to the size and misfit of a model a certain value of model quality, based on more inherent considerations. Examples of such types of trade-offs are Akaike's and Bayes information criteria, and Rissanen's minimum description length principle. We do not address these approaches, as the need for an automatic trade-off of accuracy and power in the context of modelling on the basis of one time-series is not that strong, and instead it is often preferable to consider several possibilities, corresponding to different tolerated sizes.
Concerning the ordering on sizes as defined in Definition 2.1.6, this is an obvious choice as far as it compares models. We use a partial ordering in which we do not compare models if one model has a higher degree but lower rank than the other. The reason for this is that on short intervals one model may have a higher dimension than the other, while on long intervals its dimension is lower, so that an ordering will depend on a preference for intervals of a certain length. Alternatively, we could consider the lexicographic ordering, in which a model is called smaller than another one if either its rank is smaller or if the ranks are equal but its degree is smaller. A motivation for this would be that this is a full ordering, and that 'smaller' then coincides with 'having smaller dimension on long intervals'. A disadvantage is that this inevitably leads to autonomous models, i.e., with zero rank, as it is always preferable to choose a high degree of models to taking the rank equal to one.

The motivation of the GTLS approach is summarized in Table 2.1.

## 2.3 Static Total Least Squares

As an illustration of the GTLS method we first consider the static case, i.e., we consider the approximate modelling of an observed multivariable time series by means of linear time-invariant non-dynamic relations. In fact, this amounts to a well-known method.

> *'GTLS for tolerated degree zero'* coincides with the method called (static) *'total least squares'* or *'orthogonal regression'*.

We briefly discuss the model class and model criteria for this situation, and describe how the corresponding estimation problem is solved.

**Table 2.1**: Motivation of GTLS as straightforward formalization of modelling principles.

| | **principle** | **GTLS** |
|---|---|---|
| View on Models: | approximate | allow for misfit<br>choice for deterministic models |
| | external | models as subsets of time series |
| | open | indeterminism at each time instant |
| Model Class: | pragmatism,<br>(coherence) | linearity and shift-invariance |
| Model Criteria: | accuracy | small total least squares<br>distance to the data |
| | power | small dimension |
| Model Objective: | trade-off | minimize misfit under constraint<br>on size |

According to Definition 2.1.4, the model class consists of subspaces of time series that are linear and shift-invariant. In addition we now impose that models are static, which can be formalized at the level of their behaviour by the condition

$$\{w(t) \in \mathcal{B}_{[t,t]} \text{ for all } t \in \mathsf{Z}\} \Rightarrow \{w \in \mathcal{B}\}. \tag{2.7}$$

This means that all model restrictions can be evaluated pointwise, i.e., separately for each time instant. This is in contrast with dynamic models that relate behaviour at different time instants, so that the past behaviour has implications for the future.
So a linear, shift-invariant static model takes the form

$$\mathcal{B} = \{w : \mathsf{Z} \to \mathsf{R}^q;\ w(t) \in \mathcal{M} \text{ for all } t \in \mathsf{Z}\}, \tag{2.8}$$

with $\mathcal{M}$ a linear subspace of $\mathsf{R}^q$. It is easily verified that the misfit of such a model can be evaluated pointwise, and that the misfit at each time instant $t$ is given by the orthogonal distance between data point $w(t)$ and $\mathcal{M}$. Concerning the size of static models, notice that $\dim(\mathcal{B}_{[1,t]}) = mt$ with $m$ the dimension of $\mathcal{M}$, so its rank equals $m$ and it has zero degree.
We say that $\hat{w} \in \mathsf{R}^{q \times N}$ has rank $m$ if it belongs to a static system of rank $m$ and not to one of lower rank. This corresponds to the common definition of the rank of a matrix. Now the static total least squares problem can be formulated as follows.

**Definition 2.3.1 (Static Total Least Squares)**
*Given:*

- *an observation* $w : T \to \mathsf{R}^q$*, with* $T = [1, N] \subset \mathsf{Z}$
- *tolerated rank* $m$*,*

*determine*

- *a decomposition* $w = \hat{w}^* + \tilde{w}^*$ *with* $\mathrm{rank}(\hat{w}^*) \leq m$ *and* $\|\tilde{w}^*\|$ *minimal.*

We call $\hat{w}^*$ the static rank-$m$ approximation of $w$. Generically this approximation indeed has rank $m$, but for degenerate cases its rank can be smaller. The solution to this problem is given by the singular value decomposition (SVD), see e.g. [16]. We denote the usual Euclidean norm on vectors $a$ by $|a|$.

**Definition 2.3.2 (Singular Value Decomposition)**
*Every* $w \in \mathsf{R}^{q \times N}$ *can be decomposed as* $w = \Sigma_{i=1}^{q} \lambda_i u_i v_i^\top$, *with*

1. $\lambda_1 \geq \ldots \geq \lambda_q \geq 0$, *called the singular values of* $w$;
2. $u_i \in \mathsf{R}^{q \times 1}$ *with* $|u_i| = 1$ *and* $u_i \perp u_j = 0$ *for* $i \neq j$, *called the left singular vectors of* $w$;
3. $v_i \in \mathsf{R}^{N \times 1}$ *with* $|v_i| = 1$ *and* $v_i \perp v_j$ *for* $i \neq j$, *called the right singular vectors of* $w$.

*The singular values are uniquely determined, and if they are distinct, then the singular vectors are also uniquely determined, up to a sign.*

So this decomposes a matrix with $q$ rows as the sum of $q$ matrices of rank one (or zero) that are mutually orthogonal and ordered according to their magnitude. The static total least squares problem is solved by the SVD as follows. Define $U_m := [u_1, \ldots, u_m] \in \mathsf{R}^{q \times m}$.

**Proposition 2.3.3 (Optimal Static Approximation)**
*Let* $w = \Sigma_{i=1}^{q} \lambda_i u_i v_i^\top$ *be the SVD of* $w \in \mathsf{R}^{q \times N}$. *Then*

- $\hat{w}^* := \Sigma_{i=1}^{m} \lambda_i u_i v_i^\top$ *is the static rank-m approximation of* $w$
- $\mathcal{B}_{stls}^m := \{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ w(t) \in \mathrm{im}(U_m)\}$ *is a corresponding optimal static model*
- $\mathrm{d}(w, \mathcal{B}_{stls}^m)^2 = \Sigma_{i=m+1}^{q} \lambda_i^2$

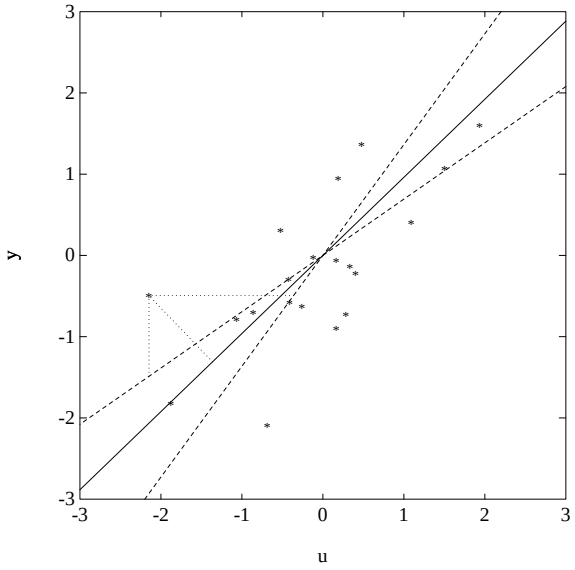
*Proof.* This result follows immediately from the properties of the SVD, cf. [16]
♣

We remark that generically $\hat{w}^*$ is unique, namely if $\lambda_m \neq \lambda_{m+1}$.

**Example.** Figure 2.2 shows the scatter diagram of measurements consisting of two components $(u, y)$ at twenty time instants, so $q = 2$ and $N = 20$. The optimal static total least squares model of rank one, $\mathcal{B}_{stls}^1$, consists of a line that minimizes the sum of the squared orthogonal (Euclidean) distances of each point to the line. The optimal rank-1 approximation $\hat{w}^* \in \mathcal{B}_{stls}^1$ of the data consists of the orthogonal projection of each point on this line.
For comparison we also depict two regression lines, denoting optimal models if no approximation is allowed in either $u$ or $y$. These lines minimize respectively the sum of squared vertical and horizontal distances to the line, and are obtained by regression of respectively $y$ on $u$ and $u$ on $y$.

◇

**Figure 2.2**: Total Least Squares compared to regression.




In the scatter diagram of measurements, the solid line denotes the Static Total Least Squares model, and the dashed lines the two regression models; the dotted lines indicate which distance is minimized in the three different models.

## 2.4 Alternative Formulations

In the previous sections we have given an abstract definition of the GTLS method at a purely set-theoretic level. For the case of static models this leads to the well-known method of orthogonal regression. In this section we give some alternative formulations of the GTLS problem in terms of concrete numerical representations of models, which may be helpful to gain further intuition for the GTLS method. However, we first have to address the notion of *completeness*, which plays a role in the representation of GTLS models.

### 2.4.1 Completeness

Before we discuss various descriptions of the GTLS problem in terms of numerical representations, we address a subtlety concerning the properties of systems over infinite lag. This may be considered as a small price we have to pay for defining models on the infinite time axis $\mathsf{Z}$.
The problem is that models that coincide on all finite time intervals may be different in their behaviour on the infinite time axis $\mathsf{Z}$. We illustrate this by an example.

**Example.** Compare the system (2.1),

$$\mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^2;\ \hat{w}_2(t) = \alpha \hat{w}_2(t-1) + \beta \hat{w}_1(t)\},$$

with the system $\mathcal{B}^{\text{fin}}$ consisting of all time series in $\mathcal{B}$ that have finite support,

i.e.,

$$\mathcal{B}^{\text{fin}} = \{\hat{w} \in \mathcal{B};\ \exists t_0, t_1 \in \mathsf{Z} \text{ such that } \hat{w}(t) = 0 \text{ for } t < t_0 \text{ and } t > t_1\}.$$

Both systems coincide on finite intervals: every finite part of a trajectory in $\mathcal{B}$ is also a finite part of a trajectory in $\mathcal{B}_{\text{fin}}$. On the other hand, $\mathcal{B}$ contains time series with infinite support, so the systems are not equal. $\diamond$

As representations should define a model unambiguously, behind them there is a choice on the limiting properties of models. Notice that such a choice is irrelevant for the quality of systems as GTLS models, as this is entirely defined in terms of their behaviour on finite time.
Roughly speaking, we will assume that there are no restrictions that concern infinite lags, which is formalized by the property of completeness.

**Definition 2.4.1 (Completeness)** *A model $\mathcal{B} \in (\mathsf{R}^q)^{\mathsf{Z}}$ is called complete if it satisfies the condition*

$$\{w_T \in \mathcal{B}_T \text{ for all finite intervals } T \subset \mathsf{Z}\} \Rightarrow\ w \in \mathcal{B} \tag{2.9}$$

So completeness means that if a time series is in correspondence with the model on finite intervals, it is admissible as a whole. Notice that the system $\mathcal{B}$ in the preceding example is complete, while $\mathcal{B}^{\text{fin}}$ is not. This condition can indeed be imposed without loss of generality.

**Lemma 2.4.2 (Complete Models)**
*For each system $\mathcal{B}$ in the GTLS model class $\mathsf{B}^q$ there exists a unique model $\bar{\mathcal{B}} \in \mathsf{B}^q$ that coincides with $\mathcal{B}$ on finite intervals. This model is called the completion of $\mathcal{B}$ and is given by $\bar{\mathcal{B}} = \{w : \mathsf{Z} \to \mathsf{R}^q;\ w_T \in \mathcal{B}_T \text{ for all finite intervals } T \subset \mathsf{Z}\}$.*

*Proof.* See Appendix A.2.

Therefore, in the GTLS approach we can restrict the attention to complete models, and define the model class accordingly.

**Definition 2.4.3 (Model Class of Complete Models)** *$\bar{\mathsf{B}}^q$ denotes the class of complete models in $\mathsf{B}^q$, i.e., the class of linear, shift-invariant, complete subspaces in $(\mathsf{R}^q)^{\mathsf{Z}}$. If we want to leave the $q$ unspecified, we denote it by $\bar{\mathsf{B}}$. Similarly, $\bar{\mathsf{B}}^{q,m,n}$ denotes the class of complete models in $\mathsf{B}^{q,m,n}$.*

In the sequel we only impose completeness where it is relevant, e.g. in the discussion of numerical representations of systems. Examples of non-complete systems that are of practical relevance are $l_2$-systems, which are discussed in [34].

### 2.4.2 Difference Equations

The GTLS problem can be formulated in terms of linear difference equations with constant coefficients and finite lag. For time series with $q$ components such equations take the form

$$r_0\hat{w}(t) + \ldots + r_d\hat{w}(t-d) = 0,\ r_i \in \mathsf{R}^{1\times q} \tag{2.10}$$

Here $d$ denotes the *lag* of the equation, provided that $r_0 \neq 0$ and $r_d \neq 0$.
A set of $p$ of such equations can be grouped in matrix notation as follows.

$$R_0\hat{w}(t) + \ldots + R_d\hat{w}(t-d) = 0,\ R_i \in \mathsf{R}^{p\times q} \tag{2.11}$$

This equation describes a set of time series consisting of its *solution set*

$$\{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ R_0\hat{w}(t) + \ldots + R_d\hat{w}(t-d) = 0,\ \text{for all } t \in \mathsf{Z}\}, \tag{2.12}$$

which equals the intersection of the solution sets of the individual equations.
Two sets of equations are called *equivalent* if their solution sets coincide. Equations are called *independent* if they cannot be replaced by a smaller equivalent set of equations, i.e., by a smaller set of equations with the same solution set.
Another relevant feature of matrix difference equations (2.11) it their *total lag*, which is defined as the sum of the lags of the individual equations.
GTLS models can be described in terms of difference equations as follows.

**Proposition 2.4.4 (Representation by Difference Equations)**
*The class of solution sets of linear, time-invariant difference equations of finite lag coincides with $\bar{\mathsf{B}}$, the class of linear, shift-invariant complete systems. Moreover, for systems in $\bar{\mathsf{B}}^q$ it holds that*

- *their rank equals $q - p$, with $p$ the number of independent equations that describe the system, and*
- *their degree equals the minimal possible total lag in such a description.*

*Proof.* See Appendix A.2.

**Example.** We consider two equations for time series with three components,

$$\begin{aligned} \hat{w}_2(t) &= \hat{w}_2(t-1) + \hat{w}_1(t) \\ \hat{w}_3(t) &= \hat{w}_2(t) + \hat{w}_1(t-1) \end{aligned} \tag{2.13}$$

In matrix notation this gives

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \hat{w}(t) + \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \hat{w}(t-1) = 0. \tag{2.14}$$

These equations are independent, as they cannot be replaced by one equivalent equation. Both equations are first order, so their total lag equals 2. From the proposition it follows that these equations define a linear, shift-invariant, complete system of rank $3 - 2 = 1$ and degree 2. ◇

This leads to following formulation of GTLS, cf. Definition 2.1.8.

**Definition 2.4.5 (GTLS in terms of Difference Equations)**
*Given:*

- *an observation* $w : T \to \mathsf{R}^q$, *with* $T = [1, N] \subset \mathsf{Z}$
- *tolerated size* $(q - p, n)$,

*determine:*

- *a decomposition* $w = \hat{w}^* + \tilde{w}^*$, *with* $\hat{w}^*$ *satisfying (at least)* $p$ *independent linear time-invariant difference equations of total lag at most* $n$ *and with* $\|\tilde{w}^*\|$ *minimal.*

We remark that an optimal approximation $\hat{w}^*$ satisfies more than the required number $p$ of independent equations only if the data itself already does, which is proved in Section 5.1.
So GTLS corresponds to the approximation of an observed time series by a regular one that satisfies a certain number of independent equations of restricted total lag. Stated equivalently, the GTLS method for data $w$ and tolerated size $(m, n)$ amounts to:

> *determine* $p := q - m$ *independent equations of total lag at most* $n$ *for which we have to change the data* $w$ *by a minimal amount* $\tilde{w}^*$ *such that* $w - \tilde{w}^*$ *satisfies the equations exactly, throughout the whole observation interval.*

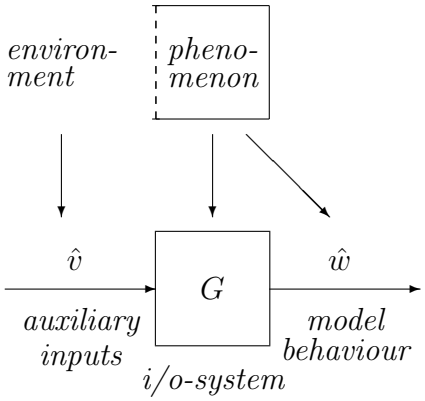
If the misfit is small, i.e., if $\tilde{w}^*$ is small, the equations can be considered as due to intrinsic properties of the phenomenon, as they reflect a structural restriction on the phenomenons external behaviour. Here $\tilde{w}^*$ corresponds to the 'unexplained details', cf. the introduction. Further, in general we impose fewer restrictions than the number of components, i.e., $p < q$, so the equations do not determine the behaviour completely, even not modulo initial conditions. This results in indeterminism at each time instant, which reflects the fact that a part of the behaviour is not so much inherent in the phenomenon but due to unmodelled external effects.
As the GTLS criterion measures the difference between the data and the behaviour that is in exact correspondence with the model laws, it is called a *behavioural* misfit. This is in contrast with *equation error* oriented criteria, which measure the deviations from the model laws at each time separately. The use of a behavioural misfit is the distinctive feature of the GTLS with respect to the deterministic approximate modelling procedures in [44, 20], and also with respect to most stochastic methods.

### 2.4.3 Auxiliary Inputs

In GTLS modelling we take account of the effect of the environment by incorporating indeterminism in the model behaviour. This is reflected by allowing for a nonzero rank of systems or equivalently by imposing less equations than

the number of components in the time series, leading to a degree of freedom at each time instant.
In linear models these external effects can be represented as *auxiliary inputs*, which is depicted in the following diagram.

**Figure 2.3**: GTLS in terms of auxiliary inputs.




In GTLS the behaviour of a phenomenon is explained approximately as the effect of free, unobserved auxiliary inputs (representing the environment) in a linear, time-invariant system.

The auxiliary inputs consist of time series over $\mathsf{Z}$, and their value determines a time series in the model behaviour, modulo initial conditions. The environment is left unmodelled, so the value of the auxiliary inputs is completely free. We use the term i/o-system if they are viewed as mappings, in order to distinguish it from systems in the GTLS model class which are defined as sets.
In order to formulate the GTLS problem according to this scheme, we have to answer the question which i/o-systems correspond to linear, shift-invariant models of a certain rank and degree. As explained before, without loss of generality we can restrict our attention to models that are complete, so this amounts to the question for which i/o-systems $G$ it holds that the set of all outputs belongs to $\bar{B}^{q,m,n}$, i.e., the class of linear, shift-invariant complete subsets of time series over $\mathsf{Z}$ with $q$ components.
This can be answered in quite conventional system theoretic notions. Namely, models of rank $m$ and degree $n$ correspond to linear, time-invariant i/o-systems with $m$ inputs and McMillan degree (or state dimension) $n$. In fact, this class of systems plays a central role in systems theory, where such systems are represented e.g. by rational transfer functions, state space descriptions or impulse response representations.
For those who are less familiar with these notions we remark that we discuss this topic in further detail in the next chapters, as the scheme in Fig. 2.3 also plays a role in our estimation procedures. The GTLS problem can be formulated in terms of auxiliary inputs as follows.

**Definition 2.4.6 (GTLS in terms of Auxiliary Inputs)**
*Given:*

- *an observation* $w : T \to \mathsf{R}^q$*, with* $T = [1, N] \subset \mathsf{Z}$

- *tolerated size* $(m, n)$,

*determine:*

- *a linear, time-invariant transfer function* $G$ *with* $m$ *inputs,* $q$ *outputs, and McMillan degree at most* $n$*, and*
- *an auxiliary input sequence* $\hat{v}^* : \mathsf{Z} \to \mathsf{R}^m$,

*such that for the corresponding output* $\hat{w}^* := G\hat{v}^*$, $\|w - \hat{w}_T^*\|$ *minimal.*

This relates the GTLS method to the output error method (see e.g. [24]) in which a system is determined for which the output to a given input approximates the data as close as possible. The difference is that in the output error method the input is supposed to be known, so it is in fact part of the data, while in the GTLS approach the input is an auxiliary notion that represents the indeterminism in the model and is hence unknown. So we can conclude that

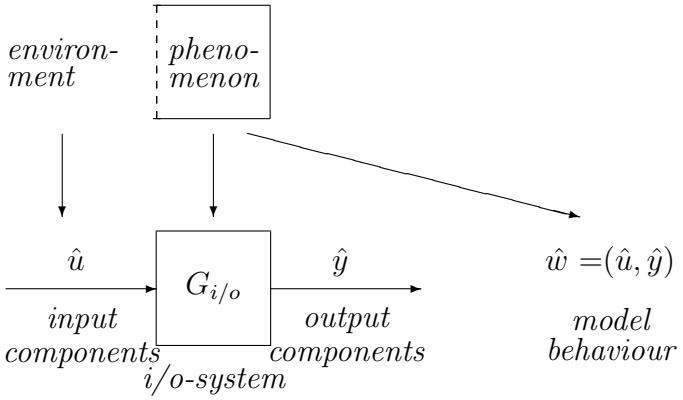
> *GTLS corresponds to the output error method, but with* unknown, artificial *inputs.*

The output error method is further discussed in Section 2.4.4 and 2.5.
We remark that the optimal $G$ and $\hat{v}^*$ are by no means uniquely determined by the GTLS criterion, even if the optimal approximation $\hat{w}_T^*$ is unique. This is discussed in detail in the next chapter. Here we only make some informal remarks. It turns out that we can always choose $G$ to be causal, i.e., such that outputs do not depend on future inputs, which means that the auxiliary inputs $\hat{v}$ only have a future effect on $\hat{w}$. Hence the values of $\hat{v}^*(t')$ for $t' > N$ are irrelevant. In fact, $\hat{w}^*$ is determined by $\hat{v}_T^*$ and an initial state, representing the degree of freedom due to initial conditions. The non-uniqueness of $G$ may be illustrated by the fact that (almost without loss of generality) we can choose $G$ to be isometric, i.e., such that $\|G\hat{v}\| = \hat{w}$. This is one of the key ideas underlying the representations of systems that we use in our estimation procedures.

### 2.4.4 Input/Output Decomposition

GTLS models have an indeterministic character in order to account for unknown external effects. In the previous interpretation of the GTLS method these effects were represented as auxiliary inputs. Another interpretation is based on assigning all indeterminism in the model behaviour to some components of the time series, and to view the remaining ones as their 'effect' or 'output'. In this way we break the symmetric treatment of all components as we divide them into two groups $\hat{w} = (\hat{u}, \hat{y})$, where $\hat{u}$ is considered to be determined by the environment, and $\hat{y}$ as the effect of the environment and of intrinsic properties of the phenomenon. This is depicted in the following scheme.

**Figure 2.4**: GTLS in terms of an input/output decomposition.




If we take our starting point in an input/output decomposition of data components, in GTLS the output behaviour of a phenomenon is explained approximately as the effect of approximate inputs in a linear time-invariant system.

Here we use the subindex in $G_{i/o}$ in order to make a distinction with the role of i/o-systems in Fig. 2.3 as a mapping from auxiliary inputs to all components in the model behaviour. The GTLS problem can be formulated in terms of input/output decompositions as follows.

**Definition 2.4.7 (GTLS in terms of I/O Decomposition)**
*Given:*

- *an observation* $w : T \rightarrow \mathsf{R}^q$*, with* $T = [1, N] \subset \mathsf{Z}$
- *tolerated size* $(m, n)$*,*

*determine:*

- *a linear, time-invariant system* $G_{i/o}$ *with* $m$ *inputs,* $q - m$ *outputs, and at most* $n$ *states,*
- *an input sequence* $\hat{u}^* : \mathsf{Z} \rightarrow \mathsf{R}^m$*, and*
- *a decomposition of the data* $w = (u, y)$*, with* $u$ *consisting of* $m$ *components,*

*such that for the corresponding output* $\hat{y}^* := G\hat{u}^*$*,* $\|u - \hat{u}_T^*\|^2 + \|y - \hat{y}_T^*\|^2$ *minimal.*

This interpretation relates the GTLS method to the output error method in yet another way, cf. Section 2.4.3.

> *GTLS corresponds to the output error method, but with also the inputs approximated.*

We conclude with some intuitive remarks. The input-output decomposition is not inherent in GTLS systems On the contrary, generically every GTLS system as defined in Definition 2.4.3 with rank $m$ admits an arbitrary decomposition into $m$ free components and $q-m$ remaining 'output' components, just as for a $(q \times m)$ matrix generically every set of $m$ rows is non-singular and hence the remaining rows are linearly dependent on these. In Section 3.4 we give a construction for making input/output decompositions for GTLS systems.
These i/o-systems $G_{i/o}$ relate to i/o-systems $G$ in terms of auxiliary inputs as introduced in Section 2.4.3 as follows. Notice that $G$ has $q$ outputs, cf. Fig. 2.3, while $G_{i/o}$ has $q-m$. An input/output decomposition corresponds to a representation in which the auxiliary inputs $\hat{v}$ equal the input components in the behaviour $\hat{u}$. For example, if we decide to consider the first $m$ components in the model behaviour as inputs, i.e. $\hat{w} = \begin{pmatrix} \hat{u} \\ \hat{y} \end{pmatrix}$, this corresponds to taking $G$ of the form $G = \begin{bmatrix} I_m \\ G_{i/o} \end{bmatrix}$. Indeed, without loss of generality we can take $G$ of this type, i.e. with $m$ rows forming an identity matrix.
So one GTLS system of rank $m$ can be described by several different i/o-systems with $m$ inputs and $q-m$ components. Further, it is not exceptional that some of these are stable, and others unstable. This means that

*stability is not an intrinsic property of GTLS models,*

but relates to a chosen input/output decomposition. Input/output decompositions are further discussed in Section 3.4.

## 2.5 Comparison with other Methods

In this section we compare GTLS with two well-known methods in time series analysis and system identification, namely ARX (Auto Regression with eXogenous variables) and OE (Output Error). In addition we also discuss a 'local total least squares' method (LTLS), which is a simple modification of the static total least squares method of section 2.3 for the dynamic case. For expository reasons we restrict ourselves to first order models for time series consisting of two components. We assume that the data consists of a time series with components called $u$ and $y$, so

$$w = \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} u(1) & \dots & u(N) \\ y(1) & \dots & y(N) \end{bmatrix} \tag{2.15}$$

The common feature of the above-mentioned methods is that they all relate to linear, time-invariant difference equations for $u$ and $y$, and moreover that they measure deviations from such equations by a least squares criterion. However, they differ in the way deviations are defined, and we will discuss these various approaches as an illustration of the 'global' and 'total' character of GTLS. We remark that ARX and OE are usually presented in a stochastic context, but

they can also be interpreted as methods for minimizing a certain least squares criterion, in which the stochastic interpretation is not essential.
As we restrict ourselves to first order models, the equations take the form

$$\hat{y}(t) = \beta\hat{y}(t-1) + \alpha_0\hat{u}(t) + \alpha_1\hat{u}(t-1). \tag{2.16}$$

Here we have scaled the coefficient of $y(t)$ to one, which means that in this discussion we do not consider the case where this would be zero. In general the data does not satisfy any relation of this type exactly, and the crucial question is, how to define the accuracy of the equation with respect to the data.
In ARX modelling one takes into account the first step ahead prediction error for one of the variables ($y$, say). This corresponds to defining deviations as

$$\epsilon_{arx}(t) := y(t) - \hat{y}(t) \text{ with } \hat{y}(t) = \beta y(t-1) + \alpha_0 u(t) + \alpha_1 u(t-1) \tag{2.17}$$

for $t = 2, \ldots, N$, so the equation is considered as a prediction mechanism. The aim is to determine parameter values such that the sum of squares of the deviations is minimized. The corresponding estimation problem is solved by regression of $y(t)$ on $y(t-1)$, $u(t)$ and $u(t-1)$.
The LTLS method also takes into account the deviations from an equation for each time separately, as in ARX, but it allows for deviations in all variables occurring in (2.16), so both in $u$ and $y$ and both in past and present. The optimal first order LTLS model equals the optimal static model for the sequence $H(t) = \begin{bmatrix} w(t-1) \\ w(t) \end{bmatrix}, t = 2, \ldots, N$, which is determined by the SVD, cf. Section 2.3. This means that the deviations are defined as a four-dimensional sequence

$$\epsilon_{ltls}(t) := \begin{bmatrix} u(t) - \hat{u}(t) \\ y(t) - \hat{y}(t) \\ u(t-1) - \hat{u}_p(t-1) \\ y(t-1) - \hat{y}_p(t-1) \end{bmatrix} \text{ with } \hat{y}(t) = \beta\hat{y}_p(t-1) + \alpha_0\hat{u}(t) + \alpha_1\hat{u}_p(t-1), \tag{2.18}$$

with $\hat{u}, \hat{y}, \hat{u}_p, \hat{y}_p$ such that the sum of squares of the deviations is minimal. Notice that in general $\hat{u}(t) \neq \hat{u}_p(t)$ and $\hat{y}(t) \neq \hat{y}_p(t)$, so that $\hat{u}, \hat{y}$ is not a *global* solution of the difference equation.
In the output error method one takes into account how much $y$ deviates from the effect of $u$ according to the equation (2.16), i.e.,

$$\epsilon_{oe}(t) := y(t) - \hat{y}(t) \text{ with } \hat{y}(t) = \beta\hat{y}(t-1) + \alpha_0 u(t) + \alpha_1 u(t-1) \tag{2.19}$$

for $t = 2, \ldots, N$. Again the sum of squares is minimized, now over the parameters and an initial value for $\hat{y}(1)$. This means that in OE equations are interpreted as input/output systems that map the input $u$ to the output $y$. A more conventional notation for (2.19) in this approach would be in transfer function notation as

$$\hat{y} = \frac{\alpha_0 + \alpha_1\sigma^{-1}}{1 - \beta\sigma^{-1}}u, \tag{2.20}$$

where $\sigma^{-1}$ denotes the inverse time shift $\sigma^{-1}f(t) = f(t-1)$. For an iterative algorithm for determining locally optimal OE models we refer to the overview in [3].
In the GTLS method we allow for deviations in both $u$ and $y$, so in this case the deviation also consists of two components, defined by

$$\epsilon_{gtls}(t) := \begin{bmatrix} u(t) - \hat{u}(t) \\ y(t) - \hat{y}(t) \end{bmatrix} \text{ with } \hat{y}(t) = \beta\hat{y}(t-1) + \alpha_0\hat{u}(t) + \alpha_1\hat{u}(t-1) \quad (2.21)$$

for $t = 1, \ldots N$. Now the aim is to minimize the sum of squares of the deviations over the parameters and over all values $\hat{u}(t)$ and $\hat{y}(t)$ satisfying the equation. So in GTLS we view the equation as the description of a solution set. This shows that from a logical point of view the GTLS approach is the most obvious choice *as a measure of the accuracy of an equation* among the three least squares methods. Namely, it measures the distance between the data and that what is in correspondence with all logical consequences of the equation.
Notice that both ARX and OE make an a priori distinction between variables into exogenous/endogenous or inputs/outputs, while GTLS treats both variables on the same footing. This corresponds to the distinction between *ordinary* and *total* least squares in static modelling, cf. Section 2.3, so we could call ARX and OE 'ordinary least squares methods'. Further, LTLS and ARX models are only based on the first step ahead predictions of equations, ignoring their implications over a longer lag, and therefore we call these local methods. For instance, by adding (2.16) for two subsequent time-instants we obtain

$$\hat{y}(t) = (\beta-1)\hat{y}(t-1)+\beta y(t-2)+\alpha_0\hat{u}(t)+(\alpha_0+\alpha_1)\hat{u}(t-1)+\alpha_1 u(t-2), \quad (2.22)$$

but this second order equation is not taken into account, although it is a logical consequence of (2.16). This is in contrast with GTLS and OE, in which respectively $(\hat{u}, \hat{y})$ as defined in (2.21) and $(u, \hat{y})$ in (2.19) are *global* solutions of the equations, and hence satisfy also all higher order implications. It seems to be inherent in this global character that the corresponding estimation problems are relatively complex. The solution of the GTLS problem is the main issue of this monograph.
We summarize the comparison in the following table.

**Table 2.2**: Comparison of least squares methods.

| least squares methods | *ordinary* | *total* |
|---|---|---|
| *local* | ARX | LTLS |
| *global* | OE | GTLS |

We remark that there are several other results on dynamic modelling with a symmetric treatment of all data components, but these methods do not give a decomposition in terms of global solutions of difference equations. For instance, the symmetric modelling approach described in [13] amounts to the

decomposition of dynamic stochastic processes into a structured part, satisfying deterministic relations, and an error system. However, this dynamic 'error in variables' approach is not compatible with GTLS, as the analysis is performed for fixed frequencies and without an order restriction on the relations. The dynamic extension of the Frisch scheme, introduced in [5], concerns the reconstruction of dynamic relations under corruption of all components by white noise, but is not applicable to general data. Also the stochastic realization problems treated in [42] and dynamic factor analysis as proposed in [9, 23] are quite different from GTLS, cf. also [21].

## 2.6 Limitations

In this section we want to delimit the scope of this monograph in order not to raise expectations that we cannot come up to. We do not address the first stages of time series analysis, involving data acquisition and preliminary manipulations, but simply take our starting point in a given time series. Theoretically the method is applicable to every multivariable time series, but this is of course no guaranty that it yields valuable models in all cases.

Not only the scope of the GTLS method is limited, we should also clearly state that we are not able to give a complete solution of the GTLS problem (Definition 2.1.7). As is common for non-linear optimization problems, a one shot algorithm that determines a globally optimal model seems inachievable. We present an iterative algorithm for determining locally optimal models. The the solution will then depend on the initial models in which the algorithms are started, and therefore we also briefly discuss a heuristic method for determining reasonable starting points. We do not address the issue of determining the number of locally optimal systems.

Secondly, for technical reasons concerning the numerical representation of systems we will exclude so called non-stabilizable systems (stabilizability is defined in Definition 3.2.3, see also the list of basic notions). In such systems an exponential growth of most system trajectories cannot be avoided, despite the degrees of freedom in the model. This is a highly specific case, as almost all systems in the GTLS model class are not of this type. Moreover, by exploiting the fact that GTLS is completely symmetric in time, in fact we only have to exclude systems that are neither stabilizable nor 'anti-stabilizable', which means stabilizable for the reversed time direction.

Finally we want to point at a limitation that is inherent in the GTLS method, but perhaps somewhat implicitly. GTLS is only applicable for data consisting of time series of two or more components, as for a time series with one component ($q = 1$), the GTLS problem becomes rather trivial. Namely, we either have to allow for systems with rank one ($m = 1$), which are trivial systems with behaviour consisting of all single component time series, or we have to impose a zero rank, which corresponds to autonomous systems whose behaviour is entirely determined by some initial conditions. The latter case might have some practical relevance, but in general these models are not flexible enough to mo-

del the data. So the simplest non-trivial case we consider concerns time series with two components ($q = 2$) and models with tolerated rank one ($m = 1$).

# Chapter 3

# Isometric State Representations

In the previous chapter we described the GTLS method in purely geometrical terms. Systems are defined as sets of time series, and we aim for a linear, shift-invariant system of restricted dimension that has minimal distance to given data.
In this chapter we introduce *isometric state representations* (ISR's), which form the main tool in the construction of GTLS models from data.
A representation of a system is a concrete description of its behaviour. In the definition of the GTLS problem we need not make any reference to representations. However, they do become relevant as soon as we want to specify a model, i.e., if we want to indicate which time series belong to it and which not. There are many ways to represent a linear, time-invariant system. They can be described as the solution set of difference equations, cf. Section 2.4.2, or by more structured representations that make explicit the memory (or state) of a system or possible input/output decompositions of the system variables. For an overview of system representations we refer to [44, 46]. As far as the specification of systems is concerned, the type of representation is hardly relevant. Different representations of the same system are just different ways of describing the same object, and moreover algorithms exist for converting one type into another.
However, from a practical point of view the type of representation is important. Representations play a basic role in the implementation of algorithms, and an appropriate choice may streamline algorithms considerably. We choose to use ISR's as they turn out to form a highly appropriate tool in the GTLS method. They are based on the following principles.

1. The memory of a system is made explicit by *state variables*.

The state of a system is one of the central notions in systems theory, which has proved its merits both in identification and control. The state represents all features of the past behaviour of a system that are relevant for its future evolution, so we may say that it represents the heart of the system dynamics. A mathematical characterization is that state representations describe systems

in terms of first order relations. In the development of the GTLS algorithms the notion of state turns out to be essential at various points.

2. Symmetry in the system variables.

The best-known type of representations with state variables are the so called input/state/output (i/s/o-) representations. These are based on an a priori decomposition of system variables between inputs and outputs. The inputs are free, unrestricted variables, which determine the output (modulo initial conditions). In the GTLS method we do not make a preliminary distinction into inputs and outputs, but we treat all components of an observed time series on the same footing. It is then natural to reflect this in the representations also.

3. Indeterminism is made explicit by *auxiliary inputs*.

The systems in the GTLS model class are indeterministic, as they contain a degree of freedom at each time instant, which is reflected by their rank. This indeterminism is reflected by *auxiliary* inputs, which are auxiliary variables whose value determines a system trajectory (again modulo initial conditions). We use the term *state representation* (SR) for descriptions that make use of state variables and auxiliary inputs. SR's are introduced in [44], where they are called state space systems. We discuss them in the next section.
The absence of an input/output decomposition is an innovative aspect of SR's with respect to the mainstream approach in system theory. In order to give some intuition on SR's, the precise relation with the classical i/s/o-representations is explained in Section 3.4.

4. Each state component and each auxiliary input has a *normalized* effect on the external behaviour of the system.

State representations are highly non-unique, and this can be exploited to impose that the auxiliary variables should have a normalized effect on the system trajectories, which means that the norm of an auxiliary input equals the norm of its effect on system trajectories. This restriction can be imposed almost completely without loss of generality, as we only have to exclude so-called non-stabilizable systems. ISR's turn out to be useful in the evaluation of the misfit of systems, and they have convenient properties that play a role in the iterative algorithms for the construction of GTLS models. We introduce ISR's in Section 3.2.
SR's with only a normalization condition on the auxiliary inputs are used in [47]. The innovative aspect of ISR's is that all auxiliary variables, both the state and the auxiliary inputs, have a normalized effect. To our knowledge their use in identification is new.

## 3.1 State Representations

State representations are defined as follows.

**Definition 3.1.1 (State Representation)**
*A state representation $(A, B, C, D)$ of a system $\mathcal{B}$ is a description of the form:*

$$\begin{aligned}\mathcal{B} = \{&\hat{w} : \mathsf{Z} \to \mathsf{R}^q; \\ &\textit{there exists a state trajectory } \hat{x} : \mathsf{Z} \to \mathsf{R}^n \textit{ and} \\ &\textit{an auxiliary input } \hat{v} : \mathsf{Z} \to \mathsf{R}^m \textit{ such that} \\ &\qquad \begin{aligned} \sigma\hat{x} &= A\hat{x} + B\hat{v} \\ \hat{w} &= C\hat{x} + D\hat{v}\}, \end{aligned} \qquad (3.1)\end{aligned}$$

*with $A \in \mathsf{R}^{n\times n}$, $B \in \mathsf{R}^{n\times m}$, $C \in \mathsf{R}^{q\times n}$, $D \in \mathsf{R}^{q\times m}$ and $n$,$m \in \mathsf{N}$. This system is denoted as $\mathcal{B}(A, B, C, D)$.*

The variables $\hat{x}$ and $\hat{v}$ are auxiliary variables with the following interpretation. The auxiliary input $\hat{v}$ is a free variable representing the indeterminism in the system at each time-instant. The state $\hat{x}$ represents the memory of the system. Two system trajectories with the same state at a certain time instant allow for exactly the same continuations in the system from that time instant on, so the state contains all information from the past that has implications for the future. Another interpretation is that the state at time $t_0$, $\hat{x}(t_0)$, represents the degrees of freedom due to initial conditions at $t_0$, i.e., the additional freedom from $t_0$ on if the value of the time series before $t_0$ is not specified.
In order to give some further intuition, we briefly consider the two simple cases where either there are no states ($n = 0$) or there are no auxiliary inputs ($m = 0$). In the first case SR's take the form $(-, -, -, D)$, with $-$ the empty matrix, so $\hat{w}(t) = D\hat{v}(t)$. This describes the static system with system trajectories taking their value in the image of $D$ at each time. This illustrates that in the dynamic case there is indeterminism in the direction of the image of $D$ at each time. If there are no auxiliary inputs, SR's are of the form $(A, -, C, -)$, describing autonomous systems in which there are only degrees of freedom due to initial conditions.
The class of systems that can be represented by SR's is precisely the class of linear, shift-invariant, complete systems. A feature of SR's that is of crucial importance in the GTLS approach is that the number of auxiliary inputs and state variables correspond to respectively the rank and degree of a system.

**Proposition 3.1.2 (State Representation of Systems)**

1. *Every linear, shift-invariant complete system admits a state representation.*

2. *Conversely, every state representation corresponds to a linear, shift-invariant complete system.*

3. *Systems $\mathcal{B} \in \bar{\mathsf{B}}^{q,m,n}$, i.e., with rank at most $m$ and degree at most $n$, are precisely those systems that can be described by a state representation with $m$ auxiliary inputs and $n$ state variables.*

*Proof.* See Appendix A.3.

SR's for a system are highly non-unique. SR's are called equivalent if they represent the same system. One source of non-uniqueness is the possible presence of ineffective auxiliary variables, i.e., state components or auxiliary inputs that have no effect on system trajectories. We call an SR minimal if both the number of states and auxiliary inputs are minimal. From Proposition 3.1.2.3 it follows that in a minimal SR the number of auxiliary inputs must equal the rank of a system, and the state dimension the degree. For future reference we formulate this fact as a corollary.

**Corollary 3.1.3 (Minimal SR)** *In a minimal state representation of a system $\mathcal{B}$ the number of auxiliary inputs equals the rank of $\mathcal{B}$ , and the number of state variables equals the degree of $\mathcal{B}$ .*

Even minimal SR's are highly non-unique. From a given SR we obtain equivalent ones as follows. For expository reasons we display the system matrices in one block-matrix.

**Proposition 3.1.4 (Equivalent State Representations)**
*The state representations*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \text{ and } \begin{bmatrix} S(A+BF)S^{-1} & SBR \\ (C+DF)S^{-1} & DR \end{bmatrix}$$

*represent the same system for all invertible $S \in \mathsf{R}^{n\times n}$, invertible $R \in \mathsf{R}^{m\times m}$, and $F \in \mathsf{R}^{m\times n}$. Moreover, if $(A, B, C, D)$ is minimal, then all minimal equivalent state representations are obtained in this way.*

*Proof.* See Appendix A.3.

So the non-uniqueness of minimal SR's can be assigned to three sources:

- a basis transformation of the state $\hat{x} \to S\hat{x}$,
- a basis transformation of the auxiliary inputs $\hat{v} \to R^{-1}\hat{v}$,
- addition of state components to the auxiliary input $\hat{v} \to \hat{v} - F\hat{x}$.

The choice of basis in the state space, corresponding to $S$, is a well-known non-uniqueness of state space representations. In our framework the auxiliary input $\hat{v}$ is merely a tool to describe the system behaviour, and need not have external significance. This allows for a basis transformation for the auxiliary input, corresponding to $R$. Further the behaviour is invariant under a static state feedback $F$ to this auxiliary input. This is in contrast with the common notion of feedback to the actual input of the system, which would affect the

set of compatible input-output pairs. As a consequence, in our framework the spectrum of the $A$-matrix is not an intrinsic property of a system. In the next section we exploit this non-uniqueness in order to obtain SR's with convenient properties for computing the misfit of a model with respect to data.
As stated before, in minimal SR's the number of auxiliary variables corresponds to the size of a system. Minimality can also be expressed more concretely in terms of rank conditions on the matrices $(A, B, C, D)$. We mention the following characterization, as it is will play a role in the construction of ISR's. The observability condition for a pair of matrices is spelled out in the proof.

**Proposition 3.1.5 (Minimal State Representations)**
*A state representation $(A, B, C, D)$ is minimal if and only if:*

1. *the matrix $[A\ B] \in \mathsf{R}^{n\times n+m}$ has full row rank $n$*
2. *for all $F \in \mathsf{R}^{m\times n}$ is $(A + BF, C + DF)$ observable*
3. *$D$ has full column rank $m$.*

*Proof.* See Appendix A.3.

The proposition indicates three types of ineffective auxiliary variables. Firstly, some components of the state might be zero, namely if the matrix $[A\ B]$ contains zero rows, and these variables can be removed. Similarly, if this matrix has not full row rank $n$, this implies that some linear combinations of the state variables are zero, and then their number can also be reduced. The second condition excludes that the future effect of the state can be cancelled by a certain combination of auxiliary inputs. The third source of non-minimality concerns the case in which some auxiliary inputs have no or only a delayed effect.

**Leading Example.** We describe a simple example, which will be used in the following sections to clarify the introduced general framework. Consider the linear time-invariant system

$$\begin{aligned}\mathcal{B}_{ex} := \{\hat{w} : \mathsf{Z} \to \mathsf{R}^2; \hat{w} = (\hat{u}, \hat{y}) \text{ with} \\ \hat{y}(t) = 2/3\,\hat{y}(t-1) + 2\hat{u}(t) - 2\hat{u}(t-1)\}.\end{aligned} \tag{3.2}$$

The state of a system should represent all information from the past that is relevant for the future. The relation between past and future in $\mathcal{B}$ can be expressed as

$$2/3\,\hat{y}(t-1) - 2\hat{u}(t-1) = \hat{y}(t) - 2\hat{u}(t), \tag{3.3}$$

so we can define for instance $\hat{x}(t) = \hat{y}(t-1) - 3\hat{u}(t-1)$, cf. also the proof of Proposition 3.1.2.1. This can be written in input/state/output form as

$$\hat{x}(t+1) = 2/3\,\hat{x}(t) - \hat{u}(t);\ \hat{y}(t) = 2/3\,\hat{x}(t) + 2\hat{u}(t). \tag{3.4}$$

From this it is easy to obtain an SR by taking the auxiliary input $\hat{v}$ to be equal to $\hat{u}$, which gives

$$\mathcal{B}_{ex} = \mathcal{B}(2/3, -1, \begin{bmatrix} 0 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}). \tag{3.5}$$

The representation (3.5) is minimal. An SR equivalent to (3.5) is given by $(1, -1/2, \begin{bmatrix} -1/3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \end{bmatrix})$, which follows from Proposition 3.1.4 by taking $S = 1$, $R = 1/2$ and $F = -1/3$. Note that in this representation the auxiliary input is given by $y$. This also illustrates that the eigenvalues of $A$ are not intrinsic for the system $\mathcal{B}$. $\diamond$

## 3.2 Isometric State Representations

In this section we define isometric state representations (ISR's), which form the cornerstone in the construction of GTLS models. ISR's are a special type of SR's defined by a local isometry property involving the state variable.

**Definition 3.2.1 (Isometric State Representation)** *A state representation $(A, B, C, D)$ is called isometric if for all $x \in \mathsf{R}^n$,$v \in \mathsf{R}^m$, $w \in \mathsf{R}^q$ and $z \in \mathsf{R}^n$ such that $z = Ax + Bv$ and $w = Cx + Dv$ it holds*

$$|v|^2 + |x|^2 = |w|^2 + |z|^2. \tag{3.6}$$

*Equivalently,*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^\top \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix}. \tag{3.7}$$

We illustrate this definition by considering the static case ($n = 0$) and the autonomous case ($m = 0$). ISR's without state variables just consist of an isometric matrix $D$, i.e. with $D^\top D = I_m$. This describes the system

$$\mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ \exists \hat{v} : \mathsf{Z} \to \mathsf{R}^m \text{ such that } \hat{w}(t) = D\hat{v}(t) \forall t \in \mathsf{Z}\}. \tag{3.8}$$

Notice that the auxiliary input has a normalized effect, as $|\hat{w}(t)| = |D\hat{v}(t)| = |\hat{v}(t)|$. In static ISR's this effect is even immediate, while in the dynamic case it is in general distributed over different time instants. ISR's without auxiliary inputs take the form $(A, -, C, -)$ with $A^\top A + C^\top C = I_n$, describing the autonomous system

$$\begin{aligned} \mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ &\exists \hat{x} : \mathsf{Z} \to \mathsf{R}^n \text{ such that} \\ &\hat{x}(t+1) = A\hat{x}(t);\ \hat{w}(t) = C\hat{x}(t) \forall t \in \mathsf{Z}\}. \end{aligned} \tag{3.9}$$

The system trajectories in this system are completely determined by the value of the state at one time instant. The isometry condition implies that each initial state has a normalized future effect on system trajectories, i.e., the norm of a system trajectory on $\mathsf{Z}^+$ equals the norm of the corresponding initial state

at $t = 0$, at least if $A$ is asymptotically stable. Notice that if $A$ is singular, some states cannot occur in (3.9). The exceptional cases in which $A$ is singular or not asymptotically stable is further discussed later on.
The local property (3.6) that characterizes ISR's induces an isometry on finite intervals that will play a central role in the sequel.

**Proposition 3.2.2 (Isometry on Finite Time)**
*Let $\mathcal{B}$ denote a system with isometric state representation $(A, B, C, D)$, and let $\hat{v}$ and $\hat{x}$ denote auxiliary input and state trajectory corresponding to a system trajectory $\hat{w} \in \mathcal{B}$ in this representation. Then for all finite intervals $[1, N] \subset \mathsf{Z}$ it holds that*

$$\begin{aligned} |\hat{x}(1)|^2 + \quad & |\hat{v}(1)|^2 + \ldots + |\hat{v}(N)|^2 \quad = \\ & |\hat{w}(1)|^2 + \ldots + |\hat{w}(N)|^2 \quad + |\hat{x}(N+1)|^2. \end{aligned} \tag{3.10}$$

*Proof.* This follows immediately from the summation of (3.6) with $x = \hat{x}(t)$ and $z = \hat{x}(t+1)$ over time instants $t \in T$. ♣

This shows that the state and the auxiliary input have a normalized effect on the system trajectories. Namely, the squared norm of $\hat{w}$ on finite intervals equals the squared norm of the auxiliary variables and initial state, minus the squared norm of the final state which will have an effect after the interval.
**Remark.** In system theoretic terminology, ISR's are realizations of inner (stable, all-pass) transfer functions with an output normal state. To our knowledge, its use in identification is new. ◇

The isometry property (3.7) can be imposed almost without loss of generality. The exceptional cases concern systems that are not stabilizable.

**Definition 3.2.3 (Stabilizability)** *A system is called stabilizable if all trajectories on finite time intervals admit a continuation within the system that converges to zero.*

This corresponds to the classical definition in terms of inputs and states, imposing that the state in a minimal input/state/output representation can be made converging to zero by choosing appropriate inputs. Examples of non-stabilizable systems are for instance $\mathcal{B}(\alpha, 0, 1, 0)$ with $\alpha \geq 1$, which consists of exponentials $\hat{w}(t) = c\alpha^t$ with $c \in \mathsf{R}$.
Notice that by choosing zero auxiliary inputs in an ISR from a certain time on the isometry property (3.6) guarantees the convergence of the corresponding system trajectory to zero. So clearly an ISR cannot represent a system that is not stabilizable. The next proposition states that this forms the only exception.

**Proposition 3.2.4 (Existence of ISR)** *A linear, shift-invariant, complete system admits an isometric minimal state representation if and only if it is stabilizable.*

*Proof.* Every system that can be described by an ISR is stabilizable, as every system trajectory in it can be made converging to zero by applying zero auxiliary inputs from a certain time on. Namely, then the lefthand side in (3.10) is bounded by a finite value for all $N \in \mathsf{Z}^+$, which implies that $\lim_{N\to\infty} \hat{w}(N) = 0$. The fact that all stabilizable systems admit an ISR is proved by construction in the next proposition. ♣

**Proposition 3.2.5 (Construction of ISR)** *Let $(A, B, C, D)$ be a minimal state representation of a stabilizable system, and let $K \in \mathsf{R}^{n\times n}$ be the unique symmetric positive definite solution of the algebraic Riccati equation*

$$K = A^\top KA - (B^\top KA + D^\top C)^\top (B^\top KB + D^\top D)^{-1}(B^\top KA + D^\top C) + C^\top C. \quad (3.11)$$

*Let the matrices $S \in \mathsf{R}^{n\times n}$, $F \in \mathsf{R}^{m\times n}$ and $R \in \mathsf{R}^{m\times m}$ be solutions of the equations:*

$$\begin{aligned} S^\top S &= K & (3.12)\\ RR^\top &= (B^\top KB + D^\top D)^{-1} & (3.13)\\ F &= -(B^\top KB + D^\top D)^{-1}(B^\top KA + D^\top C). & (3.14) \end{aligned}$$

*Then $(S(A + BF)S^{-1}, SBR, (C + DF)S^{-1}, DR)$ is an equivalent isometric minimal state representation.*

*Proof.* See Appendix A.3.
*Implementation:* `SR2ISR` in Appendix B.1.

The isometry condition (3.7) reduces the non-uniqueness of SR's considerably, but also minimal ISR's are not uniquely determined by a system, as we can apply unitary basis transformations on the auxiliary variables. A matrix $M$ is called unitary if $M^\top = M^{-1}$.

**Proposition 3.2.6 (Equivalence of ISR's)** *The isometric state representations*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \text{ and } \begin{bmatrix} UAU^\top & UBV \\ CU^\top & DV \end{bmatrix}$$

*represent the same system for all unitary matrices $U$ and $V$. Moreover, if $(A, B, C, D)$ is minimal, then all minimal equivalent isometric state representations are obtained in this way.*

*Proof.* See Appendix A.3.

So the auxiliary variables are uniquely determined by system trajectories modulo norm-preserving transformations. This illustrates the fact that the auxiliary variables in a minimal ISR do not only have a qualitative interpretation, as in every minimal SR, but also a quantitative meaning, cf. (3.10).
One might wonder if ISR's are not by definition minimal, as every auxiliary variable should have a normalized effect on the system variables. Indeed, the

number of auxiliary variables in every ISR is minimal. This is obvious for the static case ($n = 0$), and is proved below for the dynamic case. However, the isometry property (3.6) does not exclude the presence of ineffective state components that retain all 'energy' for themselves, *ad infinitum*. As an illustration, consider the trivial ISR

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{3.15}$$

which is not minimal as it represents the trivial static system consisting of all time series with one component. Similarly, all ISR's corresponding to a square unitary matrix correspond to trivial systems, as they have rank equal to the number of system variables, and hence do not represent any restriction.
In Proposition 3.1.5 minimality conditions are formulated in terms of the system matrices in an SR, and these also apply to ISR's. The combination of the isometry condition and minimality has some implications that we want to make explicit as they play a role in the sequel. We also give an alternative characterization of minimality for ISR's. A matrix is called (asymptotically) stable if it has no eigenvalues with absolute value larger than (or equal to) one.

**Proposition 3.2.7 (Minimality of ISR's)**

1. *In a state representation that is isometric the number of auxiliary inputs is minimal.*

2. *In an isometric state representation $(A, B, C, D)$, the matrix $A$ is stable. If the representation is minimal then $A$ is asymptotically stable.*

3. *Define for an ISR $(A, B, C, D)$ with state dimension $n$ the matrix $W := [A^n \vdots A^{n-1}B \vdots \ldots \vdots B]$. Then $(A, B, C, D)$ is minimal if and only if $W$ is non-singular and $WW^\top$ asymptotically stable.*

*Proof.* See Appendix A.3.

The characterization of minimality for ISR's can be interpreted as follows. The matrix $W$ denotes the mapping from an initial state ($\hat{x}(1)$, say) and auxiliary inputs on the interval $[1, n]$ to the resulting final state $\hat{x}(n + 1)$, so

$$W : \{\hat{x}(1); \hat{v}(1), \ldots, \hat{v}(n)\} \to \hat{x}(n + 1). \tag{3.16}$$

Now if $W$ is singular, this implies that some states are not reachable. Even if we would start the system in an unreachable state, i.e. a state that is not contained in the image of $W$, the system would leave this state and never return to it. If $WW^\top$ is not asymptotically stable, this would mean that there exists a combination of initial state and auxiliary inputs that store all their 'energy' in the final state, so that there is nothing left for the system trajectory.

**Leading Example - Continued.** We apply the construction of Proposition 3.2.5 to the SR (3.5) of the system $\mathcal{B}_{ex}$. This yields $K = 4/9$, $S = 2/3$, $R = 3/7$

and $F = 4/21$, resulting in the isometric representation

$$\sigma\hat{x} = 6/7\,\hat{x} - 2/7\,\hat{v};\ \hat{w} = \begin{bmatrix} -2/7 \\ 3/7 \end{bmatrix}\hat{x} + \begin{bmatrix} 3/7 \\ 6/7 \end{bmatrix}\hat{v} \tag{3.17}$$

According to Proposition 3.2.7, $\mathcal{B}_{ex}$ has a unique ISR with state dimension one, modulo sign changes for the state and for the auxiliary input. $\diamond$

## 3.3 State Representations on Finite Time

In this section we discuss a subtlety concerning the representation of systems on finite intervals. It will not play an important role in the sequel, and may be skipped in first reading.
The GTLS problem is formulated in terms of the behaviour of systems on finite time intervals. Therefore, with an SR we will also associate a behaviour on finite intervals, as follows.

**Definition 3.3.1 (State Representation for Finite Intervals)**
*A state representation $(A, B, C, D)$ of the restriction of a system $\mathcal{B}$ to a finite time interval $T = [t_0, t_1] \subset \mathsf{Z}$ is a description of the form*

$$\begin{aligned} \mathcal{B}_T = \\ &\{\hat{w} : T \to \mathsf{R}^q;\ \exists\, \hat{x}(t_0) \in \mathsf{R}^n, \hat{v} : T \to \mathsf{R}^m \text{ such that} \\ &\hat{x}(t+1) = A\hat{x}(t) + B\hat{v}(t) \text{ and } \hat{w}(t) = C\hat{x}(t) + D\hat{v}(t)\}. \end{aligned} \tag{3.18}$$

*We write $\mathcal{B}_T = \mathcal{B}_T(A, B, C, D)$.*

Roughly speaking it does not make any difference whether we consider SR's for finite time or for behaviours on $\mathsf{Z}$. The restriction of a system to a certain interval just corresponds to a restriction of the auxiliary variables to the same interval, except for some degenerate cases. This is made precise below. We remark that the exceptional cases only play a minor role in the sequel.

**Proposition 3.3.2 (SR's for $\mathsf{Z}$ and Finite Time)**
*Let $\mathcal{B}$ denote a linear, shift-invariant complete system of degree $n$. Then*

$$\mathcal{B}_{[1,n+1]} = \mathcal{B}_{[1,n+1]}(A, B, C, D) \ \Rightarrow\ \mathcal{B} = \mathcal{B}(A, B, C, D). \tag{3.19}$$

*Conversely, if $(A, B, C, D)$ is minimal, then*

$$\mathcal{B} = \mathcal{B}(A, B, C, D) \ \Rightarrow\ \mathcal{B}_T = \mathcal{B}_T(A, B, C, D) \tag{3.20}$$

*for all finite intervals $T \subset \mathsf{Z}$.*

*Proof.* See Appendix A.3.

**Leading Example - Continued.** Consider again the SR (3.5). From (3.20) it follows that this also represents the behaviour on all finite intervals, as described in Definition 3.3.1. Conversely, (3.19) states that every SR that describes $\mathcal{B}_{[1,2]}$ also

describes $\mathcal{B}$. Clearly this is not true for SR's that describe $\mathcal{B}_{[1,1]}$. To illustrate the role of minimality in (3.20) we consider the non-minimal SR for $\mathcal{B}$ given by

$$(\begin{bmatrix} 0 & 0 \\ 0 & 2/3 \end{bmatrix}, \begin{bmatrix} 0 \\ -2/3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix})$$

The first state component is ineffective in representing behaviours on $\mathsf{Z}$. However, in (3.18) we could choose an arbitrary value for this component at $t = 1$, and this would yield additional freedom in the behaviour. It is easily verified that the behaviour described by (3.18) for the non-minimal SR and $T = [1, 2]$ has dimension 4, while $\dim(\mathcal{B}_{[1,2]}) = 3$. For instance, the first space contains the trajectory given by $\hat{u}(1) = 1, \hat{u}(2) = \hat{y}(1) = \hat{y}(2) = 0$, which is not contained in $\dim(\mathcal{B}_{[1,2]})$. $\diamond$

## 3.4 Relation with i/s/o Representations

Input/state/output (i/s/o-) representations are the counterparts of SR's in which there is an explicit partition of system variables into inputs and outputs. In this section we describe how to convert one into the other. The reason to discuss this issue is twofold. By explaining the relation between SR's and i/s/o-representations we aim to give additional intuition for SR's. Further, input/output decompositions play a role in some applications. The translation of SR's into i/s/o-representations shows how to determine such a decomposition. I/s/o-representations are defined as follows.

**Definition 3.4.1 (i/s/o Representations)**
*An input/state/output representation* $(A', B', C', D')$ *of a system* $\mathcal{B}$ *is a description of the form:*

$$\begin{aligned} \mathcal{B} = \{\hat{w} = (\hat{u}, \hat{y}) \text{ with } \hat{u} : \mathsf{Z} \to \mathsf{R}^m \text{ and } \hat{y} : \mathsf{Z} \to \mathsf{R}^p; \\ \text{there exists a state trajectory } \hat{x} : \mathsf{Z} \to \mathsf{R}^n \text{ such that} \end{aligned}$$

$$\begin{aligned} \sigma\hat{x} &= A'\hat{x} + B'\hat{u} \\ \hat{y} &= C'\hat{x} + D'\hat{u}, \end{aligned} \tag{3.21}$$

*with* $A' \in \mathsf{R}^{n\times n}$, $B' \in \mathsf{R}^{n\times m}$, $C' \in \mathsf{R}^{p\times n}$, $D' \in \mathsf{R}^{p\times m}$ *and* $n, m, p \in \mathsf{N}$. *This system is denoted as* $\mathcal{B}^{i/o}(A', B', C', D')$.

The inputs $\hat{u}$ are unrestricted or free variables, that determine the outputs $\hat{y}$ of the system modulo initial conditions. Notice that the difference with SR's as defined in Definition 3.1.1 is not the shape of the system matrices, but their interpretation. An i/s/o-representation of a system is also an SR for a different system. In order to avoid confusion we use primes for the system matrices in i/s/o-representations.
The construction of SR's from i/s/o-representations is really simple. The explicit input/output decomposition can be eliminated by taking the auxiliary input equal to the actual input $\hat{u}$ .

**Proposition 3.4.2 (From i/s/o to SR)** *Let $(A', B', C', D')$ denote an input/state/output representation of a system $\mathcal{B}$ . Define $C := \begin{bmatrix} 0 \\ C' \end{bmatrix}$ and $D := \begin{bmatrix} I_m \\ D' \end{bmatrix}$. Then $(A', B', C, D)$ is a state representation for $\mathcal{B}$ .*

*Proof.* This follows immediately from the fact that the auxiliary input in the SR equals the input in the i/s/o-representation. ♣
*Implementation:* `ISO2SR` in Appendix B.1.

The converse question, how to determine an i/s/o-representation from a given SR, is only slightly more involved. This amounts to determining a valid input/output decomposition of the system variables $\hat{w} = (\hat{u}, \hat{y})$, i.e. with unrestricted input components $\hat{u}$ that determine the remaining output components $\hat{y}$, up to initial conditions. In view of Proposition 3.1.5 we can restrict the attention to SR's with an injective $D$-matrix, and for simplicity we assume that the first $m$ rows of $D$ are independent.

**Proposition 3.4.3 (From SR to i/s/o)** *Let $(A, B, C, D)$ be a minimal state representation of a system $\mathcal{B}$ , with $C = \begin{bmatrix} C_u \\ C_y \end{bmatrix}$, $D = \begin{bmatrix} D_u \\ D_y \end{bmatrix}$, and $D_u$ invertible. Then $(A - BD_u^{-1}C_u, BD_u^{-1}, C_y - D_yD_u^{-1}C_u, D_yD_u^{-1})$ is an input/state/output representation of $\mathcal{B}$ .*

*Proof.* The auxiliary input $\hat{v}$ in the SR and the input $\hat{u}$ in the i/s/o-representation are related by $\hat{u} = C_u\hat{x} + D_u\hat{v}$. Substituting $\hat{v} = D_u^{-1}(C_u\hat{x} - \hat{u})$ in the SR equations (3.1) gives the result. ♣.
*Implementation:* `SR2ISO` in Appendix B.1.

This explains how the system matrices in an SR are related to i/s/o-representations. Perhaps even more enlightening is the relation between the auxiliary variables in both type of representations. Notice that the state variables in both representations completely coincide, i.e. $\hat{x}$ is a state trajectory for $\hat{w} \in \mathcal{B}$ in a minimal i/s/o-representation if and only it is one in a minimal SR. Further, the auxiliary input in an SR can be chosen equal to the actual input components of a system trajectory. The extra freedom in the auxiliary inputs in SR's as compared to the actual inputs in i/s/o-representations is the addition of state components and a basis transformation, i.e., $\hat{v} = R(\hat{u} - F\hat{x})$ for some $F$ and invertible $R$.
As the number of (auxiliary) variables in minimal representations of both types are equal, this means that the rank and degree of a system equal the number of inputs and states in a minimal i/s/o-representation.

**Leading Example - Continued.** An example of the construction of an i/s/o-representation from an SR has already been given in the leading example of Section 3.1. We illustrate the converse direction by applying Proposition 3.4.3

to the ISR given in (3.17). For input $u$ the construction yields the i/s/o-representation $(2/3, -2/3, 1, 2)$, which is indeed in correspondence with (3.4). Both entries in $D$ are non-zero, so we could also take $y$ as input variable. $\diamond$

# Chapter 4

# Evaluation of the Misfit

In this chapter we treat the first part of the GTLS problem: how to evaluate the misfit of a given system with respect to an observed time series. Stated otherwise, we determine how to change the data by a minimal amount in order to make it satisfy the laws of a given system. The distinctive aspect of the GTLS approach is that we consider the *global* approximation of the data by one system trajectory that belongs to the system on the entire observation interval, instead of taking into account only local equation misfits as in prediction error oriented methods. Another appealing feature of the GTLS misfit is that it measures the accuracy of the system laws *simultaneously*, not in terms of errors with respect to the individual equations.
Due to this global character the evaluation of the misfit for a given system is not a trivial problem, in contrast to for instance the evaluation of the first step ahead prediction errors of a given system. On the other hand, it concerns a linear optimization problem that can be solved in a straightforward way.
We present an algorithm in terms of ISR's that were introduced in the previous chapter. The algorithm is based on a simple formula for the orthogonal projection onto linear spaces in terms of isometric operators. We also describe the orthogonal complement of systems by ISR's, which reveals the structure of the approximation error.

## 4.1 Orthogonal Projection onto Systems

In this section we consider the first part of the GTLS problem, namely the evaluation of the misfit (Definition 2.1.5). We determine this misfit by constructing the optimal approximation of an observed time series by a system trajectory in a given linear, shift-invariant system. In fact this amounts to the simple mathematical problem of projecting a point (the data) onto a linear space (the system), but we pay some extra attention to the use of ISR's in this.
Our approach is based on the following result on the orthogonal projection onto linear spaces. Let $G : \mathsf{R}^k \to \mathsf{R}^\ell$ denote a linear function. The *adjoint* of $G$ is a linear function $G^* : \mathsf{R}^\ell \to \mathsf{R}^k$ defined by the condition $\langle w, Gv \rangle = \langle G^* w, v \rangle$ for all $w \in \mathsf{R}^k$ and $v \in \mathsf{R}^\ell$, where $\langle ., . \rangle$ denotes the inner product both on $\mathsf{R}^k$

and $\mathsf{R}^\ell$. So the adjoint of a matrix is given by its transpose. The function $G$ is called isometric or norm-preserving if $\langle Gv, Gv\rangle = \langle v, v\rangle$. It follows that $G$ is isometric if and only if $G^*G = I_k$.
The orthogonal projection onto a linear space admits a simple expression in terms of an isometric function with image equal to that space.

**Lemma 4.1.1 (Orthogonal Projection)** *Let $\mathcal{M}$ denote a linear subspace of $\mathsf{R}^\ell$, and $G : \mathsf{R}^k \to \mathsf{R}^\ell$ a linear isometric function with image equal to $\mathcal{M}$. The orthogonal projection of $w \in \mathsf{R}^\ell$ onto $\mathcal{M}$ is given by*

$$\hat{w} = GG^*w. \tag{4.1}$$

*Further, $\hat{w}$ is the optimal approximation within $\mathcal{M}$ of $w$ in the Euclidian metric.*

*Proof.* See Appendix A.4.

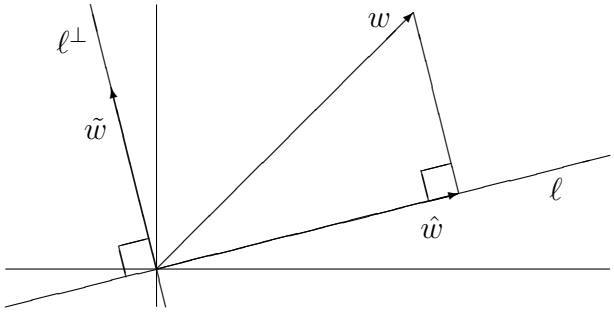
We remark that $GG^*$ is comparable to the so-called 'hat-matrix', a term used in statistical literature to indicate a mapping from data to estimate.

**Example.** We give a simple illustration. Let $\ell$ denote a line in the real plane, described by

$$\ell = \{\hat{w} \in \mathsf{R}^2;\ \text{ there exists } \hat{v} \in \mathsf{R} \text{ such that } \hat{w} = \begin{bmatrix} a \\ b \end{bmatrix} \hat{v}\}$$

with $a, b$ fixed real numbers. By appropriate scaling we can obtain $a^2 + b^2 = 1$, which yields the description of the line as the image of an isometric function $M := \begin{bmatrix} a \\ b \end{bmatrix}$. The orthogonal projection of a point $w \in \mathsf{R}^2$ onto $\ell$, depicted in Fig. 4.1 is given by $\hat{w} = MM^\top w$.

**Figure 4.1**: Orthogonal projection.




The figure clearly shows that this is coincides with the optimal approximation of $w$ by a point on the line. Further notice that, independent of the value of $w$, the approximation error $\tilde{w} = w - \hat{w}$ lays on the line orthogonal to $\ell$, which is the image of $\tilde{M} = \begin{bmatrix} -b \\ a \end{bmatrix}$. ◇

For the orthogonal projection onto systems we take our starting point in an observed time series

$$w : T \to \mathsf{R}^q \text{ with } T = [1, N]. \tag{4.2}$$

For a given system $\mathcal{B}$ we have to determine the element in $\mathcal{B}_T$ that is the closest to $w$. Notice that $\mathcal{B}_T$ is a linear space in $(\mathsf{R}^q)^N$, which can be identified with $\mathsf{R}^{qN}$ in the obvious way. So in order to determine the optimal approximation, within a given system $\mathcal{B}$ it suffices to determine a linear isometric operator $G$ with image equal to $\mathcal{B}_T$. For this purpose we use the parametrization of $\mathcal{B}_T$ in terms of initial states $\hat{x}_1$ and auxiliary inputs $\hat{v} : T \to \mathsf{R}^m$ in an SR as described in (3.18). This indeed corresponds to the representation of $\mathcal{B}_T$ as the image of a linear operator, but it is in general not isometric. As a first attempt to obtain the isometry property let us consider the operator induced by an ISR. We assume that $\mathcal{B}$ is stabilizable, which guarantees the existence of an ISR, cf. Proposition 3.2.4. So we consider the mapping

$$(\hat{x}_1, \hat{v}) \xrightarrow{\text{ISR}} \hat{w}$$

with $\hat{w}$ defined as in (3.18) for an ISR of $\mathcal{B}$ . It follows from Proposition 3.2.2 that then

$$|\hat{x}_1|^2 + \|\hat{v}\|^2 = \|\hat{w}\|^2 + |\hat{x}_{N+1}|^2, \tag{4.3}$$

with $\hat{x}_{N+1}$ the end state corresponding to $\hat{w}$. Notice that if the last term would be absent, this mapping would be an isometry, for the obvious choice of norms. The equation also shows that ISR's induce an isometry from the pair $(\hat{x}_1, \hat{v})$ to the pair $(\hat{w}, \hat{x}_{N+1})$, for the obvious choice of norms. Therefore we make a slight adaptation by also taking into account the end state at time $N + 1$ for trajectories $\hat{w}$, so we define the operator $G$ by

$$G : (\mathsf{R}^n \times \mathsf{R}^{m\times N}) \to (\mathsf{R}^{q\times N} \times \mathsf{R}^n),\ G(\hat{x}_1, \hat{v}) := (\hat{w}, \hat{x}_{N+1}), \tag{4.4}$$

with $\hat{w}, \hat{x}_{N+1}$ defined by $\hat{x}(t + 1) = A\hat{x}(t) + B\hat{v}(t);\ \hat{w}(t) = C\hat{x}(t) + D\hat{v}(t)$, $t \in [1, N]$ for an ISR $(A, B, C, D)$ of $\mathcal{B}$ .
This is an isometric operator, with image equal to $\mathcal{B}_T$, albeit that it also attaches the corresponding end state to trajectories. It turns out that we can use the result in Lemma 4.1.1 for $G$ with only a slight modification, as follows.

**Proposition 4.1.2 (Orthogonal Projection onto Systems)**
*Let $w : T \to \mathsf{R}^q$ denote an observed time series,and $\mathcal{B}$ a given stabilizable system. Define*

$$(\hat{w}, \hat{x}_{N+1}) := GG^*(w, x_{N+1}) \tag{4.5}$$

*with $G$ the operator induced by an ISR of $\mathcal{B}$ as defined in (4.4), and $x_{N+1} \in \mathsf{R}^n$ chosen such that $\|w - \hat{w}\|$ is minimal. Then $\hat{w}$ is the orthogonal projection of $w$ onto $\mathcal{B}_T$.*

*Proof.* See Appendix A.4.

Summarizing, the optimal approximation in a given system equals the orthogonal projection onto that system, which can be determined a backward and

forward recursion in terms of an ISR, and an optimization problem over $\mathsf{R}^n$. In order to streamline the implementation of this procedure we first derive some additional results concerning the orthogonal complement of systems in the next section.

**Leading Example - Continued.** We illustrate Proposition 4.1.2 by considering the orthogonal projection of the time series

$$w_{ex} = \begin{pmatrix} u_{ex} \\ y_{ex} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 1 & 0 \end{bmatrix} \tag{4.6}$$

onto the system $\mathcal{B}_{ex}$, defined by

$$\hat{y}(t) = 2/3\,\hat{y}(t-1) + 2\hat{u}(t) - 2\hat{u}(t-1)\}, \tag{4.7}$$

cf. (3.2). First we construct $G$ as defined in (4.4). Let $(A, B, C, D)$ denote the ISR for $\mathcal{B}_{ex}$ computed in (3.17), then $G$ can be specified in matrix notation by

$$G : \begin{pmatrix} \hat{x}_1 \\ \hat{v}_1 \\ \hat{v}_2 \\ \hat{v}_3 \end{pmatrix} \to \begin{bmatrix} C & D & 0 & 0 \\ CA & CB & D & 0 \\ CA^2 & CAB & CB & D \\ A^3 & A^2B & AB & B \end{bmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{v}_1 \\ \hat{v}_2 \\ \hat{v}_3 \end{pmatrix} =: \begin{pmatrix} \hat{w}_1 \\ \hat{w}_2 \\ \hat{w}_3 \\ \hat{x}_4 \end{pmatrix} \tag{4.8}$$

which maps the initial state and auxiliary inputs to the system trajectory and end state corresponding to the ISR. From (4.3) it follows that this mapping is isometric. The adjoint $G^*$ is given by the transpose of this matrix, for which we give a system theoretic interpretation in Section 4.3. The evaluation of equation (4.5) amounts to a simple linear optimization problem over $\hat{x}_4 \in \mathsf{R}$, which yields $\hat{x}_4 = 0.55$ and

$$\hat{w}_{ex} = \begin{bmatrix} 1.10 & 0.67 & 0.24 \\ 2.97 & 1.13 & -0.12 \end{bmatrix} \tag{4.9}$$

This is the closest time series to $w$ that satisfies the system law (4.7) of $\mathcal{B}_{ex}$, so the misfit of $\mathcal{B}_{ex}$ with respect to the data equals $\|w_{ex} - \hat{w}_{ex}\| = 0.45$. Stated otherwise, the data is 0.45 away from what is in exact correspondence with the system laws.

In order to illustrate the global character of the GTLS misfit, we compare it with the first step ahead prediction error induced by $\mathcal{B}_{ex}$. These are simply evaluated by substituting the data in the right hand side of (4.7), *separately* for $t = 2$ and $t = 3$, and subtracting the prediction for $y_{ex}(t)$ from its actual value. This gives predictions 2 and $-4/3$ for respectively $y(2)$ and $y(3)$, so the prediction errors equal $-1$ and $4/3$ for $t = 2, 3$. Replacing $y(2)$ and $y(3)$ by these predictions does not result in a time series that satisfies (4.7).

We remark that in principle this procedure for determining the GTLS misfit is also applicable for long time series, but then this involves very large matrices. Therefore we will develop in Section 4.3 an algorithm in terms of recursive formulas. $\diamond$

## 4.2 The Orthogonal Complement of Systems

In the previous section we have seen that the optimal approximation of data within a system $\mathcal{B}$ equals the orthogonal projection of the data onto that system. This implies that the approximation error is contained in the orthogonal complement of the system. In this section we present some results on the representation of the orthogonal complement by ISR's. Although these results are not essential for the construction of the optimal approximation, they play a crucial role in the construction of GTLS models in the next chapter.
There is a close relationship between the ISR's of a system and its orthogonal complement. For a system $\mathcal{B}$ and a finite time interval $T$ the orthogonal complement of $\mathcal{B}_T$ is defined by

$$(\mathcal{B}_T)^\perp := \{\tilde{w} : T \to \mathsf{R}^q;\ \langle w, \tilde{w}\rangle = 0 \text{ for all } w \in \mathcal{B}_T\}, \tag{4.10}$$

where the inner product is defined by $\langle w, \tilde{w}\rangle := \Sigma_{t\in T} w(t)^\top \tilde{w}(t)$. Let $\mathcal{B}_T^0$ denote the time series in $\mathcal{B}_T$ that can be preceded and continued by zeros in the system $\mathcal{B}$ . The orthogonal complement of systems can be described in terms of ISR's as follows.

**Proposition 4.2.1 (Orthogonal Complement)**
*Let $(A, B, C, D)$ denote an ISR of a system $\mathcal{B}$ , and let $\tilde{B}$ and $\tilde{D}$ such that $\begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix}$ is a unitary matrix. Then $\mathcal{B}_T(A, B, C, D)^\perp$ is given by*

$$\begin{aligned}&\mathcal{B}_T^0(A, \tilde{B}, C, \tilde{D}) = \\ &\quad \{\tilde{w} : T \to \mathsf{R}^q;\ \{\ldots, 0, 0, \tilde{w}, 0, 0, \ldots\} \in \mathcal{B}(A, \tilde{B}, C, \tilde{D})\}.\end{aligned} \tag{4.11}$$

*Proof.* See Appendix A.4.

In order to get some intuition we suggest first to consider the static case, i.e., with $A, B, C$ empty matrices and $[D\ \tilde{D}]$ a unitary matrix.
Notice that $(A, \tilde{B}, C, \tilde{D})$ is also isometric, so we may call this an ISR of the orthogonal complement of $\mathcal{B}$ . To give this a formal meaning, we define the orthogonal complement $\mathcal{B}^\perp$ of $\mathcal{B}$ on the infinite time axis equal to the time series in $\mathcal{B}(A, \tilde{B}, C, \tilde{D})$ with finite support, i.e., which are zero outside a finite interval.
The fact that time series in $(\mathcal{B}_T)^\perp$ can be extended by zeros in $\mathcal{B}(A, \tilde{B}, C, \tilde{D})$ can be interpreted as follows. If the error would not have zero boundary states in this system, this would imply that for all extensions of the data the misfit with respect to $\mathcal{B}$ would increase. Clearly this is not true for an extension that consists of a forward or backward continuation within $\mathcal{B}$ of the optimal approximation of a time series in $\mathcal{B}_T$. Such an extension of the data corresponds to an two-sided extension of the approximation error by zeros.

**Remark.** The proposition can be used to derive the following characterization of minimal ISR's, which is an alternative to Proposition 3.2.7. The controllability condition for a pair of matrices is spelled out in the proof.

**Proposition 4.2.2 (Minimality of ISR)** *An isometric state representation is $(A, B, C, D)$ is minimal if and only if*

1. *$[A\ B]$ has full row rank*
2. *$(A, \tilde{B})$ controllable, with $\tilde{B}$ as in Proposition 4.2.1.*

*Proof.* See Appendix A.4.

The first condition has already been discussed in Proposition 3.1.5. The second condition can be interpreted as follows. Only the system trajectories in $\mathcal{B}(A, \tilde{B}, C, \tilde{D})$ with finite support play a role in (4.11). Uncontrollable (autonomous) modes correspond to exponentials, which have infinite support, so they are superfluous in describing the orthogonal complement. ◇

**Leading Example - Continued.** We use Proposition 4.2.1 to determine an ISR for the orthogonal complement of $\mathcal{B}_{ex}$ on the basis of the ISR (3.17). This gives $(6/7, 3/7, \begin{bmatrix} 2/7 \\ -3/7 \end{bmatrix}, \begin{bmatrix} -6/7 \\ 2/7 \end{bmatrix})$, which corresponds to the equation

$$y(t) = y(t-1) - 1/3\, u(t) + 1/2\, u(t-1). \tag{4.12}$$

So every time series with finite support that is orthogonal to $\mathcal{B}_{ex}$ satisfies this difference equation. This holds in particular for the approximation error $\tilde{w}_{ex} := w_{ex} - \hat{w}_{ex}$, with data $w_{ex}$ given by (4.6), and its optimal approximation in $\mathcal{B}_{ex}$ given by (4.9). This approximation error equals

$$\tilde{w}_{ex} = \begin{bmatrix} -0.096 & 0.331 & -0.235 \\ 0.032 & -0.126 & 0.118 \end{bmatrix}. \tag{4.13}$$

According to Proposition 4.2.1, this sequence, extended by zeros, satisfies (4.12). (Proposition 4.2.2 states that the ISR for the orthogonal complement is controllable.) ◇

## 4.3 Projection Algorithm

In this section we present an algorithm for the orthogonal projection of an observed time series onto a given system. The algorithm is an implementation of Proposition 4.1.2, and also exploits the characterization of the orthogonal complement of systems in terms of ISR's, as developed in the previous section. First we determine an expression for the adjoint of $G$.

**Lemma 4.3.1 (Adjoint)** *The adjoint of $G$ is given by $G^*(w, x_{N+1}) = (x_1, \hat{v})$ with $\hat{v}, x_1$ defined by the backward recursive formulas with $t \in \{N, \ldots, 1\}$,*

$$\begin{aligned} x(t) &= A^\top x(t+1) + C^\top w(t) \\ \hat{v}(t) &= B^\top x(t+1) + D^\top w(t), \end{aligned} \tag{4.14}$$

*Proof.* This is immediately clear from the representation of $G$ as a matrix, cf. (4.8), and the fact that the adjoint of a matrix is given by its transpose. ♣

Now the optimal approximation in a given system can be determined as follows. In order to increase the readability, the details of the simple linear optimization problem in the second step is described in the proof of correctness of the algorithm.

**Algorithm 1 (Optimal Approximation in Given System)**

**Data:**
- *An observation* $w : T \to \mathsf{R}^q$, $T = \{1, \ldots, N\}$
- *A stabilizable system* $\mathcal{B}$ *with minimal ISR* $(A, B, C, D)$.

**Step 1:** *Determine* $\tilde{B}$ *and* $\tilde{D}$ *such that* $\begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix}$ *is a unitary matrix.*

**Step 2:** *Define* $\hat{v}$, $\tilde{v}$ *and* $x$ *by the backward recursive equations*

$$\begin{aligned} x(t) &= A^\top x(t+1) + C^\top w(t) \\ \hat{v}(t) &= B^\top x(t+1) + D^\top w(t) \\ \tilde{v}(t) &= \tilde{B}^\top x(t+1) + \tilde{D}^\top w(t), \end{aligned} \tag{4.15}$$

*with the end state* $x(N+1)$ *determined such that* $\|\tilde{v}\|$ *is minimal.*

**Step 3:** *Define* $\hat{w}$ *by*

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + B\hat{v}(t) \text{ with } \hat{x}(1) = x(1) \\ \hat{w}(t) &= C\hat{x}(t) + D\hat{v}(t), \end{aligned} \tag{4.16}$$

*and define* $\tilde{w}$ *by*

$$\begin{aligned} \tilde{x}(t+1) &= A\tilde{x}(t) + \tilde{B}\tilde{v}(t) \text{ with } \tilde{x}(1) = 0 \\ \tilde{w}(t) &= C\tilde{x}(t) + \tilde{D}\tilde{v}(t), \end{aligned} \tag{4.17}$$

**Result:**
- $\hat{w}$ *is the optimal approximation of* $w$ *in* $\mathcal{B}$ .
- $\tilde{w} = w - \hat{w}$ *is the corresponding approximation error.*
- $\mathrm{d}(w, \mathcal{B}) = \|\tilde{w}\| = \|\tilde{v}\|$.

*Proof.* See Appendix A.4.
*Implementation:* `Alg1` in Appendix B.2.

The relation with Proposition 4.1.2 is as follows. Let $G$ be the operator induced by $(A, B, C, D)$, as described in (4.4), and define $\tilde{G}$ similarly as the operator induced by $(A, \tilde{B}, C, \tilde{D})$. The second step is the combined evaluation of $G^*(w, x_{N+1})$ and $\tilde{G}^*(w, x_{N+1})$ denoted respectively by $(x_1, \hat{v})$ and $(x_1, \hat{v})$. So

in the third step we determine $\hat{w} = GG^*(w, x_{N+1})$ (and $\tilde{w} = \tilde{G}\tilde{G}^*(w, x_{N+1})$). This is in correspondence with Proposition 4.1.2 as $x_{N+1}$ is determined such that $\|w - \hat{w}\| = \|\tilde{v}\|$ is minimal.
We remark that for the mere evaluation of the misfit the third step is superfluous, as it is already determined in the second step, by the norm of $\tilde{v}$ .
This algorithm is a modified version of the projection algorithm in [47], which concerns the optimal approximation of infinite square summable time series. New aspects of Algorithm 1 are the estimation of boundary states, the use of ISR's and the results on the approximation error.

**Leading Example - Continued.** In order to illustrate the algorithm we consider the optimal approximation in the system $\mathcal{B}_{ex}$, defined by (4.7) for a system trajectory in $\mathcal{B}_{ex}$ that is corrupted by white noise. The regular part $w_r$ consists of two components, $u_r$ and $y_r$, where $u_r$ is the realization of a white noise process with unit variance, and $y_r$ is satisfies the system law of $\mathcal{B}_{ex}$, so

$$y_r(t) = 2/3\, y_r(t-1) + 2u_r(t) - 2u_r(t-1), \tag{4.18}$$

for $t \in [1, 100]$. The observation $w$ consists of two components $u$ and $y$, with

$$u(t) = u_r(t) + \eta(t);\ y(t) = y_r(t) + \varepsilon(t), \tag{4.19}$$

where $\eta$ and $\varepsilon$ are independent white noise processes with variance 0.25, both independent of $u_r$.
The misfit of $\mathcal{B}_{ex}$ equals $\|\tilde{w}\| = 5.20$, as compared to $\|w\| = 22.72$, so the *relative misfit* is given by $5.20/22.72 = 0.23$. This is considerably smaller than the Euclidian norm of the white noise by which the observation was corrupted (recall that $w = w_r + w_n$, with $w_r \in \mathcal{B}_{ex}$ and $w_n$ white noise; in our example $\|w_n\| = 7.62$). This can be explained as follows. The optimal approximation is simply obtained by projecting the noise $w_n$ on $\mathcal{B}_{ex}$, with a resulting decomposition $w_n = \hat{w}_n + \tilde{w}_n$, where $\hat{w}_n \in \mathcal{B}_{ex}$ and $\tilde{w}_n \in \mathcal{B}_{ex}^{\perp}$. In our case, $\|\hat{w}_n\| = 5.56$ and $\|\tilde{w}_n\| = \|\tilde{w}\| = 5.22$. So the noise is distributed nearly equally over the system $(\mathcal{B}_{ex})_{[1,100]}$ and its orthogonal complement, which corresponds to the fact that they are nearly equally large linear subspaces of $\mathsf{R}^{2\times 100}$, of dimension respectively 101 and 99. $\diamond$

## 4.4 Recursive evaluation of misfit

The algorithm in Section 4.3 does not only determine the misfit of a system with respect to a time series, but also determines the approximation error (or 'residuals') at each time instant. The approximation error indicates how the misfit is distributed over the observation interval. If it consist of some high peaks, this might point exceptional deviations at the corresponding time instants, while a 'white noise like' error indicates that the misfit is equally distributed over time. Notice, however, that a high peak in the approximation

error at time $t$, $\tilde{w}(t)$, is not necessarily due to a large deviation in the data at exactly that moment, but may also be the effect of observations in the neighbourhood of $t$.
In order to gain additional insight in the location of deviations, we analyse the effect of each individual observation on the misfit *recursively*, i.e. given the observations in the past, but independent of observations in the future. Let $\hat{w}_t$ denote the optimal approximation, within a given system $\mathcal{B}$ , of $w_{[1,t]}$, so

$$\hat{w}_t : [1,t] \to \mathsf{R}^q \in \mathcal{B}_{[1,t]} \text{ minimizes } \|w_{[1,t]} - \hat{w}\| \tag{4.20}$$

Then the problem of determining these optimal approximations $\hat{w}_t$ recursively in time can be formulated as follows.

**Definition 4.4.1 (Recursive Evaluation of Misfit)**
*Given:*

- *an observation* $w : T \to \mathsf{R}^q$*, with* $T = [1, N] \subset \mathsf{Z}$
- *tolerated size* $(m, n)$*.*

*determine:*

- $\hat{w}_t(t)$ *as a function of the previous approximations* $\hat{w}_{t-1}$ *and the new observation point* $w(t)$*.*

We make use of the following notation. Let $\mathcal{B}$ denotes a given system with ISR $(A, B, C, D)$, and let $\tilde{B}, \tilde{D}$ be defined as in Algorithm 1. The observation interval is given by $[1, N]$ . Further

| | |
|---|---|
| $\hat{w}_t$: | the optimal approximation within $\mathcal{B}$ of the observation up to and including time $t$, cf. (4.20). |
| $\tilde{w}_t$: | the approximation error corresponding to $\hat{w}_t$. |
| $\hat{x}_t$, $\hat{v}_t$: | the state and auxiliary input corresponding to $\hat{w}_t$ in $(A, B, C, D)$. |
| $\tilde{x}_t$, $\tilde{v}_t$ | the state and auxiliary input corresponding to $\tilde{w}_t$ in $(A, \tilde{B}, C, \tilde{D})$. |
| $x_t$: | $\hat{x}_t + \tilde{x}_t$. |

Notice that $\hat{w}_t$, $\tilde{w}_t$, $\hat{v}_t$ and $\tilde{v}_t$ are time series defined on $\{1, \ldots, t\}$, while $\hat{x}_t$, $\tilde{x}_t$ and $x_t$ are also defined for time $t + 1$.
In this notation, the optimal approximation determined in Algorithm 1 is denoted by $\hat{w}_N$ and its corresponding state by $\hat{x}_N$. In the recursive algorithm we determine the values for $\hat{x}_t(t + 1)$ and $\hat{w}_t(t)$ recursively for $t = 1, \ldots, N$. Comparing $\hat{w}_t(t')$ and $\hat{w}_N(t')$, both are approximations of $w(t')$, where the first one is optimal given the observations up to time $t$, while the latter is based on the whole observation. Stated according to the common terminology in the literature, Algorithm 1 determines the smoothed values of the optimal approximation and the corresponding state, while the recursive algorithm computes their past-induced values.
The algorithm is based on the following considerations. Suppose we have observed a time series $\{w(1), \ldots, w(t - 1)\}$ and determined its optimal approximation in $\mathcal{B}$ ,

$$w(t') = \hat{w}_{t-1}(t') + \tilde{w}_{t-1}(t'), \text{ for } t' \in \{1, \ldots, t - 1\},$$

where $\hat{w}_{t-1} \in \mathcal{B}_{[1,t-1]}$ with final state $\hat{x}_{t-1}(t)$. First consider the case that the next observation is a continuation of $\hat{w}_{t-1}$ within $\mathcal{B}$ , i.e., $w(t) = C\hat{x}_{t-1}(t) + D\hat{v}(t)$ for some $\hat{v}(t) \in \mathsf{R}^m$. Obviously, this will not increase the misfit, as there is no reason to change the approximation before $t$, and no approximation error has to be made at $t$.
Next suppose that $w(t)$ is not compatible with $\hat{w}_{t-1}$ in $\mathcal{B}$ , i.e., $\epsilon(t) := w(t) - C\hat{x}_{t-1}(t) \notin \mathrm{im}D$. Then we have to approximate the observation by

$$\hat{w}_t(t) = C\hat{x}_t(t) + D\hat{v}_t(t), \tag{4.21}$$

where $\hat{x}_t(t)$ and $\hat{v}_t(t)$ have to be determined such that the increase of the squared misfit, denoted by $m(t)^2$, is as small as possible. This increase consists of two parts:

- the approximation error at time $t$, i.e., $|w(t) - \hat{w}_t(t)|$, denoted by $m_0^2(t)$
- an increase of misfit over the past, denoted by $m_-^2(t)$, due to changing the state $\hat{x}_{t-1}(t)$, which is optimal for $\{w(1), \ldots, w(t-1)\}$, into $\hat{x}_t(t)$.

In general, keeping the past approximation $\hat{w}_{t-1}$ fixed, hence keeping $\hat{x}_{t-1}(t)$ fixed, would lead to a large approximation error at time $t$, while minimizing the error at $t$ alone would lead to a large increase of the misfit over the past due to the large change in the state at time $t$. The optimal strategy is a trade off between both approaches. The main result underlying the algorithm is that the only feature of the past observations that determines this trade off is the final state $\hat{x}_{t-1}(t)$ of the optimal approximation. Equivalently, given $\hat{x}_{t-1}(t)$, the optimal values for $\hat{x}_t(t)$ and $\hat{v}_t(t)$ in (4.21) are independent of the past observations. This opens the way to compute the optimal values for $\hat{x}_t(t+1)$ and $\hat{v}_t(t)$ in (4.21) recursively.

**Algorithm 2 (Recursive Approximation within a Given System)**

**Data:**
- *An observation* $w : T \to \mathsf{R}^q$, $T = \{1, \ldots, N\}$
- *A stabilizable system* $\mathcal{B}$ *with minimal ISR* $(A, B, C, D)$.

**Step 1:** *Determine* $\tilde{B}$ *and* $\tilde{D}$ *such that* $\begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix}$ *is a unitary matrix.*

**Step 2:** *Define* $\tilde{W}_t$, $\tilde{F}_t$, $\tilde{G}_t$ *and* $\tilde{H}_t$ *for* $t \in T$ *by the equations*

$$\begin{aligned} \tilde{W}_{t+1} &= A\tilde{W}_t A^\top + \tilde{B}\tilde{B}^\top;\ \tilde{W}_1 = 0 \\ \tilde{F}_t &= A\tilde{W}_t C^\top + \tilde{B}\tilde{D}^\top, \\ \tilde{G}_t &= C\tilde{W}_t C^\top + \tilde{D}\tilde{D}^\top \\ \tilde{H}_t &= \tilde{G}_t - \tilde{F}_t^\top \tilde{W}_{t+1}^{-1} \tilde{F}_t \end{aligned}$$

*If the inverse does not exists, replace by a pseudo-inverse.*

**Step 3:** *Define for $t \in T$,*

$$\begin{aligned} \epsilon(t) &= w(t) - C\hat{x}_{t-1}(t) && (4.22)\\ \hat{x}_t(t+1) &= A\hat{x}_{t-1}(t) - \tilde{W}_{t+1}^{-1}\tilde{F}_t\epsilon(t); \hat{x}_0(1) = 0 && (4.23)\\ \hat{w}_t(t) &= C\hat{x}_t(t) + (I - \tilde{H}_t)\epsilon(t). && (4.24) \end{aligned}$$

*and further*

$$\begin{aligned} m(t)^2 &= \epsilon(t)^\top \tilde{H}_t \epsilon(t); && (4.25)\\ m_0(t)^2 &= \epsilon(t)^\top \tilde{H}_t^2 \epsilon(t) && (4.26)\\ m_-(t)^2 &= \epsilon(t)^\top (\tilde{H}_t - \tilde{H}_t^2)\epsilon(t) && (4.27) \end{aligned}$$

**Result:**

- *$\hat{x}_{t+1}(t)$ is the end state of the optimal approximation of $\{w(1), \ldots, w(t)\}$ in $\mathcal{B}$.*
- *$\hat{w}_t(t)$ is the optimal approximation of $\{w(1), \ldots, w(t)\}$ at time $t$.*
- *$m(t)^2$ is the increase of squared misfit due to observation $w(t)$.*
- *$m(t)^2 = m_-(t)^2 + m_0(t)^2$, where $m_-^2(t)$ denotes the increase of the squared misfit due to observation $w(t)$ on the past $\{1, \ldots, t-1\}$, and $m_0(t)$ the misfit at $t$.*

*Proof.* See Appendix A.4.
*Implementation:* `Alg2` in Appendix B.2.

Concerning the interpretation of this algorithm, we make the following remarks.

- The matrix $\tilde{W}_t$ is the finite time controllability gramian corresponding to $(A, \tilde{B})$. It measures the cost of changing the optimal end state: a change into $\hat{x}_{t-1}(t) + \bar{x}(t)$ requires an increase of the squared norm of the auxiliary variables $\tilde{v}_t$ by $\bar{x}(t)^\top \tilde{W}_t \bar{x}(t)$, as shown in the proof.

- The algorithm is comparable to Kalman filtering in the sense that it concerns the recursive reconstruction of the (unobserved) state from an observed time series. For a description of the Kalman Filter we refer to [2]. In the sequel we assume some familiarity with this subject. The main difference is that in Kalman filtering the estimation of the state is based on (a priori) stochastic assumptions on the data generating process for $w$, while the 'GTLS-filter' is entirely based on the (also a priori) chosen deterministic misfit criterion. In addition, GTLS models contain some degrees of freedom at each time instant, while Kalman Filtering takes its starting point in a *complete* (stochastic) specification of a process.

- The matrix $\tilde{W}_{t+1}^{-1}\tilde{F}_t$ in the third step plays the role of the so-called Kalman gain, indicating the weights by which the deviations $\epsilon(t)$ influence the state evolution.

- One may compare $\epsilon(t)$ to an innovation process, representing the 'unexpected' part of a new observation $w(t)$ given the final state $\hat{x}_{t-1}(t)$ of the approximation $\hat{w}_{t-1}$. A zero innovation $\epsilon(t)$ corresponds to an observation $w(t)$ that is in perfect correspondence with the implications of the system $\mathcal{B}$ , and hence does not lead to an increase of misfit, cf. (4.25).

- The effect of a nonzero innovation at time $t$ on the misfit is measured by the matrix $\tilde{H}_t$. Equation (4.25) states that the increase of squared misfit at time $t$, $m(t)^2$, equals $\epsilon(t)^\top \tilde{H}_t \epsilon(t)$. For a good understanding of the algorithm it is important to notice that $\tilde{H}_t$ is singular: $\tilde{H}_t D = 0$. This reflects the fact that values of $\epsilon(t)$ in the image of $D$ do not cause an increase of misfit. Such an 'innovation' corresponds to an observation of the form $w(t) = C\hat{x}_{t-1}(t) + Dz$, with $z \in \mathsf{R}^m$, which is a continuation of the approximation $w_{t-1}$ in $\mathcal{B}$, and hence causes no error at time $t$. Stated otherwise, the kernel of $\tilde{H}_t$ represents the degrees of freedom in the system $\mathcal{B}$ at time $t$.

  An observation $w(t)$ for which $\epsilon(t)$ is not contained in the kernel of $\tilde{H}_t$ leads to an increase of misfit. The equations (4.26) and (4.27) show how the approximation error is divided over the past (due to a revision of the approximation before $t$) and present.

- The system matrices involved in Step 3 are time dependent, but independent of the observation. For $N \to \infty$, $\tilde{W}_t$ converges to the controllability gramian $\tilde{W}$ of $(A, \tilde{B})$, which is well defined as $A$ is asymptotically stable, cf. Proposition 3.2.7.1. Consequently, if $\tilde{W}_t$ has become equal to $\tilde{W}$ within a sufficiently high precision level, it can be replaced by $\tilde{W}$, resulting in a time-invariant filter for the optimal approximation in Step 3. This is comparable to the so-called steady state Kalman filter.

We conclude this chapter by some reflections on the following question: *What does a model $\mathcal{B}$ imply for an observation $w(t)$ given the past observations $\{w(1), \ldots, w(t-1)\}$?* Clearly the answer depends on the relationship between model and data we assume. If $\mathcal{B}$ is known to be an *exact* model for $w$, the answer is obvious: $w(t)$ belongs to the set

$$\{w(t);\ w(t) = C\hat{x}_{t-1}(t) + Dz \text{ for some } z \in \mathsf{R}^m\}, \tag{4.28}$$

where $\hat{x}_{t-1}(t)$ is the optimal end state constructed on basis of the previous observations. Equivalently, the 'innovation' $\epsilon(t)$ is then contained in the kernel of $\tilde{H}_t$, which equals the image of $D$ for $t > n$, as explained before.
On the other hand, if no assumptions are made on the quality of $\mathcal{B}$ for $w$, of course *nothing* can be said about the value $w(t)$. The results of the algorithm are still valid in this case, as it is obtained without any assumption on the relation between system and data, but nothing can be said about $\epsilon(t)$ and hence nothing about $w(t)$.
The most interesting question from a practical point of view concerns situations in between these extremes, in which we know that the system $\mathcal{B}$ is an *accurate*

model for $w$, i.e. with a low misfit. On basis of the foregoing results we can determine the set of all values for $w(t)$ that lead to a certain increase of squared misfit $\mu$. From (4.25) it follows that this set is given by

$$\{w(t);\ |\tilde{H}_t^{1/2}(w(t) - C\hat{x}_{t-1}(t)| = \sqrt{\mu}\}.$$

So the level curves of the GTLS criterion consist of elliptic 'cones' around the 'exact' data points (4.28). This gives a rather detailed answer to the question for approximate models. We refer to Section 6.4.1 for an illustration of this fact.

# Chapter 5

# GTLS Models

GTLS models are systems of restricted size that have minimal misfit with respect to an observed time series. In the previous chapter we described an algorithm for the evaluation of the misfit for a given system, in terms of isometric state representations. Now we come to the central problem in the GTLS approach, namely to determine a system of restricted rank and degree that has minimal misfit with respect to the data.

This is a relatively complex problem as compared to the evaluation of the misfit, as it concerns optimization over a space that is non-linear, even non-convex. For the static case this is solved by the singular value decomposition (SVD), for the dynamic case we have to resort to iterative procedures that determine locally optimal models.

Another complication is the fact that neither the existence nor the uniqueness of solutions is guaranteed. Therefore, most results in this chapter are valid for *generic* data, i.e. with the exclusion of a set that has zero measure.

We will consider the GTLS problem at three levels:

| | |
|---|---|
| GTLS: | at the level of trajectories, as in Definition 2.1.8, |
| GTLS-SR: | as an optimization problem in terms of SR's, |
| GTLS-ISR: | as an optimization problem in terms of ISR's. |

We start with an analysis of the GTLS problem at the first level, by discussing the existence and uniqueness of solutions, and the occurrence of local optima. This is followed by two algorithms for the construction of locally optimal models, based on the formulation of the GTLS problem in terms of respectively state representations (GTLS-SR) and isometric state representations (GTLS-ISR).

The first algorithm consists of three model improvement constructions to lower the misfit. From these constructions we derive a characterization of the stationary points of the GTLS criterion in terms of orthogonality conditions. This is used to determine the *optimality margin* of an approximation of the data, which indicates how much the data has to be changed to make the approximation a stationary point. The algorithm also shows how to exploit the SVD for the dynamic case.

The second algorithm is a Gauss-Newton algorithm in terms of ISR's.

The result of the iterative algorithms depends on the chosen initial model, as there may be several local optima. One way to account for this is to start the algorithms in several models and to compare the results, in the hope to find the global optimum. In fact this seems to work quite well in applications that concern small systems. A more sophisticated approach is to use a heuristic method to determine 'good' initial values that may be expected to be close to the global optimum. In Section 5.6 we present such a procedure, based on canonical correlation analysis of the data.

## 5.1 On Existence and Uniqueness

In this section we discuss the existence and uniqueness of solutions for the GTLS problem. For the analysis of these issues we take our starting point in the formulation of the GTLS problem as the approximation of the data by regular time series,

$$\text{GTLS} : \text{minimize } \|w - \hat{w}\| \text{ over } \hat{w} \in \mathsf{B}_T^{q,m,n}, \tag{5.1}$$

as formulated in Definition 2.1.8. As before, $T$ denotes the observation interval, $q$ the number of components in the data, and $m$ and $n$ denote the tolerated rank and degree. In the sequel we exclude the trivial case $m = q$.
First we discuss some properties of the set $\mathcal{B}_T^{q,m,n}$ that indicate the difficulty of the GTLS problem. This set is not linear, and even not convex. In general the sum of two of its elements belong to a system with rank $2m$ and degree $2n$, or it is even not regular if $2m \geq q$. For instance, the sum of two singular matrices in $\mathsf{R}^{2\times N}$ is in general not singular. Moreover $\mathcal{B}_T^{q,m,n}$ is not a closed set (in the 'standard' topology of pointwise convergence), and related to this, the optimal approximation of a data point in $\mathcal{B}_T^{q,m,n}$ need not exist. We illustrate this by a simple example.

**Example.** (No solution for GTLS.)
In order to illustrate the exceptional cases, we first consider the simple case with tolerated size at most $(m, n) = (0, 1)$, i.e., autonomous models of degree at most one, and the following data sequence

$$w \;=\; [0, 0, 1]. \tag{5.2}$$

The optimal approximation of this data sequence does not exist, which can be seen as follows. Clearly $w$ can not be generated by an autonomous model of rank one, so a zero misfit cannot be achieved. On the other hand, arbitrarily low misfits can be achieved for the system $\{w : \mathsf{Z} \to \mathsf{R};\ w(t) = c\alpha^t \text{ with } c \in \mathsf{R}\}$, by taking $\alpha$ sufficiently large. This shows that there does not exist an optimal autonomous system of degree one for $w$.
Similarly, there does not exist an optimal model for

$$w \;=\; [1, 0, 0] \tag{5.3}$$

of this size. Arbitrarily low misfits are obtained by taking $\alpha$ sufficiently close to zero, but $w$ is not a part of an exponential time series on $\mathsf{Z}$.
These examples are easily extended to the non-autonomous case. For instance, from the same type of reasoning it follows that there does not exist an optimal system of rank one and degree one for the vector time series

$$w = \begin{bmatrix} 0 & 0 & 1 \\ a & b & c \end{bmatrix}$$

if $a \neq 0$. ◇

Not only existence of solutions, also uniqueness is not guaranteed.

**Example.** (Non-uniqueness of solution)
As in the previous example we consider tolerated rank zero and degree one, now for the data $w = [0, 1, 0]$. Approximations take the form $c[1\ \lambda\ \lambda^2]$ for $c, \lambda \in R$. It is easily verified that for given $\lambda$ the optimal value for $c$ is given by $\frac{\lambda}{1+\lambda^2+\lambda^4}$, and that the corresponding GTLS criterion equals $1 - \frac{\lambda^2}{1+\lambda^2+\lambda^4}$. This achieves its optima in $\lambda = \pm 1$ and $c = \pm 1/3$, so both $\hat{w}_1 = [1/3, 1/3, 1/3]$ and $\hat{w}_2 = [-1/3, 1/3, -1/3]$ are globally optimal approximations in $\mathcal{B}^{1,0,1}$. ◇

Summarizing, the GTLS problem amounts to an optimization over a non-convex set, and neither the existence nor the uniqueness of solutions is guaranteed.

However, things are not as bad as this might suggest. The previous examples are degenerate cases, and for generic data the situation seems much simpler. In the following conjecture we list some statements that fall in the class 'obvious but not obvious to prove'.

**Conjecture 5.1.1 (Existence and Uniqueness of Solutions)**

1. *For generic data there exists a unique solution of the GTLS problem (5.1).*

2. *For generic data there is a unique stabilizable system of full tolerated size that contains the optimal approximation.*

3. *The solution is continuous in generic data.*

For a motivation we refer to Appendix A.5.

On the basis of this conjecture we speak in the sequel of *the* GTLS approximation of the data or *the* GTLS system (of a certain tolerated size), without claiming that there could not happen to be another global optimum. Both an optimal trajectory and the corresponding system are called the solution of the GTLS problem.
The second conjecture implies that the GTLS problem can be treated in terms of (minimal) ISR's, which only exist for stabilizable systems. Continuity of

the solution indicates that small changes in the data, such as round-off errors, do not have a dramatic effect.

The non-linear character of the optimization problem causes serious problems from a practical point of view. As mentioned before, we are not able to develop an algorithm that is guaranteed to determine the globally optimal solution of the GTLS problem. Instead, we have to set out to the less ambitious goal of determining locally optimal solutions, i.e., approximations that are the closest to the data at least in a small neighbourhood. This can be formalized as follows. We call a set $V \subset \mathcal{B}_T^{q,m,n}$ *open in* $\mathcal{B}_T^{q,m,n}$ if there is an open set $R$ in $\mathsf{R}^{q\times N}$ such that $V = \mathcal{B}_T^{q,m,n} \cap R$. It is easily verified that this induces a topology on $\mathcal{B}_T^{q,m,n}$. A neighbourhood of a data point is an open set that contains that point.

**Definition 5.1.2 (Local Optimality)** *A trajectory $\hat{w} \in \mathcal{B}_T^{q,m,n}$ is called locally optimal if it is the closest approximation to the data in a neighbourhood of $\hat{w}$. A system in $\mathsf{B}^{q,m,n}$ is called locally optimal if it contains a trajectory that is locally optimal (in $\mathcal{B}_T^{q,m,n}$).*

Clearly a global solution is also locally optimal, but in nonlinear problems the occurrence of several models that are only locally optimal is a rule rather than an exception. We conjecture that there exists finitely many local optima, and that upper bounds for their number can be estimated in terms of the length of the data and the tolerated size, but we have to leave this issue for further research.
Summarizing, although there are exceptional cases in which there exists no or more than one solution to the GTLS problem, we conjecture that this only concerns non-generic data. The algorithms we present only determine locally optimal models, which are in general not unique. This means that they do not necessarily yield the optimal solution, and in order to account for this fact, we accompany the iterative algorithms by a heuristic method to determine good initial values.

**Remark.** It is of some interest to know that from every initial point in $\mathcal{B}^{q,m,n}$ every other element can be reached by a continuous path, i.e., the set $\mathcal{B}_T^{q,m,n}$ is *pathwise connected*. A trivial way to see this is the connection of two points $z_1, z_2$ via zero by $\{\lambda z_1\} \cup \{\mu z_2\}$ with $\lambda, \mu \in [0,1] \subset \mathsf{R}$.
Of course it is not true that there exists a path from every point to the global optimum on which the misfit is decreasing, which means that the global optimum might be unreachable from some initial points along iterative improvements. So despite the fact that $\mathcal{B}_T^{q,m,n}$ is pathwise connected, it is still important to vary initial models in order to find the global optimum. $\diamond$

## 5.2 Model Improvement Constructions

The basic idea of our algorithm is to determine repeatedly local improvements of a given approximation of the data, until no direction can be found in which

the distance to the data decreases. This results in a sequence of models with decreasing misfit that, roughly speaking, converges to a local optimal solution. The first algorithm we present is based on the expression of the GTLS criterion as a parameter optimization problem in terms of SR's. We make use of the result in Proposition 3.3.2, which describes how SR's represent system behaviours on finite time intervals. Let $T = [1, N]$ denote the observation interval, and let $\hat{w}$ denote the outcome corresponding to the initial state $\hat{x}_1$ and auxiliary inputs $\hat{v} : T \to \mathsf{R}^m$ in the SR $(A, B, C, D)$, so

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + B\hat{v}(t) \\ \hat{w}(t) &= C\hat{x}(t) + D\hat{v}(t) \end{aligned} \tag{5.4}$$

for $t \in T$, cf. (3.18) with $\hat{x}(1) = \hat{x}_1$. According to this we define the function $G$ that maps an SR, an initial state and an auxiliary input sequence to the corresponding system trajectory $\hat{w} : T \to \mathsf{R}^q$, i.e.,

$$G(A, B, C, D, \hat{x}_1, \hat{v}) := \hat{w} \tag{5.5}$$

The GTLS problem (5.1) is (practically) equivalent to

$$\text{GTLS-SR} : \text{minimize } \|w - G(A, B, C, D, \hat{x}_1, \hat{v})\|. \tag{5.6}$$

**Remark.** In fact there is a slight difference with the GTLS problem as formulated in (5.1), related to the difference between SR's on finite time and on $\mathsf{Z}$ as discussed in Section 3.3.
As an illustration of the exception case, we consider the data $w$ as given in (5.3). This belongs to $\mathcal{B}_T(0, -, 1, -)$ for initial state $\hat{x}(1) = 1$, so (5.6) has a solution with zero misfit for $w$, while it has been proved that (5.1) has not, see Section 5.1.
This means that (5.6) may have some 'pseudo-solutions' that consist of system trajectories starting at $t = 1$ that cannot be extended to the past $\{t < 0\}$. Stated otherwise, some states of the system are not reachable from the past. As this plays no role in all our simulations, we ignore this detail in the sequel. ◇

We follow an iterative approach for the nonlinear problem (5.6). In each step we keep some parameters fixed, such that the resulting subproblem becomes sufficiently simple. For instance, for fixed $(A, B, C, D)$ the resulting problem in $\hat{v}$ and $\hat{x}(1)$ is solved by the projection scheme discussed in the foregoing chapter. We consider the following subproblems.

**Problem 1: Optimal $C$, $D$**
For given $A$, $B$, $\hat{v}$ and $\hat{x}(1)$, solve the GTLS problem for $C$ and $D$.

**Problem 2: Optimal $B$ and $D$ and $\hat{x}(1)$**
For given $A$, $C$ and $\hat{v}$, solve the GTLS problem for $B$, $D$ and $\hat{x}(1)$.

**Problem 3: Optimal $B$, $D$ and $\hat{v}$ in an ISR**
For given $A$ and $C$ with $A^\top A + C^\top C = I_n$ and given $\hat{x}(N+1)$, solve the GTLS problem for $B$, $D$ and $\hat{v}$ under the restriction that $(A, B, C, D)$ is isometric.

Notice that for the first two problems the approximation as defined in (5.6) is linear in the parameters. This means that the misfit is quadratic in these parameters, and hence these problems have a unique solution that is relatively easy to determine. The third problem is precisely the problem that can be solved by the SVD. In comparison with the static case, for which the SVD yields a complete solution, cf. Section 2.3, we can say that it also solves the GTLS problem for 'fixed memory structure', as $A$ and $C$ determine the effect of the state on system trajectories. In the next theorem we give constructive solutions of these problems. The variables $\hat{w}$, $\hat{x}$ and $\hat{v}$ are defined as in (5.4). In order to shorten the notation, we use $\hat{x}$ and $\sigma\hat{x}$ for the state on respectively time interval $[1, N]$ and $[2, N+1]$.

**Theorem 5.2.1 (Model Improvement Constructions)**

1. *Construction 1 (Projection of the Approximation Error)*
   *Let $\mathcal{E}$ denote the linear space*

$$\begin{aligned}\mathcal{E} := \quad \{\hat{w} : T \to \mathsf{R}^q;\ \exists C \in \mathsf{R}^{q\times n} \text{ and } D \in \mathsf{R}^{q\times m} \\ \text{such that with } \sigma\hat{x} = A\hat{x} + B\hat{v} \\ \text{it holds that } \hat{w} = C\hat{x} + D\hat{v}\}.\end{aligned} \tag{5.7}$$

   *Let $P$, $Q$ and $x_1$ denote the coefficients of the orthogonal projection of $w$ onto $\mathcal{E}$, say $\hat{w}' = P\hat{x} + Q\hat{v}$ with $\hat{x}(1) = \hat{x}_1$. Then $\mathcal{B}_1 := \mathcal{B}(A, B, P, Q)$ solves problem 1.*

2. *Construction 2 (Dual version of construction 1)*
   *Define the linear space $\mathcal{F}$ by*

$$\begin{aligned}\mathcal{F} := \quad \{\hat{w} : T \to \mathsf{R}^q;\ \exists B \in \mathsf{R}^{n\times m}, D \in \mathsf{R}^{q\times m} \text{ and } \hat{x}_1 \in \mathsf{R}^n \\ \text{such that with } \sigma\hat{x} = A\hat{x} + B\hat{v}, \hat{x}(1) = \hat{x}_1, \\ \text{it holds that } \hat{w} = C\hat{x} + D\hat{v}\}.\end{aligned} \tag{5.8}$$

   *Let $P, Q$ denote the coefficients of the orthogonal projection $\hat{w}'$ of $w$ onto $\mathcal{F}$, i.e., $\hat{w}' = C\hat{x} + Q\hat{v}$ and $\hat{x} = A\hat{x} + P\hat{v}$. Then $\mathcal{B}_2 := (A, P, C, Q)$ solves problem 2.*

3. *Construction 3 (SVD on auxiliary inputs)*
   *Let $\bar{B} \in \mathsf{R}^{n\times q}$, $\bar{D} \in \mathsf{R}^{q\times q}$ be such that $\begin{pmatrix} A & \bar{B} \\ C & \bar{D} \end{pmatrix}$ is a unitary matrix.*
   *Let $v : [1, N] \to \mathsf{R}^m$ be defined by*

$$\begin{aligned} x &= A^\top\sigma x + C^\top w;\ x(N+1) = \hat{x}(N+1) \\ v &= \bar{B}^\top\sigma x + \bar{D}^\top w, \end{aligned} \tag{5.9}$$

   *with SVD $v = \Sigma_{i=1}^q \lambda_i u_i z_i$, and let $U_m := [u_1, \ldots u_m] \in \mathsf{R}^{q\times m}$. Then $\mathcal{B}_3 := \mathcal{B}(A, \bar{B}U_m, C, \bar{D}U_m)$ solves problem 3.*

*Proof.* Parts *1* and *2* follow immediately from the definitions of $\mathcal{E}$ and $\mathcal{F}$. For part *3*, observe that $(A, B, C, D)$ is an ISR if and only if there exist $\tilde{B}$, $\tilde{D}$ and a unitary $V$ such that $[B\ \tilde{B}] = \bar{B}V$ and $[D\ \tilde{D}] = \bar{D}V$. For the representation $(A, BV, C, DV)$ Algorithm 1 gives as auxiliary input for the approximation and error $\hat{v}'$ and $\tilde{v}' \begin{pmatrix} \hat{v}' \\ \tilde{v}' \end{pmatrix} = V^\top v$, with misfit $\|\tilde{v}\|$. By taking $V = [u_1, \ldots, u_q]$, this misfit is determined by the $q - m$ smallest singular values of $v$, which is minimal. ♣

*Implementation:* `mic1`, `mic2`, `mic3` in Section B.2.

From this theorem it follows that the iterative application of these three constructions leads to a sequence of models with monotonically decreasing misfit. The question is whether these constructions are sufficient to determine locally optimal models. Stated otherwise, if an approximation is not locally optimal, does then at least one of the constructions yield an improvement? Roughly speaking, this is indeed the case. The proof of this result is based on a characterization of local optimality, which is the topic of the next section, where we also present a GTLS algorithm in terms of the constructions in Theorem 5.2.1.

**Leading Example - Continued.** In order to illustrate the foregoing, we consider the data $w = (u, y)$ described in Section 4.3, see (4.19). We apply the model improvement constructions to the model $\mathcal{B}_{ex}$ as defined in (3.2), and to a randomly chosen model $\mathcal{B}_{rand}$. The parameters of an SR of $\mathcal{B}_{rand}$ were obtained by a random sample from the standard normal distribution.
In Table 5.1 the resulting relative misfits are listed, i.e. the misfits divided by the norm of the data. The first row shows the initial misfits. The next three rows contain the misfits of the models obtained by applying each of the constructions separately and only once. This shows that each individual construction can give a significant decrease of the misfit. The last row shows the misfit resulting from applying these constructions iteratively until convergence.

**Table 5.1**: Model improvements.

| model | $\mathcal{B}_{ex}$ | $\mathcal{B}_{rand}$ |
|---|---|---|
| relative misfit | 0.2287 | 0.7368 |
| construction 1 | 0.2249 | 0.2340 |
| construction 2 | 0.2251 | 0.2420 |
| construction 3 | 0.2286 | 0.2333 |
| limit | 0.2236 | 0.2236 |

Construction 1 transforms model $\mathcal{B}_{rand}$ with relative misfit 0.7368 into an improved model with relative misfit 0.2340. Applying construction 1, 2 and 3 iteratively leads to a sequence of relative misfits that converges to 0.2236, both for $\mathcal{B}_{ex}$ and $\mathcal{B}_{rand}$ as initial model.

For both initial models the decrease of misfit becomes below $10^{-10}$ after about 20 iterations. Later on we compare the performance of the algorithm with the

Gauss-Newton algorithm, in Section 5.5. The limiting model is the same in both cases, which suggests that it is (at least locally) optimal. It is given by

$$\mathcal{B}_{gtls} = \{(u,y) : \mathsf{Z} \to \mathsf{R}^2 \; y(t) = 0.67y(t-1) + 1.84u(t) - 1.96u(t-1)\}. \quad (5.10)$$

The parameters of this system are relatively close to those of the data generating system, cf. equation (3.2). ◇

## 5.3 Optimality Conditions

In this section we investigate whether the model improvement constructions in the previous section are powerful enough to determine locally optimal models. First we derive necessary conditions for optimality from the fact that for an optimal model the model improvement constructions can give no improvement. Although we conjecture that the following result is valid for all stabilizable systems, we only managed to prove it for *controllable* systems, i.e. systems for which all trajectories on finite time admit a continuation that becomes zero within finite time. We express the optimality conditions in terms of empirical covariances. For two sequences $a : T \to \mathsf{R}^k$ and $b : T \to \mathsf{R}^l$ this is defined as $\mathrm{cov}(a,b) := \Sigma_{t\in T} a(t)b(t)^\top \in \mathsf{R}^{k\times l}$. Further, by $\mathrm{cov}([a_1,a_2],[b_1,b_2],)$ we denote the covariance matrix of the combined trajectories $[a_1^\top \; a_2^\top]^\top$ and $[b_1^\top \; b_2^\top]^\top$.

**Theorem 5.3.1 (Optimality Conditions)** *Let $\mathcal{B}$ denote a controllable GTLS model for an observation $w$. Let $\hat{w} \in \mathcal{B}$ denote the optimal approximation of $w$, and $\tilde{w} = w - \hat{w}$ the corresponding approximation error. Let $\hat{x}$, $\hat{v}$ denote respectively the state and auxiliary input corresponding to $\hat{w}$ in a minimal state representation of $\mathcal{B}$, and let $\tilde{x}$, $\tilde{v}$ be defined analogously for $\tilde{w}$ in $\mathcal{B}^\perp$. Then the following equivalent conditions hold:*

1. $\mathrm{cov}(\hat{v},\tilde{v}) = 0, \mathrm{cov}(\hat{v},\tilde{x}) = 0$ *and* $\mathrm{cov}(\hat{x},\tilde{v}) = 0$;
2. $\mathrm{cov}([\hat{v},\hat{x}],[\tilde{v},\tilde{x}]) = 0$;
3. $\mathrm{cov}([\hat{w},\sigma\hat{x}],[\tilde{w},\sigma\tilde{x}]) = 0$;
4. $\mathrm{cov}([\hat{v},\hat{x},\hat{w},\sigma\hat{x}],[\tilde{v},\tilde{x},\tilde{w},\sigma\tilde{x}]) = 0$;

*Proof.* See Appendix A.5.

In practice, in order to evaluate to what extent these conditions are satisfied it may be useful to consider the empirical correlations, i.e., the covariances scaled by the magnitude of the variables.
Next we investigate to what extent these conditions are sufficient for optimality. It is not difficult to check that the number of free parameters in $(A,B,C,D)$ modulo the equivalence of Proposition 3.1.4, is given by $nq+m(q-m)$. This is precisely the number of equations in Theorem 5.3.1.1, and the other conditions

can be derived from these. This indicates that the conditions determine a finite number of models.
In fact, these conditions characterize the *stationary points* with respect to the GTLS criterion (5.6). Stationary points are those values of parameters in which a differentiable criterion function has zero derivative. Notice that $G$ as defined is (5.5) is a polynomial function, so this notion is well-defined for $G$. With a slight abuse of terminology, we also call the approximation and system corresponding to a stationary point of $G$ a stationary point.

**Theorem 5.3.2 (Characterization of Stationary Points)** *A system $\mathcal{B}$ satisfies the optimality conditions of Theorem 5.3.1 if and only if $\mathcal{B}$ is a stationary point of the GTLS criterion.*

*Proof.* See Appendix A.5.

This establishes the fact that only for stationary points none of the improvement constructions yields a decrease of misfit. So by iteratively applying these constructions stationary points can be obtained. In order to be specific we formulate the following algorithm.

**Algorithm 3 (Model Improvement Algorithm)**

**Data:**
- *An observation $w : T \to \mathsf{R}^q$, $T = \{1, \ldots, N\}$*
- *An isometric state representation $(A, B, C, D)$ with $m$ auxiliary inputs and $n$ states, corresponding to an initial model $\mathcal{B} \in \mathsf{B}^{q,m,n}$*

**Step 1:** *Determine optimal $\hat{v}$ and $\hat{x}(1)$ by Algorithm 1.*

**Step 2:** *Determine optimal $C$ and $D$ by the first construction in Theorem 5.2.1, and transform the result to an equivalent ISR.*

**Step 3:** *Determine optimal $B$, $D$ and $\hat{x}(1)$ by the second construction in Theorem 5.2.1, and transform the result to an equivalent ISR.*

**Step 4:** *Determine optimal $B$, $D$ and $\hat{v}$ in an ISR, i.e., such that (3.7) remains valid, by the third construction in Theorem 5.2.1, and redefine $\mathcal{B}$ as the system corresponding to the resulting ISR.*

**Step 5:** *If the misfit has decreased go to step 1, else stop.*

**Result:** *$\mathcal{B}$ is a stationary point with respect to the GTLS criterion.*

*Proof.* If the misfit has not decreased in Step 2, 3 and 4, then none of the model improvement constructions yields an improvement. This implies that the optimality conditions in Theorem 5.3.1 are valid. Now the result follows from Theorem 5.3.2. ♣
*Implementation:* `GTLS` in Section B.2.

We remark that we do not have theoretical or practical reason for the order in which the constructions are applied.
Of course, in practice the algorithm is stopped when the decrease of misfit has come below a certain threshold, depending on the required accuracy. So the result should be interpreted as that a stationary point is approximated within arbitrary precision. In the next section we develop a measure for the distance of models to an exact stationary point. Although the algorithm leads to a converging sequence of misfits, it is conceivable that it might yield a non-converging series of stationary points. This seems a very unlikely situation, at least for generic data. Moreover, also in that case the algorithm does yield a stationary point within arbitrary accuracy. As much as it is worth, we mention that we never encountered convergence problems in any of our simulations. We leave a thorough analysis of the convergence properties of the algorithm as an issue for further research.

**Leading Example - Continued.** The covariances of Theorem 5.3.1 give a first indication of the optimality of a model. In order to make this scale invariant we consider the correlations in an ISR. For the nominal model $\mathcal{B}_{ex}$ they are around 0.3, for the randomly chosen model $\mathcal{B}_{rand}$ correlations of 0.8 occur, while for the model $\mathcal{B}_{gtls}$ in (5.10) the correlations are approximately zero, below $10^{-5}$. This confirms that $\mathcal{B}_{gtls}$ is indeed (locally) optimal.
Next we investigate whether the model order can be deduced from the data. For this purpose we determine models of various degree by Algorithm 3. Their misfits are listed in Table 5.2. The misfit of the optimal static model is given by the smallest singular value of $w$.

**Table 5.2**: Order selection.

| order | $n=0$ | $n=1$ | $n=2$ | $n=3$ | $n=4$ |
|---|---|---|---|---|---|
| rel. misfit | 0.295 | 0.224 | 0.218 | 0.210 | 0.201 |

The first order GTLS model for the data $w = (u, y)$ given by (4.19) is substantially more accurate than the static GTLS model, and only slightly less accurate than higher order models.

This motivates the choice of a first order model. It is substantially better than the static model, and a further increase of the order gives only small improvements. ◇

## 5.4 Optimality Margin

The foregoing results can be used to analyse whether a proposed system $\mathcal{B}$ is close to optimality, or more precisely, whether it is close to a stationary point. This is for example relevant in the formulation of stopping criteria for the iterative algorithm of Section 5.3. Perhaps the most convincing way to evaluate optimality is to consider the distance between $\mathcal{B}$ and a GTLS model

$\mathcal{B}^*$, as defined in Definition 2.1.7. However, this is in general not feasible as it would require the knowledge of $\mathcal{B}^*$. Instead of asking how much the system should be changed to become optimal for the observed data $w$, we will consider the question how much these data should be changed to make the given system optimal. For pragmatic reasons we consider the distance to the nearest stationary point, defined as

$$\min\{\|\bar{w}\|;\ \mathcal{B} \text{ is stationary for } w - \bar{w}\}. \tag{5.11}$$

Because it seems difficult to evaluate this distance exactly, we present an upper bound that is relatively easy to compute. This upper bound is obtained by allowing only adjustments of the data that belong to $(\mathcal{B}_T)^\perp$, so that the optimal approximation of the data within $\mathcal{B}_T$ is not affected. This leads to the following definition of the optimality margin.

**Definition 5.4.1 (Optimality Margin)** *The optimality margin of a system $\mathcal{B}$ with respect to an observation $w$ is defined as*

$$\min\{\|\bar{w}\|;\ \bar{w} \in (\mathcal{B}_T)^\perp \text{ and } \mathcal{B} \text{ is stationary for } w - \bar{w}\}. \tag{5.12}$$

The following result shows that the computation of the optimality margin is relatively easy.

**Proposition 5.4.2 (Optimality Margin)** *Let $\hat{w}$ denote the optimal approximation of $w$ in $\mathcal{B}_T$ and let $\tilde{w} := w - \hat{w}$ denote the corresponding approximation error. Further define $\mathcal{Z} := \{z \in (\mathcal{B}_T)^\perp;\ \mathcal{B} \text{ is stationary for } \hat{w} + z\}$. Then $\mathcal{Z}$ is a linear space, and the optimality margin is given by $\|\tilde{w} - \tilde{w}'\|$, where $\tilde{w}'$ is the orthogonal projection of $\tilde{w}$ on $\mathcal{Z}$.*

*Proof.* See Appendix A.5.

Under the assumption that a model $\mathcal{B}$ is globally optimal for adjusted data $w - \bar{w}$, the optimality margin can also be used to determine a lower bound for the minimal misfit for $w$, as follows.

**Proposition 5.4.3 (Bounds for the Minimal Misfit)** *Let be given an observation $w$, and let $e^*$ be the minimally achievable misfit under a certain size constraint. Further let $\mathcal{B}$ be a GTLS model of tolerated size for adjusted data $w - \bar{w}$ with misfit $e := \mathrm{d}(w, \mathcal{B})$. Then it holds*

$$e - 2\|\bar{w}\| \leq e^* \leq e \tag{5.13}$$

*Proof.* Let $\mathcal{B}^*$ denote a GTLS model for the original data $w$, so that $\mathrm{d}(w, \mathcal{B}^*) = e^*$. Then the upper bound follows from the optimality of $\mathcal{B}^*$. For the lower bound we use the properties of the misfit that $\mathrm{d}(w, \mathcal{B}) \leq \|w\|$ and $\mathrm{d}(w_1 + w_2, \mathcal{B}) \leq \mathrm{d}(w_1, \mathcal{B}) + \mathrm{d}(w_2, \mathcal{B})$, so that $e = \mathrm{d}(w, \mathcal{B}) \leq \mathrm{d}(\bar{w}, \mathcal{B}) + \mathrm{d}(w - \bar{w}, \mathcal{B}) \leq \|\bar{w}\| + \mathrm{d}(w - \bar{w}, \mathcal{B}^*) \leq \|\bar{w}\| + \mathrm{d}(w, \mathcal{B}^*) + \mathrm{d}(\bar{w}, \mathcal{B}^*) \leq 2\|\bar{w}\| + e^*$. Here we have used the optimality of $\mathcal{B}$ for $w - \bar{w}$ in the second inequality. □

**Leading Example - Continued.** In the previous part of the leading example we evaluated the optimality of $\mathcal{B}_{gtls}$ on basis of correlations. From the optimality margins we obtain more precise information about the optimality of the systems. They are listed in Table 5.3.

**Table 5.3**: Optimality margins.

| model | $\mathcal{B}_{gtls}$ | $\mathcal{B}_{ex}$ | $\mathcal{B}_{rand}$ |
|---|---|---|---|
| optimality margin | $1.46 * 10^{-6}$ | 1.04 | 15.06 |

The model $\mathcal{B}_{gtls}$, given by (5.10), is exactly a stationary point for $w - \bar{w}$, with $w$ the data generated according to (4.19), and $\bar{w}$ a time series with $\|\bar{w}\| = 1.46 * 10^{-6}$. There does not exist a smaller change $\bar{w}$ for which $\hat{w}$ (the optimal approximation in $\mathcal{B}_{gtls}$ of $w$) is exactly a stationary point of $w - \bar{w}$.

This shows that it requires only a change of the observation of the order $\|\bar{w}\| \approx 10^{-5}$ to make $\mathcal{B}_{gtls}$ a stationary point, cf. (5.12). Now assume that $\mathcal{B}_{gtls}$ is globally optimal for $w - \bar{w}$, which is a reasonable assumption. The evaluation of the bounds in Proposition 5.4.3 for the data in this example with $e = 5.0792$ and $2\|\bar{w}\| \approx 10^{-5}$ shows that the optimal misfit $e^* \approx 5.0792$ is determined within an accuracy of $10^{-5}$. This also shows that $\mathcal{B}_{gtls}$ is optimal within this accuracy level. $\diamond$

## 5.5 A Gauss-Newton Algorithm for GTLS

In this section we present a Gauss-Newton algorithm for the construction of stationary points of the GTLS criterion. As already mentioned in the previous section, it is in general much faster than Algorithm 3 in Section 5.3. Moreover, the results on system representations underlying this algorithm might be of independent interest.

The starting point of Gauss-Newton methods is the function ($H$, say) that maps model parameters to the sequence of corresponding residuals whose sum of squares has to be minimized. In each iteration the parameter values are improved on basis of the derivative of this function. More precisely, for given parameters $\theta_0$ new parameter values $\theta_0 + \bar{\theta}$ are determined that solve the quadratic optimization problem in terms of the Jacobian of $H$,

$$\text{minimize} \|H(\theta_0) + \left( \frac{dH}{dt} |_{\theta_0} \right) \bar{\theta}\|. \tag{5.14}$$

This is a quadratic optimization problem, as the derivative of $H$ in the point $\theta_0$ is a linear function of $\bar{\theta}$. The details are discussed below. Descriptions of Gauss-Newton methods can be found in many textbooks on system identification and econometrics, e.g. [24, 8]. In the statistical literature it is called the method of scoring, cf. [31]. In [30] an interpretation is given of the Gauss-Newton method as a Riemannian steepest descent algorithm, cf. also [18].

The Gauss-Newton algorithm for GTLS is based on a formulation in terms of ISR's, based on equation (4.15) in Algorithm 1. Let $\tilde{v}$ denote the effect of final state $\hat{x}_{N+1}$ in (4.15) for ISR's $(A, \tilde{B}, C, \tilde{D})$, so

$$\begin{aligned} x(t) &= A^\top x(t+1) + C^\top w(t) \\ \tilde{v}(t) &= \tilde{B}^\top x(t+1) + \tilde{D}^\top w(t), \end{aligned} \tag{5.15}$$

So, although $\tilde{v}$ are not residuals in the literal sense, it is indeed their sum of squares that have to be minimized. According to this we define the function $H$ that maps an ISR and final state to the corresponding sequence $\tilde{v} : T \to \mathsf{R}^{q-m}$, i.e.,

$$H((A, \tilde{B}, C, \tilde{D}), x_{N+1}) = \tilde{v}. \tag{5.16}$$

The ISR $(A, \tilde{B}, C, \tilde{D})$ represents the orthogonal complement of a system $\mathcal{B}$ with ISR $(A, B, C, D)$, cf. Proposition 4.2.1.
One of the results in Algorithm 1 it that for an appropriate choice of $x_{N+1}$, $\|\tilde{v}\|$ equals the misfit of $\mathcal{B}$, so the GTLS problem amounts to

$$\text{GTLS-ISR: minimize} \|H((A, \tilde{B}, C, \tilde{D}), x_{N+1})\| \tag{5.17}$$

for $\hat{x}_{N+1} \in \mathsf{R}^n$, and for $A, \tilde{B}, C, \tilde{D}$ satisfying

$$\begin{bmatrix} A & \tilde{B} \\ C & \tilde{D} \end{bmatrix}^\top \begin{bmatrix} A & \tilde{B} \\ C & \tilde{D} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_{q-m} \end{bmatrix}$$

.

**Remark.** There is a minor difference between (5.17) and the GTLS problem in terms of SR's as formulated in (5.6), as non-stabilizable systems do not have an ISR, cf. Proposition 3.2.4. In view of our conjecture 5.1.1.2, we ignore this difference in the sequel.
Secondly, we briefly discuss the existence of a solution for (5.17), cf. also 5.1.1.1. Notice that $H$ is a polynomial function, hence continuous. If we bound the norm of $x_{N+1}$ by a constant $c \in \mathsf{R}$, its domain is a compact set, so $H$ achieves its minimum on this domain. This implies that if there does not exist a solution, the infimum of $H$ corresponds to an infinite norm of $x_{N+1}$, cf. also the example (5.2), which seems to be a non-generic case. Moreover, this problem is easily detected in numerical simulations, by considering $\|x_{N+1}\|$ in each iteration. $\diamond$

Notice the remarkable reduction of number of parameters with respect to the formulation in terms of SR's (5.6) which also concerns the optimization over auxiliary inputs. The price we have to pay for this is that the parameter space is of a more complex nature, as the isometry condition is quadratic in the system parameters. So we have obtained a simpler formula for the misfit on a more complex parameter set.

With a slight abuse of terminology we call $((A, \tilde{B}, C, \tilde{D}), x_{N+1})$ a stationary point of $H$ if the derivative of $\|H((A, \tilde{B}, C, \tilde{D}), x_{N+1})\|$ is zero. Our aim is to determine stationary points by Gauss-Newton iterations. We first introduce some notation.

- $w$ is a given observation on the time interval $[1, N]$ with $q$ components
- $(m, n)$ denotes the tolerated size
- $(A, B, C, D)$ denotes an ISR with $m$ auxiliary inputs and $n$ state variables
- $\mathcal{I}$ is the space of ISR's with $q-m$ auxiliary inputs and $n$ state variables, so $H$ is a function on $\mathcal{I} \times \mathsf{R}^n$.
- Elements $(A, \tilde{B}, C, \tilde{D}) \in \mathcal{I}$ represent the orthogonal complement of models $\mathcal{B}(A, B, C, D)$ as described in Proposition 4.2.1.2.
- $\mathcal{U}_k$ denotes the space of $k \times k$ unitary matrices, so that

$$\begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix} \in \mathcal{U}_{n+q} \tag{5.18}$$

- $T_M\mathcal{I}$ is the tangent space of $\mathcal{I}$ in the ISR $M \in \mathcal{I}$. This consists of a linear space of SR's through $M$ tangent to $\mathcal{I}$. Elements $\Delta_M \in T_M\mathcal{I}$ are called a variation of $M$ in the direction of other ISR's, or simply a variation of $M$. Similarly, elements of $T_M\mathcal{U}_k$ are called variations of $M$ in the direction of other unitary matrices.

Our Gauss-Newton algorithm for determining stationary points of the function $H$ has the following structure.

1. Choose initial values in the domain of $H$, i.e., choose an ISR $M := (A, \tilde{B}, C, \tilde{D}) \in \mathcal{I}$ and a value for the end state $x_{N+1} \in R^n$. Let $H_0$ denote the value of $H$ in the initial point.

2. Determine how $\tilde{v}$ varies if the initial values are varied, i.e., if the initial ISR is varied in the direction of another ISR, and if the end state is varied. This amounts to determining the derivative $H'$ of $H$ in the initial point, or, in geometrical terms, determining the tangent space of the image of $H$ in that point.

3. Determine the optimal variation of the initial ISR and end state on the basis of this derivative, i.e., determine $(\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}})$ and $\bar{x}_{N+1}$ that minimizes $\|H_0 + H'((\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}}), \bar{x}_{N+1})\|$[1].

   Stated in geometrical terms, determine the optimal variation on the basis of the tangent space of im$H$ which is considered as an approximation of im$H$ in $H_0$.

[1] $H'((\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}}), \bar{x}_{N+1})$ is the derivative of $H$ in the (fixed) point $((A, B, C, D), x_{N+1})$, in the direction $(\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}}), \bar{x}_{N+1})$. This is often denoted as $H' \mid_{(\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}}), \bar{x}_{N+1})} ((A, B, C, D), x_{N+1})$, but we consider the derivative as a mapping of parameter variations with $((A, B, C, D), x_{N+1})$ fixed.

4. If the resulting model is equal to the initial one, it is a stationary point, so stop. If the resulting model improves the previous one, repeat the algorithm for this model, otherwise first halve the size of the variation repeatedly until it corresponds to a better model.

In deriving a Gauss-Newton algorithm for GTLS there are two aspects that deserve additional attention.

- The domain of $H$, $\mathcal{I} \times \mathsf{R}^n$, is determined by a non-linear restriction, cf. (5.18).
- ISR's for a given system are not unique, cf. Proposition 3.2.6

Concerning the first point, we have to analyse how to vary an ISR into the direction of other ISR's. For expository reasons we not only consider variations of $(A, \tilde{B}, C, \tilde{D})$, but also the corresponding variations of $(A, B, C, D)$. So this amounts to the question how to vary a unitary matrix as given in (5.18) into the direction of another unitary one, which is answered in the following lemma.

**Lemma 5.5.1** *The tangent space of the space of unitary matrices $\mathcal{U}$ in a point $M$ is given by $T_M\mathcal{U} = \{MK;\ K + K^\top = 0\}$.*

*Proof.* See Appendix A.5.

Notice that for a variation $MK \in T_M\mathcal{U}$, the matrix $M + MK$ is unitary up to a quadratic term in $K$ as $(M + MK)^\top(M + MK) = I + K^\top K$. Although the quadratic term is negligible if $K$ is sufficiently small, this implies that the state representation corresponding to $M+K$ is not exactly isometric anymore. Therefore, after computing the optimal variation in Step 3 of the Gauss-Newton algorithm, we have to transform the resulting state representation to an equivalent ISR.
The non-uniqueness of ISR's can be exploited to decrease the number of variations that we have to take into account in the Gauss-Newton algorithm. More precisely, we will show that without loss of generality we can restrict our attention to variations of (5.18) of the form

$$\begin{bmatrix} \Delta_A & \Delta_B & \Delta_{\tilde{B}} \\ \Delta_C & \Delta_D & \Delta_{\tilde{D}} \end{bmatrix} = \begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix} \begin{bmatrix} 0 & L & P \\ -L^\top & 0 & Q \\ -P^\top & -Q^\top & 0 \end{bmatrix}, \tag{5.19}$$

This means that for all variations as described in Lemma 5.5.1, there exists an equivalent variation corresponding to zero diagonal blocks of $K$. Moreover, if $(A, B, C, D)$ is minimal, different variations of this form are not equivalent, i.e., they all correspond to changes into the direction of different systems. This means that generically we have taken full account of the non-uniqueness of ISR's in this way, and that (5.19) are canonical forms for the perturbations of ISR's.

We use this result in the following Gauss-Newton algorithm for $H$. In each iteration we determine the variation of parameters of the form (5.19) (and a variation $\bar{x}_{N+1}$ of $x_{N+1}$) that minimizes

$$\|H((A,B,C,D),x_{N+1}) + H'((\Delta_A,\Delta_{\tilde{B}},\Delta_C,\Delta_{\tilde{D}}),\bar{x}_{N+1})\|. \tag{5.20}$$

**Algorithm 4 (Gauss-Newton for GTLS)**

**Data:**
- *An observation* $w : T \to \mathsf{R}^q$, $T = \{1,\ldots,N\}$
- *A bound* $(m,n)$ *for the model size*

**Step 1:** *Choose an initial isometric state representation* $(A,\tilde{B},C,\tilde{D}) \in \mathcal{I}$ *and determine* $B$ *and* $D$ *such that* $\begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix}$ *is a unitary matrix. Construct* $x$, $\hat{v}$ *and* $\tilde{v}$ *according to Algorithm 1 in Section 4.3. This gives an initial value for* $x_{N+1}$ *that is optimal for the chosen initial ISR, so that* $\tilde{v}$ *in Algorithm 1 equals* $H(M,x_{N+1}) =: H_0$.

**Step 2:** *The derivative* $\bar{v} := H'(\Delta_M,\bar{x}_{N+1})$ *of* $H$ *in* $(M,x_{N+1})$ *for* $\Delta_M := (\Delta_A,\Delta_{\tilde{B}},\Delta_C,\Delta_{\tilde{D}})$ *of the form (5.19) is given by*

$$\begin{aligned} \bar{x}(t) &= A^\top \bar{x}(t+1) + L\hat{v}(t) + P\tilde{v}(t); \bar{x}(N+1) = \bar{x}_{N+1} \\ \bar{v}(t) &= \tilde{B}^\top \bar{x}(t+1) - P^\top x(t) - Q^\top \hat{v}(t)\}. \end{aligned} \tag{5.21}$$

*Define*

$$\mathcal{T} := \{\bar{v};\ \exists L,P,Q,\bar{x}(N+1) \text{ such that } \bar{v} \text{ satisfies } (5.21)\}, \tag{5.22}$$

*which can be considered as the relevant part of the tangent space of* $\mathrm{im}H$ *in* $\tilde{v}$.

**Step 3:** *Compute the orthogonal projection of* $\tilde{v}$ *onto* $\mathcal{T}$, *denoted by* $\tilde{v}'$, *and let* $L$, $P$, $Q$ *and* $\bar{x}_{N+1}$ *denote the corresponding values of the parameters in (5.22). Define* $\Delta_A,\Delta_B,\Delta_C,\Delta_D,\Delta_{\tilde{B}},\Delta_{\tilde{D}})$ *according to (5.19). Redefine* $(A,B,C,D)$ *as an ISR for* $\mathcal{B}(A-\Delta_A,B-\Delta_B,C-\Delta_C,D-\Delta_D)$, *and redefine* $\tilde{B},\tilde{D}$ *by (5.18).*

**Step 4:** *If the orthogonal projection* $\tilde{v}$ *is zero, then* $\mathcal{B}(A,B,C,D)$ *is a stationary point, so stop. If the misfit* $\mathrm{d}(w,\mathcal{B}(A,B,C,D))$ *has decreased, repeat the algorithm for this model, otherwise first halve* $L,P,Q$ *and* $\bar{x}(N+1)$ *repeatedly, until the corresponding model improves the initial one.*

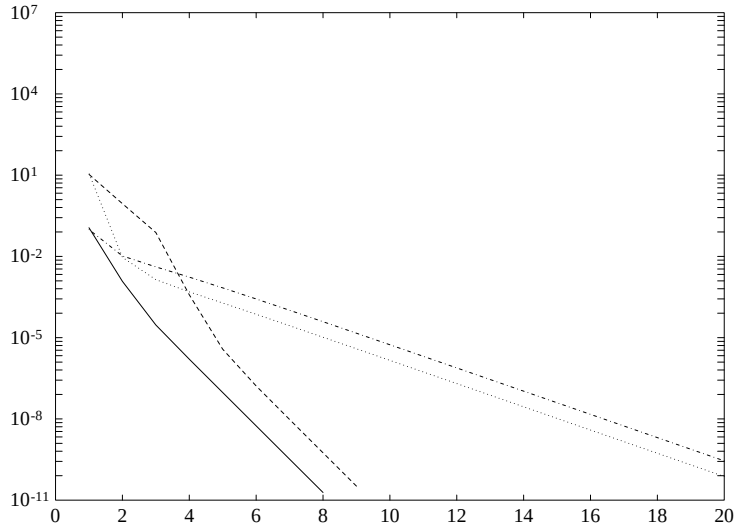
**Result:** $\mathcal{B}(A,B,C,D)$ *is a stationary point with respect to the GTLS criterion.*

*Proof.* See Appendix A.5.
*Implementation:* `GTLS` in Appendix B.2.

The implementation concerns the following issues. In addition to Algorithm 1, which is used in Step 1 and 4, Step 3 involves the computation of an orthogonal projection in $\mathsf{R}^{p\times N}$ to determine the optimal variation, and solving an $n \times n$ Riccati equation for the transformation to isometric form, as is described in Proposition 3.2.5.

**Leading Example - Continued.** We apply Algorithm 4 for the data (4.19) and initial models $\mathcal{B}_{ex}$ and $\mathcal{B}_{rand}$, cf. Section 5.3. This results in the same model determined by Algorithm 3 given by (5.10). In Fig. 5.1 the decrease of misfit in each iteration is shown, and compared with the results of Algorithm 3. This shows that the Gauss-Newton algorithm is considerably faster than Algorithm 3.

**Figure 5.1**: Performance of the Gauss-Newton algorithm and the model improvement constructions compared.




Decrease of misfit in each iteration in the Gauss-Newton algorithm (solid line: initial model $\mathcal{B}_{ex}$, dashed line: initial model $\mathcal{B}_{rand}$), and in Algorithm 3 (dash/dotted line: initial model $\mathcal{B}_{ex}$, dotted line: initial model $\mathcal{B}_{rand}$).

$\diamond$

## 5.6 Initial Models

As the iterative algorithms only determine locally optimal solutions, their outcome may be determined by the choice of initial models. In order to increase the chance to converge to a globally optimal solution of the GTLS problem it is of importance to determine good initial models that are close to the global optimum. In this section we present a heuristic method for this purpose.

The procedure consists of two steps. First we construct an approximate state trajectory for the observed time series. Our method is an approximate analogue of an exact realization algorithm in [44, part II, algorithm 6], and has been described in [33]. The main idea is to construct a state trajectory *directly from the data*, without first determining an (approximate) model. The state is expressed in terms of canonical correlations between past and future of the observations. These correlations also give an indication for a reasonable dimension of the state space. This step resembles the stochastic realization algorithm described in [1], see also [4].
As a second step we use the constructed approximate state trajectory to determine an SR of an approximate model for the observed time series. The system matrices are determined as a total least squares solution of equations in terms of the constructed state trajectory, by means of the SVD.
Nowadays there has been developed a whole class of identification algorithms in which first an (approximate) state space is constructed from the data, and then the system matrices in a state representation. For an overview on these so-called *subspace methods* we refer to [11, 40, 41]. The distinctive feature of our approach is that we consider all variables in a symmetric way, i.e., the decomposition into inputs and outputs need not be given.
A disadvantage of the canonical correlation approach is that it does not take into account the size or energy content of the data in different directions, as correlations are scale-invariant. This implies that highly correlated directions in past and future with only a tiny energy content are preferred over slightly less correlated dominant components. This may lead to less satisfactory results, and therefore we describe how this method can be modified in order to suppress this effect.

### 5.6.1 The Exact Case

As an introduction to the procedure we first describe the relatively simple case in which the data consists of an exact system trajectory

$$\hat{w} : \mathsf{Z} \to \mathsf{R}^q \in \mathcal{B}. \tag{5.23}$$

Let $(m, n)$ denote the size of $\mathcal{B}$ , so $\mathcal{B}$ is the behaviour of a system with $m$ (auxiliary) inputs and state dimension $n$. Further we suppose that there does not exist a system of smaller size that contains $\hat{w}$. From the existence of SR's, cf. Proposition 3.1.2, it follows that such a trajectory takes the form

$$\sigma\hat{x} = A\hat{x} + B\hat{v};\ \hat{w} = C\hat{x} + D\hat{v}, \tag{5.24}$$

with $\hat{x}$ an $n$-dimensional state trajectory, $\hat{v}$ an $m$-dimensional auxiliary input, and $(A, B, C, D)$ a minimal SR. The basic question is how to construct (5.24) from (5.23), including $m$ and $n$. The idea is first to determine a valid state trajectory $\hat{x}$ , and then the remaining parameters in (5.24).
This raises the question to what extent $\hat{x}$ is unique for $\hat{w}$ . Clearly it is not completely unique, as with a solution $\hat{x}$ , also $S\hat{x}$ is a solution for all

non-singular matrices $S \in \mathsf{R}^{n\times n}$, cf. Proposition 3.1.4. This choice of basis, however, is the only source of non-uniqueness.

**Lemma 5.6.1** *The state trajectory $\hat{x}$ in (5.24) is uniquely determined by the data $\hat{w}$, modulo a basis transformation of the state space.*

*Proof.* In [44, Section 17] it is shown that the system $\mathcal{B}$, which we have defined as the smallest system containing $\hat{w}$, is uniquely determined by $\hat{w}$ (it is called the 'most powerful unfalsified model' for $\hat{w}$). The uniquenes of $\hat{x}$ modulo a basis transformation $S$ then follows from Proposition 3.1.4. ♣

So we conclude that the state trajectory for $\hat{w}$ is 'essentially' unique. In [44, Section 17] it is described how to construct the state trajectory directly from an observed time series.
The key idea behind this realization algorithm is that the state is characterized by the property that it is both a linear function of the finite past and of the finite future of $\hat{w}$.

**Lemma 5.6.2** *Let $\mathcal{H}^-$ and $\mathcal{H}^+$ denote the space of time series spanned by the finite past and future of $\hat{w}$ respectively, i.e.,*

$$\begin{array}{l}\mathcal{H}^- = \{z : \mathsf{Z} \to \mathsf{R};\\ \quad z(t) = L^-\mathrm{col}(\hat{w}(t-1),\ldots,\hat{w}(t-k)) \text{ for some } k \in \mathsf{N}, L^- \in \mathsf{R}^{1\times qk}\}\\ \mathcal{H}^+ := \{z : \mathsf{Z} \to \mathsf{R};\\ \quad z(t) = L^+\mathrm{col}(\hat{w}(t),\ldots,\hat{w}(t+k-1)) \text{ for some } k \in \mathsf{N}, L^+ \in \mathsf{R}^{1\times qk}\}\end{array} \tag{5.25}$$

*and define $\mathcal{X}$ as their intersection. Then $\mathcal{X}$ is a finite dimensional space, and a state trajectory $\hat{x}$ for $\hat{w}$ consists of an (arbitrary) basis of $\mathcal{X}$.*

*Proof.* See Appendix A.5.

The actual reconstruction of the state trajectory is performed in terms of Hankel matrices of the data, which consist of blockrows containing time shifts of the data. We make use of the following notation.

$$H_{2k} := \left[\begin{array}{ccccc} \ldots & \hat{w}(-k) & \hat{w}(1-k) & \hat{w}(2-k) & \ldots\\ & \vdots & \vdots & \vdots & \\ \ldots & \hat{w}(-1) & \hat{w}(0) & \hat{w}(1) & \ldots\\ \ldots & \hat{w}(0) & \hat{w}(1) & \hat{w}(2) & \ldots\\ & \vdots & \vdots & \vdots & \\ \ldots & \hat{w}(k-1) & \hat{w}(k) & \hat{w}(k+1) & \ldots \end{array}\right] \begin{array}{l}\left.\begin{array}{l}\\ \\ \end{array}\right\} =: H_k^-\\ \left.\begin{array}{l}\\ \\ \end{array}\right\} =: H_k^+\end{array} \tag{5.26}$$

The interpretation is that in each column of $H_{2k}$ the upper part represents the past and the lower part the future at the time instant corresponding to the first row of $H_k^+$. So by convention the present is included in the future. If the degree of the system containing $\hat{w}$ is known, then we can choose $k = n$. We remark that in fact every value $k \geq \nu$ will do with $\nu$ the observability index of the system containing $\hat{w}$, which can be considerably smaller than its degree.

If $n$ is unknown, in fact we should take $k = \infty$, which means in practice that $k$ should be chosen 'large enough'.
From Lemma 5.6.2 it follows that a state trajectory for $\hat{w}$ is characterized by the fact that it is a linear combination both of the rows in $H_k^-$ and $H_k^+$ if $k \geq n$. So if $n$ is known, the construction of a minimal state trajectory amounts to determining $L^-, L^+ \in \mathsf{R}^{nq \times n}$, both of rank $n$, such that

$$L^- H_n^- = L^+ H_n^+ =: \hat{x} \tag{5.27}$$

If $n$ is unknown, we should determine $L^-$ and $L^+$ of maximal rank, say $n_k$. Clearly $n_k$ is non-decreasing with $k$, and bounded by the actual minimal state dimension for $\hat{w}$. So $n$ can be determined as the dimension of the intersection of the rowspan of $H_k^-$ and $H_k^+$ for large enough $k$.
As a final step the system matrices $A, B, C, D$ and the auxiliary input $\hat{v}$ in (5.24) can be determined from the static linear relations that are satisfied by $(\hat{w}(t), \hat{x}(t), \hat{x}(t+1))$. We discuss the details in the context of non-exact data.

### 5.6.2 Approximate State Trajectories

We have described how to construct a state trajectory for an observed time series that belongs to an (unknown) system as the intersection of the rowspan of the past and future Hankel matrices. However, if a time series does not satisfy any linear time-invariant relationship exactly, then this intersection is zero, as there are no exact linear relations between past and future. The question arises how to define and construct an approximate state trajectory for such time series. An obvious choice is to consider a state trajectory in the optimal GTLS system for the data (of reasonable size), but a direct construction of this from the data seems not feasible. Several methods for constructing approximate states from observations have been developed. We refer to [10, 39] for an overview.
In our procedure we construct the state by determining an approximate intersection of past and future in the following way. Let $H^-$ and $H^+$ denote respectively a past and future Hankel matrix for $\hat{w}$ with a certain number of blockrows $k$, defined as in (5.26) but now for finite data. We discuss the choice of $k$ later on. Let $\mathcal{H}^-$ and $\mathcal{H}^+$ denote the *rowspaces* of respectively $H^-$ and $H^+$, i.e., the linear space that is spanned by the rows in these matrices. The first direction $\ell_1$ in the approximate intersection is chosen such that the maximum of the angles with $\mathcal{H}^-$ and $\mathcal{H}^+$ is minimal. For one dimensional $\mathcal{H}^-$ and $\mathcal{H}^+$ this would give the bisecting line. The $i$-th direction $\ell_i$ is determined by the same criterion, under the extra condition that $\ell_i$ is orthogonal to the previously constructed directions $\ell_1, \ldots, \ell_{i-1}$. The dimension of the intersection can be based on the corresponding angles. In the next section we discuss the implementation of this procedure.
This approximate intersection is closely related to the canonical variables of the pair $(\mathcal{H}^-, \mathcal{H}^+)$. These are defined as follows. The first canonical angle is the smallest possible angle between two lines $\ell_1^-$ and $\ell_1^+$ in the two spaces respectively. The second canonical angle is the smallest possible angle between

two lines in the two spaces, under the extra condition that these lines are orthogonal to the previous ones. The other canonical angles are determined in a similar way. The pair of unit vectors $x_i^- \in \ell_i^-$ and $x_i^+ \in \ell_i^+$ are called the $i$-th canonical variables. The angle $\alpha_i$ between $\ell_i^-$ and $\ell_i^+$ is called the $i$-th canonical angle, and $\cos(\alpha_i)$ the $i$-th canonical correlation. Now the $i$-th direction in the approximate intersection $\ell_i$ is given by the span of $(x_i^- + x_i^+)$. For $i = 1$ this is obvious, the proof for $i > 1$ is left as an easy exercise.
The construction can also be motivated as follows. Let $x_1$ be the unit vector such that the sum of squares of the distances to $\mathcal{H}^-$ and $\mathcal{H}^+$ is minimal, i.e., such that it minimizes $\mathrm{d}(x_1, \mathcal{H}^-)^2 + \mathrm{d}(x_1, \mathcal{H}^+)^2$, where $d$ denotes the Euclidean distance. Determine $x_2, \ldots, x_n$ in the same way, under orthogonality conditions as before. Then $\ell_i = \mathrm{span}(x_i)$.
This construction of the state shows absolute preference for directions in the exact intersection. Hence for an exact time series it yields an exact state trajectory.

One of the main disadvantages of this procedure is that it treats every direction in the past and future in the same way, i.e., disregarding the corresponding energy content in the data. We define the energy of a direction $a^\top H$ as

$$\frac{\|a^\top H\|^2}{|a|^2}, \tag{5.28}$$

with $a$ a vector of appropriate length. If the Hankel matrices $H^-$ and $H^+$ are nearly singular their rowspan contains directions with almost zero energy. The canonical correlations approach, which only takes angles into account and not sizes, causes enormous amplification in these directions, with a disturbing effect on the identified model. One way to avoid nearly singular Hankel matrices is to choose a relatively small number $k$ of block rows in (5.26), but this limits the order of the resulting system and the global character of the procedure, as only relations over a small lag are considered.
In fact we want to determine those directions which are approximately common in past and future and which have, at the same time, a sufficient energy in both past and future. We propose the following somewhat rough but simple modification of the method. First we delete all directions in past and future corresponding to an energy below a certain threshold value $\theta$. More precisely, all singular values of the past and future Hankel matrices of the data below $\sqrt{\theta}$ are replaced by zero. Then we perform canonical correlation analysis on the remaining directions in the past and future. Simulations indicate that this method is indeed superior to direct canonical variable analysis, cf. Section 6.1 and a more extensive discussion in [33].

### 5.6.3 The Modified Canonical Correlation Algorithm

In the previous section we sketched a method for determining an approximate state trajectory from an observed time series. Here we describe how to translate this to a concrete algorithm for the construction of initial models for the

GTLS procedure. First we discuss the structural aspects of the algorithm, including the choice of the size of the Hankel matrices and the threshold value for the directional energy, and we conclude by an implementation of the procedure in terms of SVD.
Let the observation be denoted by $w : T \to \mathsf{R}^q$ with $T = [1, N]$. The procedure computes an SR of an approximate model for $w$ based on the approximate intersection of the matrices

$$\begin{aligned} H^- &:= \begin{bmatrix} w(1) & \dots & w(N-2k+1) \\ \vdots & & \vdots \\ w(k) & \dots & w(N-k) \end{bmatrix}, \\ H^+ &:= \begin{bmatrix} w(k+1) & \dots & w(N-k+1) \\ \vdots & & \vdots \\ w(2k) & \dots & w(N) \end{bmatrix}. \end{aligned} \tag{5.29}$$

This intersection is computed by a modified canonical correlation analysis with a threshold $\theta$ for the directional energy content in $H^-$ and $H^+$ as defined in (5.28).

**Initialization**
Reasonable values for the rank $m$ and degree $n$ of the approximate model have to be determined. Also the procedure parameters $k$ (the size of the Hankel matrices) and $\theta$ (the threshold for the directional energy) must be chosen. This can be done by computing models for different values of the parameters and comparing the resulting modelling errors. Various approaches are possible. For example, we can assume $m$ and $n$ as fixed, which determines the desired size of the model, or we can fix the $k$ and $\theta$, and then determine $n$ as the number of canonical correlations close to one and consider the results for several values of $m$.
It is difficult to give general rules which apply for all kinds of applications. We only give some rough guidelines based on our experience from simulations.

**Step 1. Construction of an Approximate State Trajectory**
Construct the Hankel matrices $H^-$ and $H^+$ as given in (5.29), with $k$ reasonably large. For an observation of length about 100 we typically take $k$ between 8 and 20. Determine a value for the threshold $\theta$, on the basis of the singular values of $H^-$ and $H^+$. This threshold corresponds to putting all singular values below $\sqrt{\theta}$ to zero, so $\theta$ should be chosen such that all 'nearly zero' singular values are cancelled. Let $H^-_\theta$ and $H^+_\theta$ denote the matrices obtained by removing low energy directions in this way in respectively $H^-$ and $H^+$.
If $n$ is not specified, determine $n$ as the number of small canonical angles. Let $(x_i^-, x_i^+), i = 1, \dots, n$ denote the first $n$ pairs of canonical variables, and let $L_i^-, L_i^+$ denote the corresponding canonical coefficients, so that $x_i^- = L_i^- H_\theta^-$ and $x_i^+ = L_i^+ H_\theta^+$. By definition the canonical variables are scaled to one, so

$$E_i := 1/(|L_i^-|^2 + |L_i^+|^2) \tag{5.30}$$

is a natural measure for the total directional energy in past and future, cf. (5.28). Define $\bar{x}_i = \mathrm{col}(\sqrt{E_i}(x_i^- + x_i^+)), i = 1, \ldots, n$ as an approximate state trajectory for $w$. The scaling by $E_i$ is imposed in order to make the norm of the state components in accordance with the data. Notice that $\bar{x}$ is defined for $t \in [k+1, N-k+1]$, which is the time axis for the upper row in $H^+$ in (5.29)

**Step 2. Determination of an Approximate System**
Construct the matrix

$$M := \begin{bmatrix} \bar{x}(k+1) & \ldots & \bar{x}(N-k) \\ \bar{x}(k+2) & \ldots & \bar{x}(N-k+1) \\ w(k+1) & \ldots & w(N-k) \end{bmatrix}, \tag{5.31}$$

and compute its SVD $M = USV^\top$. For an exact time series in a system of size $(m, n)$ as given in (5.23) the matrix $M$ would have rank $n + m$, which follows immediately from (5.24). Hence, if $m$ is not specified, $m$ is determined such that the last $n + q - m$ singular values are small compared to the first $n + m$ ones. We remark that in most of our simulations the only relevant value for $m$ is one. By putting the last $n+q-m$ singular values equal to zero we obtain an approximation $M_r$ of $M$ with rank $n+m$. In fact $M_r$ is the optimal rank $m+q$ approximation of $M$ for the static total least squares criterion, cf. Proposition 2.3.3. Now determine a matrix of the form $\begin{bmatrix} I_n & 0 \\ A & B \\ C & D \end{bmatrix} \in \mathsf{R}^{(2n+q)\times(n+m)}$ with image equal to $M_r$. Then $(A, B, C, D)$ is taken as an SR of an approximate model for $w$, which is motivated as follows. Let $\bar{x}'$ denote the first $n$ rows in $M_r$, which is the approximation of $\bar{x}$ in $M$. Assume that $\bar{x}'$ is non-singular. As $M_r$ is of rank $n+m$, there exists an $m$-dimensional sequence $\hat{v}(k+1), \ldots \hat{v}(N-k)$ such that $M_r = \begin{bmatrix} I_n & 0 \\ A & B \\ C & D \end{bmatrix} \begin{pmatrix} \bar{x}' \\ \hat{v} \end{pmatrix}$. So there exists a sequence $\hat{v}$ such that the equations induced by the SR hold approximately for the data $w$ and the state trajectory $\bar{x}$ constructed in the first step.

We conclude by an implementation of these steps in terms of the SVD, for given values of the procedure parameters $k$ and $\theta$ and desired size $(m, n)$ of the approximate system.

**Algorithm 5 (Modified Canonical Variables Method)**

**Data:**
- *An observation* $w : T \to \mathsf{R}^q$, $T = [1, N]$
- *Tolerated size* $(m, n)$
- *Values for the procedure parameters* $k$ *(the size of the Hankel matrices) and* $\theta$ *(the threshold for the directional energy)*

**Step 1:** *Determine the SVD's* $H^- = U_- S_- V_-^\top$ *and* $H^+ = U_+ S_+ V_+^\top$ *for* $H^-, H^+$ *given in (5.29).*
*Remove the columns in* $V_-, V_+$ *corresponding to singular values that are smaller or equal to* $\sqrt{\theta}$.
*Let* $USV^\top$ *denote the SVD of* $V_-^\top V_+$. *Define for* $i = 1, \ldots, n$, $E_i := 1/(\|U_i^\top S_-^{-1}\|^2 + \|V_i^\top S_+^{-1}\|^2)$, *according to (5.30), and* $\bar{x}_i := \sqrt{E_i}(U_i^\top V_-^\top + V_i^\top V_+^\top)$, *with* $U_i, V_i$ *equal to the* $i$*-th column of* $U, V$.

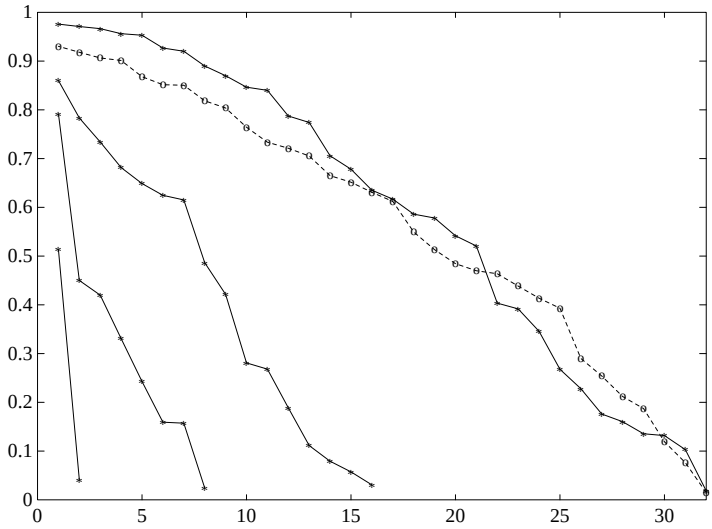
**Step 2:** *Determine the SVD* $M = USV^\top$ *for* $M$ *as defined in (5.31). Define* $M_r := U_r S_r V_r^\top$ *with* $U_r, V_r$ *the first* $n+m$ *columns of* $U, V$ *and* $S_r$ *the upper left square block of* $S$ *of size* $n+m$.
*Define* $K$ *as the first* $n$ *rows of* $U_r$, *and determine the SVD* $K = \bar{U}[\bar{S}\ 0]\bar{V}^\top$ *with* $\bar{S}$ *square. Define*
$Z := \bar{V} \begin{bmatrix} \bar{S}^{-1}\bar{U}^\top & 0 \\ 0 & I_m \end{bmatrix}$, *then* $U_r Z$ *takes the form* $\begin{bmatrix} I_n & 0 \\ A & B \\ C & D \end{bmatrix}$.

**Result:** $(A, B, C, D)$ *is an SR of the approximate model for* $w$ *as described by the modified canonical approach.*

*Proof.* See Appendix A.5.
*Implementation:* `Alg5` in Appendix B.2.

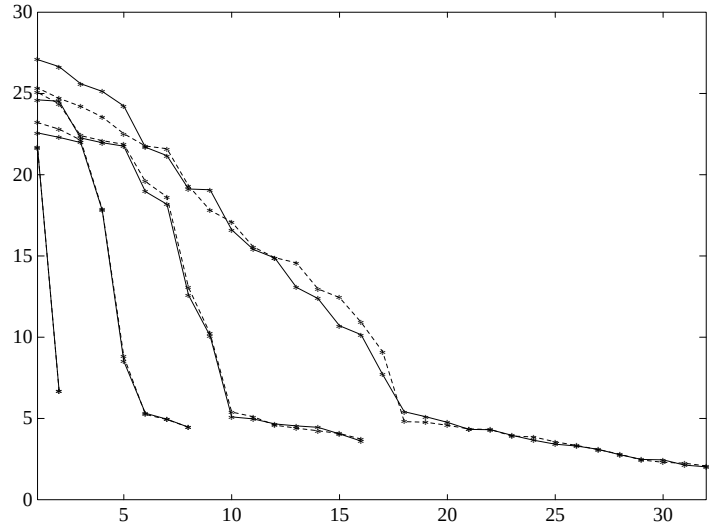
**Leading Example - Continued.** We determine the canonical correlations between past and future for data $w$ defined by (4.19). These are depicted in Fig. 5.2 for several sizes of the Hankel matrices in (5.29).
These correlations do not point unambiguously in the direction of a first order model, as there is no clear gap between the first and second correlation. The reason is that for a large number of blockrows $k$ in (5.29) coincidential relations occur between the past and future noise. As an illustration of this fact, we also depict the canonical correlations for the white noise component in the data in Fig. 5.2, for $k = 16$. This also shows that the relations between past and future corresponding to the highest canonical correlation may be affected by white noise.
Therefore we repeat the computation for an energy threshold $\theta = 100$, which means that all singular values in the Hankel matrices below 10 are removed. This value is chosen as there appears to be a 'tail' of relatively small singular values below 10 for all values of $k$, cf. Fig. 5.3. For comparison we remark that the total energy content for all directions in past and future Hankel matrices for the white noise component in the data is below 8.
The resulting matrices $H_\theta^-, H_\theta^+$ have dimension 1,4,9 and 16 for respectively $k$ = 1,4,8 and 16, which means the removal of respectively 1,4,7 and 16 low singular values. The canonical correlations for $\theta = 100$ are depicted in Fig. 5.4. This shows a larger gap between first and second correlation, so after removing the low energy directions the 'true' model order $n = 1$ becomes more

**Figure 5.2**: Canonical correlations between past and future in the simulated data of the leading example.




For the data $w = (u, y)$ generated according to (4.19), the canonical correlations between past and future (the rowspan of respectively $(H^-)$ and $(H^+)$ in (5.29)) are given for $k = 1, 4, 8, 16$. The number of correlations equals $2k$. The dashed line corresponds to the canonical correlation of the white noise component in the data for $k = 16$.

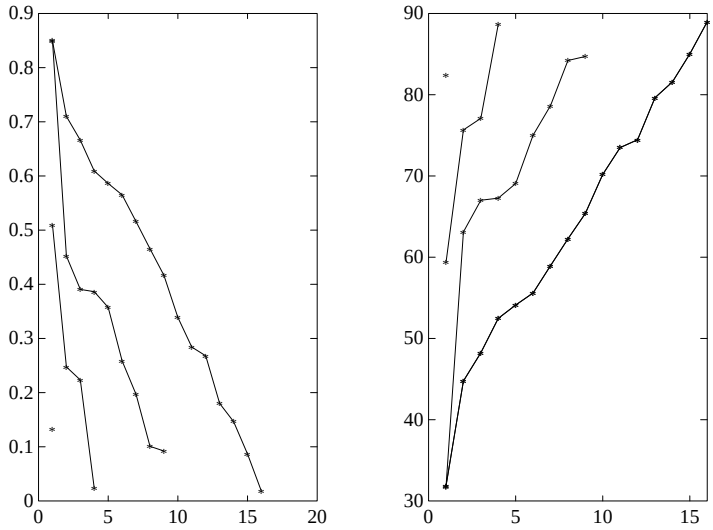
**Figure 5.3**: Singular values in past and future Hankel matrices for the leading example.




The singular values in past and future Hankel matrices (solid line and dashed line respectively) for $k = 1, 4, 8, 16$. The number of singular values equals $2k$.

clear. This is most obvious from the canonical angles, which are related with the correlations by $\cos\alpha_i = c_i$, with $\alpha_i$ the angles and $c_i$ the correlations.

**Figure 5.4**: Canonical correlations after removing small energy directions.




Left: As in Fig. 5.2, now after removing directions in past and future with energy threshold (5.28) below $\theta = 100$, i.e., with all singular values in $H^-, H^+$ in (5.29) below $\sqrt{\theta}$ put to zero. For $k = 1, 4, 8, 16$ there are resp. $1, 4, 9, 16$ directions left.
Right: the corresponding canonical angles (in degrees). The angles $\alpha_i$ are related to the correlations $c_i$ by $\alpha_i = \frac{180}{\pi}\arccos(c_i)$. The gap between first and second correlations or angles is now much clearer than in Fig. 5.2, even for $k = 16$.

In Table 5.4 we list the relative misfits for all constructed models.

**Table 5.4**: Relative misfits of initial models obtained by Algorithm 5.

| rel. misfit | $k = 1$ | $k = 4$ | $k = 8$ | $k = 16$ |
|---|---|---|---|---|
| $\theta = 0$ | 0.267 | 0.227 | 0.226 | 0.291 |
| $\theta = 100$ | 0.299 | 0.225 | 0.224 | 0.232 |

For the data generated by (4.19), the best first order initial model is obtained by applying Algorithm 5 with procedure parameters $k = 8$ (determining the size of the Hankel matrices in (5.29)) and $\theta = 100$ (denoting the energy threshold (5.28)).

This shows that the use of a threshold for the energy improves the results. For a more convincing illustration of this fact we refer to [33] and Section 6.1. $\diamond$

# Chapter 6

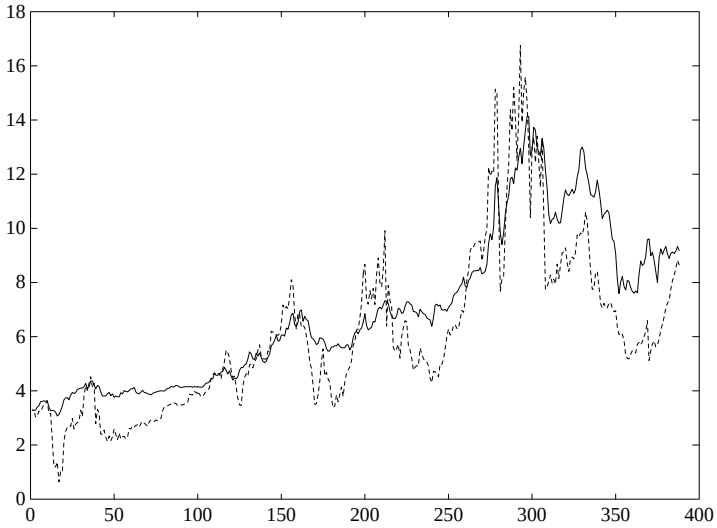
# Applications and Extensions

We illustrate the GTLS method by several applications. Our aim is to show what type of results can be obtained by this approach for real data as well as in simulation experiments, and to put this in contrast with some conventional methods.
The first example concerns long and short term interest rates in the United States. We determine linear, time-invariant relations with minimal misfit, and compare the results with those obtained by the local methods ARX and LTLS, described in Section 2.5.
As an extension to the GTLS method we consider the incorporation of constants and trends in models, in Section 6.2.
In Section 6.3 we extend the theory to periodic models, i.e., models corresponding to different relations in each period. As an illustration we consider seasonal models for the consumption and income in the German Federal Republic (GFR).
In order to illustrate the simultaneous character of the GTLS criterion for more than one system equation, we describe a simulation experiment with a system with two outputs, i.e., described by two equations. We also discuss the use of the recursive evaluation of the misfit, as described in Section 4.4, for detecting outliers in the data.
The fourth group of experiments shows the use of the GTLS method for model reduction, i.e., the approximation of a system by one of smaller size. We describe how to determine the optimal impulse response approximation for a system, and extend this result to optimal frequency-weighted approximation. We also show the use of GTLS for modelling input/output data.

## 6.1 US Interest Rates

In this section we analyse monthly observations of the long and short term interest rates in the United States from in January 1957 through April 1989. The short term interest rate is the 3-month US treasury bill rate, and the second series consists of interest rates with a maturity of 10 years. These series are depicted in Fig. 6.1.

**Figure 6.1**: Long and short term interest rates in the United States.




The interest rates (solid line: long term, dashed line: short term) in the US from Januari 1957 through April 1984.

Our goal is to gain insight in the joint behaviour of these interest rates. The ultimate goal of modelling would be an exact and complete explanation of the data in terms of underlying economic processes, involving e.g. the general condition of the economy and the strategies of banks and governments. However, this seems not feasible. In the GTLS approach we aim at a much more modest goal, namely to determine partial approximate relationships between the interest rates that are linear and time-invariant. We do not concentrate on underlying mechanisms, but we try to recognize patterns in the data that represent typical behaviour of the interest rates.

In this example we split the observation interval into two parts, the sample interval $T_s$ and the validation interval $T_v$, given by

$$\begin{array}{lll} \text{sample interval:} & T_s = [1, 200] & \text{(Jan 57 – Aug 73)} \\ \text{validation interval:} & T_v = [201, 388] & \text{(Sep 73 – Apr 89).} \end{array} \tag{6.1}$$

We write $w_s$ for the part of the data in the sample, and $w_v$ for the part out of sample.

We determine GTLS models for $w_s$, i.e., models of tolerated rank $m$ and degree $n$ for which we have to change the data by a minimal amount in order to make it satisfy the system laws throughout the whole sample interval. As the data consists of two components ($q = 2$), the only reasonable choice for the tolerated rank is one ($m = 1$), so we consider models described by one difference equation. Reasonable values for the model order are based on the canonical correlations of the data, as described in Section 5.6.

### 6.1.1 Initial Models

First we determine initial models from canonical correlation analysis of the data by Algorithm 5, which also gives an indication of a reasonable choice of the model order. After a first rough analysis we decided to apply the algorithm with $k = 8, 14, 20$ in (5.29), and for the threshold for the directional energy $\theta = 0$ and $\theta = 10$ (which is $0.001 * \|w_s\|^2$), cf. (5.28). In Table 6.1 some of the canonical correlations for $k = 20$ are listed. For $\theta = 0$ the number of rows in past and future Hankel matrices equals $2 * 20 = 40$, so there are 40 canonical correlations. For $\theta = 10$, 27 directions in past and future are removed, so there are 13 canonical correlations. We also list the energy contents corresponding to the canonical variables, as defined by (5.30).

**Table 6.1**: Canonical correlations between past and future of the US interest rates, with and without removing small energy directions.

| | $\theta = 0$ | | $\theta = 10$ | |
|---|---|---|---|---|
| | can. corr. | energy | can. corr | energy |
| 1 | 0.9999 | 1183 | 0.9995 | 11202 |
| 2 | 0.97 | 11 | 0.94 | 93 |
| 3 | 0.89 | 3.2 | 0.65 | 16 |
| 4 | 0.85 | 1.6 | 0.54 | 12 |
| 13 | | | 0.005 | 10 |
| 40 | 0.01 | 0.2 | – | – |

Canonical correlation between past and future (the rowspan of resp. $H^-$ and $H^+$ in (5.29) with $k = 20$). For $\theta = 0$ (ordinary canonical correlation analysis) there are 40 canonical correlations. For $\theta = 10$ there are 27 singular values below $\sqrt{10}$ put to zero, both in $H^-$ as in $H^+$, so there are 13 canonical correlations. The energy is a measure for the size of the data corresponding to the canonical variables, as defined by (5.30).

There are two correlations very close to one, indicating two nearly deterministic linear relations between past and future. For $\theta = 10$ the corresponding energy content in the data is clearly higher than for the ordinary canonical variables corresponding to $\theta = 0$. For $k = 8, 12$ the pattern is similar. This motivates to choose model order $n = 2$. For comparison we also determine models for $n = 0, 1, 3, 8$. The misfit of the optimal static model ($n = 0$) is computed by the SVD, as described in Proposition 2.3.3. For $n = 1, 2$ the procedure parameters $k = 14, \theta = 0$ give the lowest misfit, while for $n = 3, 8$ the values $k = 20, \theta = 10$ are optimal. In Table 6.2 we list the resulting relative misfits, i.e., the misfit divided by the Euclidian norm of the data.
This confirms our choice for a second order model, as for higher order the misfit does not decrease considerably.

### 6.1.2 GTLS Models

We use Algorithm 4 for the construction of GTLS models for tolerated rank $n = 1, 2, 3, 8$ with the initial models obtained as just described. We remark

**Table 6.2**: Relative misfits of initial models for the US interest rates.

| order | $n=0$ | $n=1$ | $n=2$ | $n=3$ | $n=8$ |
|---|---|---|---|---|---|
| relative misfit | 0.100 | 0.049 | 0.036 | 0.032 | 0.032 |

The relative misfits of initial models obtained by Algorithm 5 with procedure parameters $(k=14,\theta=0)$ for $n=1,2$, and $(k=20,\theta=10)$ for $n=3,4$.

that the incorporation of intercepts is discussed in Section 6.2.1. The iterations are stopped when the decrease of misfit has come below $10^{-5}$. The number of iterations needed to achieve convergence of the misfit at this level of precision varies from 5 (for $n=1$) to about 30 (for $n=2,3,8$). The resulting misfits are listed in Table 6.3.

**Table 6.3**: Relative misfits of GTLS models for the US interest rates.

| order | $n=0$ | $n=1$ | $n=2$ | $n=3$ | $n=8$ |
|---|---|---|---|---|---|
| relative GTLS misfit | 0.100 | 0.035 | 0.024 | 0.021 | 0.018 |

The models with these misfits are computed by the iterative Gauss-Newton algorithm (Algorithm 4), with initial models obtained by Algorithm 5 as described in Section 6.1.1.

So the GTLS algorithm gives an improvement with respect to the initial models of about 50%, cf. Table 6.2.
The GTLS models $\mathcal{B}^n_{US}$ for $n=0,1,2,3$ are given by the following equations for the long and short term interest rates $\ell, s$.

$$\begin{aligned}
\mathcal{B}^0_{US}:&\quad \ell(t) = 1.11 s(t)\\
\mathcal{B}^1_{US}:&\quad \ell(t) - 0.992\ell(t-1) = 0.28\{s(t) - 0.96 s(t-1)\}\\
\mathcal{B}^2_{US}:&\quad \ell(t) - 1.9660\ell(t-1) + 0.9665\ell(t-2) =\\
&\qquad\qquad 0.2787\{s(t) - 1.8910 s(t-1) + 0.8930 s(t-2)\}\\
\mathcal{B}^3_{US}:&\quad \ell(t) - 3.28\ell(t-1) + 3.56\ell(t-2) - 1.28\ell(t-3) =\\
&\qquad 0.11\{s(t) + 3.74 s(t-1) + 4.41 s(t-2) - 1.67 s(t-3)\}
\end{aligned} \tag{6.2}$$

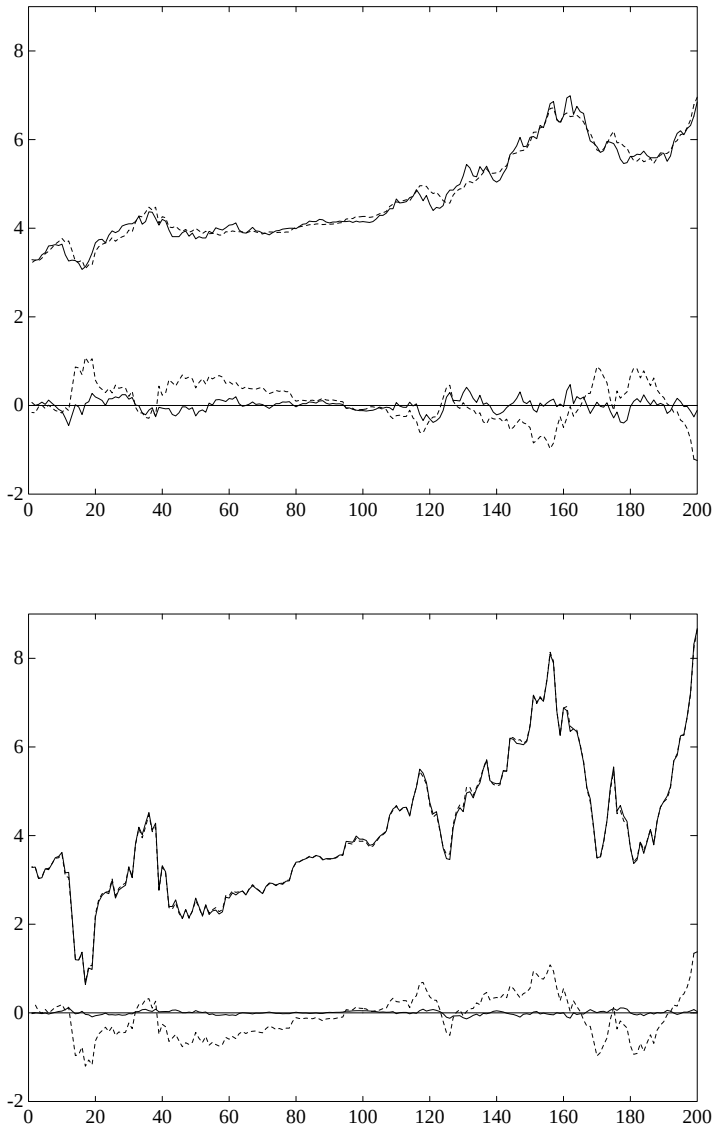
The static equation shows that during the sample period the return on loans over 10 years is about 11% higher than for short term loans. Notice the substantial decrease in misfit for the dynamic models, which indicates a structural dynamic relation.

**Remark on taking first differences.**
The first order model expresses a relation between the short term interest rates and the increase of the long term interest rates. The coefficient 0.992 is close to 1, which seems to indicate a 'unit root' in the long term interest rates. The question arises if this model is close to $\Delta\ell(t) = 0.28(s(t) - 0.96 s(t-1))$ with $\Delta$ the difference operator. Perhaps surprisingly, this is not at all the case: replacing 0.992 by 1 makes a lot of difference with respect to the *global* GTLS criterion. The model with a unit root has relative misfit 0.26, so it requires a change of 26% of the data to make it satisfy the equation with the unit root, which is even much more than for the static model. This can be explained as follows. Indeed 1 is close to 0.992, but $1^t$ and $0.992^t$ deviate more and more for increasing $t$, for instance, $0.992^{87} = 0.5$ (lag 87 corresponds to about 7 years). In the GTLS method this difference plays an important role, as also

all higher order implications of difference equations are taken into account. This indicates that the GTLS criterion is quite sensitive (hence informative) near unit roots.
This is in sharp contrast with e.g. the first step ahead prediction criterion, as this would only change the induced predictions by $(1-0.992)\ell(t-1)$, which is less than one percent of the size of the data.
Similarly, if we replace in the second order model the left hand side by $\Delta^2\ell(t)$, the relative misfit is 0.42. This means an increase of misfit by a factor 17, while it has a minor effect on the first step ahead predictions.
Even if we only change one root of the second order model to a unit root, the misfit increases dramatically. For instance, if we replace the left hand side in the second order model by $\Delta\ell(t) - 0.9663\Delta\ell(t-1)$, the relative misfit becomes 0.28, and replacing the righthandside by $0.2787(\Delta s(t) - 0.8920\Delta s(t-1))$ yields a relative misfit of 0.31.
These results should make us very careful in taking first differences of the data if our goal is to determine (long term) relations for the original series, even if the model contains 'nearly unit roots', as the results for the original and differences series are not compatible. A good model (in some sense) for the first differences need not induce a good model (in that sense) for the original series. A further analysis of the effect of differencing in GTLS modelling, e.g. by means of simulation studies, is left as a topic for further research. ◇

The list of misfits in Table 6.3 again motivates to choose $n = 2$ as a reasonable value for the model order. In the sequel we concentrate on this case. In Fig. 6.2 the corresponding optimal approximation is shown. For comparison we also depict the approximation error of the static model.
The optimal approximation in this figure satisfies the second order equation in (6.2) exactly, throughout the whole sample interval, and there is no time series closer to the data that satisfies a second order difference equation. As the approximation error is relatively small (about 2%), the data confirms this equation at each time instant, hence 198 times, modulo some unexplained details. Moreover, also all logical implications of this law are confirmed. For instance, all tenth order equations that can be derived from this law are confirmed 190 times, modulo the small approximation error. This is rather strong evidence for viewing this law as an intrinsic property of the interest rates, that also holds approximately outside the sample interval.
As an additional interpretation of the GTLS model $\mathcal{B}^2_{US}$, we consider its orthonormal basis induced by an ISR. Let $b : \mathsf{Z} \to \mathsf{R}^q$ denote the impulse response corresponding to an ISR of the GTLS system, i.e., the effect of an auxiliary input consisting of a pulse at $t = 1$. From the properties of ISR's it follows that $b$ and its shift form an orthonormal basis for $\mathcal{B}^2_{US}$, i.e., $b$ has unit norm, it is orthogonal to all its shifts $\sigma^k b$, and the span of $\{\sigma^k b; k \in \mathsf{Z}\}$ equals $\mathcal{B}^2_{US}$. This $b$ is depicted in Fig. 6.3, together with $b^\perp$, which is the orthonormal basis vector of the orthogonal complement of the model.
As the second order model has misfit 0.024, 98% of the data on $T_s$ consists of linear combinations of this basis vector and its shifts. Conversely, only 2%

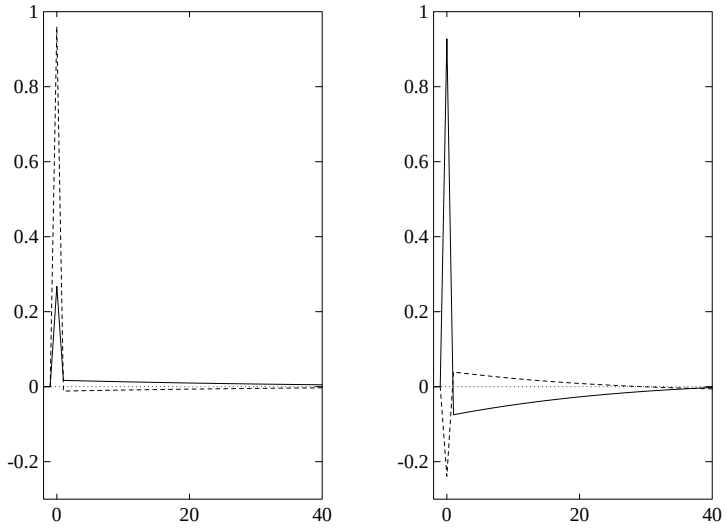
**Figure 6.2**: Optimal second order approximation of US interest rates on the sample interval.




The solid lines are the long term (upper plot) and short term (lower plot) monthly US interest rates on the sample interval (Jan 57 – Aug 73). The optimal approximation in the second order GTLS model $\mathcal{B}^2_{US}$ in (6.2) is given by the dashed lines in both plots (the approximation of the short term rates is so close that it is hardly visible). This approximation satisfies the model equation exactly, and there is no time series closer to the data that satisfies a second order equation. For comparison we also depict the corresponding approximation error and the error of the optimal static model $\mathcal{B}^0_{US}$ (solid and dashed lines around zero in both plots).

of the data is composed of linear combinations of $b^\perp$. Hence we may consider $b$ as a 'stylized' fact that is dominant throughout the complete observation interval, and $b^\perp$ as nearly absent.

**Figure 6.3**: Stylized fact in the US interest rates and its orthogonal complement.

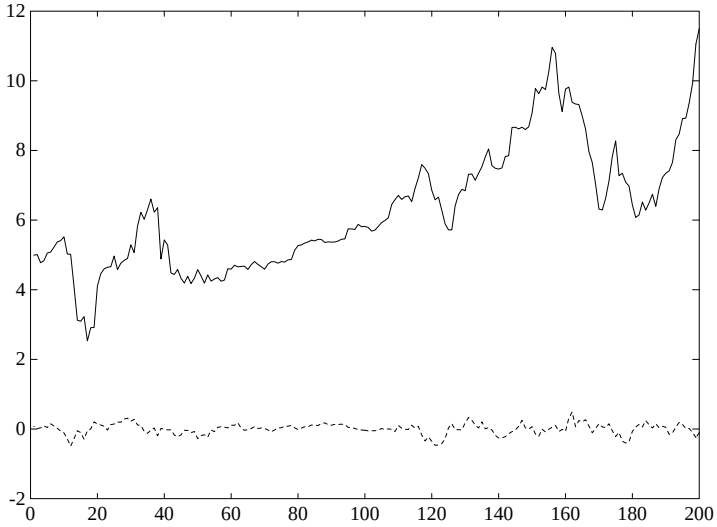



Left: basis vector for $\mathcal{B}^2_{US}$ (solid line: long term interest, dashed line: short term interest). This vector and its shifts form an orthonormal basis for $\mathcal{B}^2_{US}$. About 98% of the interest rates on the sample interval (6.1) consists of linear combinations of this basis vector and its shifts (restricted to the sample interval).
Right: basis vector for the orthogonal complement of $\mathcal{B}^2_{US}$. This type of behaviour is almost completely absent in the data.

The typical behaviour consists of a local shock in the short term interest rates, combined with a relatively small change in the long term rates with a positive after effect. This reflects the fact that the long term interest rate is less volatile, as it reacts more slowly on external effects.
The coefficients of the data with respect to the basis $b, b^\perp$ equal respectively the auxiliary input for the approximation and the error, cf. (4.15), which is depicted in Fig. 6.4.
The auxiliary input for the approximation is quite close to the short term interest rate, although we treated it similar to the long term interest rates. This motivates to consider the long term as the effect of the short term.

**Remark on cointegration analysis.**
The aim of cointegration analysis is to determine *cointegration relations* for the data, i.e. linear combinations of non-stationary components that yield a stationary series. This topic is nowadays part of every modern textbook on econometrics. A typical outcome of cointegration analysis for long and short term interest rates is that both series individually are viewed as random walks, while the difference between both is considered as a stationary process. This means that there is a common non-stationary part in both interest rates.

**Figure 6.4**: Representation of US interest rates with respect to the the orthonormal basis of the second order GTLS model.




The solid line is the auxiliary input of the optimal second order GTLS approximation on the sample interval (6.1) depicted in Fig. 6.2 in an ISR of $\mathcal{B}^2_{US}$ in (6.2). Equivalently, these are the coefficients of the approximation with respect to the orthonormal basis for $\mathcal{B}^2_{US}$ depicted in the Fig. 6.3 (left). The dashed line depicts the coefficients of the approximation error with respect to the orthonormal basis for the orthogonal complement of $\mathcal{B}^2_{US}$ depicted in Fig. 6.3 (right).

There are some similarities with the results of GTLS modelling. The approximation $\hat{w}_s$ on the sample interval might be considered as a 'non-stationary' part in the interest rates. The individual components of $\hat{w}_s$ are closely related, as they are both the effect of one common 'non-stationary' (hidden) factor, namely the auxiliary input of the approximation depicted in Fig. 6.4. The approximation error is comparable to a 'stationary' part, as it is the effect of the rather stationary looking factor depicted as the dashed line in Fig. 6.4.

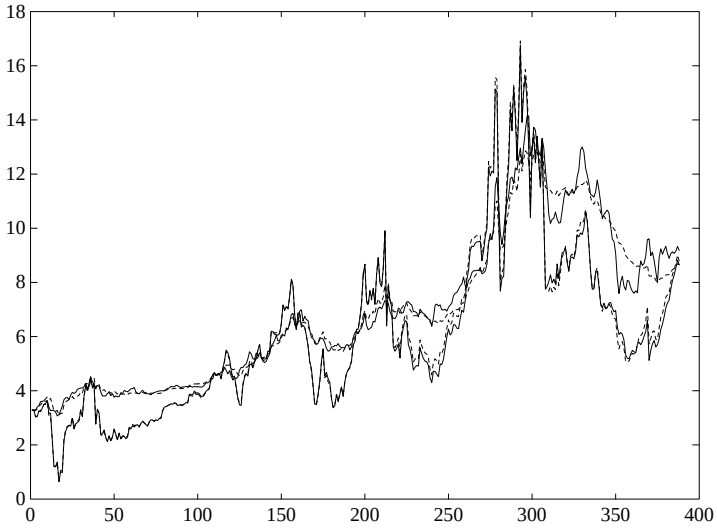
More generally, if we relate 'large' to 'non-stationary' and 'small' to 'stationary', the GTLS method leads to the decomposition of the data into a stationary and non-stationary part, like cointegration analysis. However, the non-stationary parts in each component are not litteraly 'common' (so that they would cancel in some linear (static) combination of components, as in cointegration analysis), but they are *dynamically* related. ◇

### 6.1.3 Validation

We use the remainder of the data for validating the GTLS model. The question is, whether the GTLS model is in good correspondence with the behaviour of the interest rates during the next 188 quarters after the sample interval. Stated otherwise, we have ruled out those time series that deviate considerably from the solution set of the equation as unrealistic behaviour of the interest rates,

and we check whether this is true on the validation interval. We also consider the validation on [201 : 260] separately, excluding the relatively volatile behaviour during the last part of the observation interval.
We first determine the optimal approximation, within $\mathcal{B}^2_{US}$ in (6.2), on the complete observation interval $T = [1, 388]$, which is depicted in Fig. 6.5.

**Figure 6.5**: Validation of the second order GTLSmodel for the US interest rates.




The solid lines are the long and short term monthly US interest rates on the complete observation interval (Jan 57 – April 89), also depicted in Fig. 6.1. The dashed lines depict its optimal approximation in the second order GTLS model based on the data on the sample interval (– Aug 73).

We see that the misfit remains small at the validation interval, although it is somewhat higher than on the sample interval. The increase in misfit can be due to a general increase of volatility of the interest rates, or to a change of the regularity towards a different second order equation. In the first case the GTLS model still would be nearly optimal, while in the latter case there would be a substantially better second order model. Therefore we also determine the GTLS model for the data on the validation interval. This is given by

$$\begin{aligned} &\ell(t) - 1.9274\ell(t-1) + 0.9286\ell(t-2) = \\ &\quad 0.4050(s(t) - 1.8808s(t-1) + 0.8842s(t-2)) \end{aligned} \tag{6.3}$$

In Table 6.4 we compare the relative misfit of the GTLS model for $w_v$ (the data on the validation interval) and that of $w_s$ (the data on the sample interval). This shows that the largest part of the increase of the misfit on the validation interval is due to an increase of volatility, as the minimal relative misfit of the second order GTLS model based on $w_v$ (0.042) is only about 15% lower than of the GTLS model based on $w_s$. In the first part of the validation interval [201 : 260], the relative misfit has hardly increased. As the misfit of the model $\mathcal{B}^2_{US}$ remains low at the validation interval, and is comparable to the misfit of

**Table 6.4**: Validation of second order GTLS model for the US interest rates in terms of relative misfits.

| time interval | on [1,200] | on [201:260] | on [201:388] |
|---|---|---|---|
| GTLS model on [1,200] | 0.0242 | 0.0259 | 0.0483 |
| GTLS model on [201:388] | 0.0322 | 0.0298 | 0.0418 |

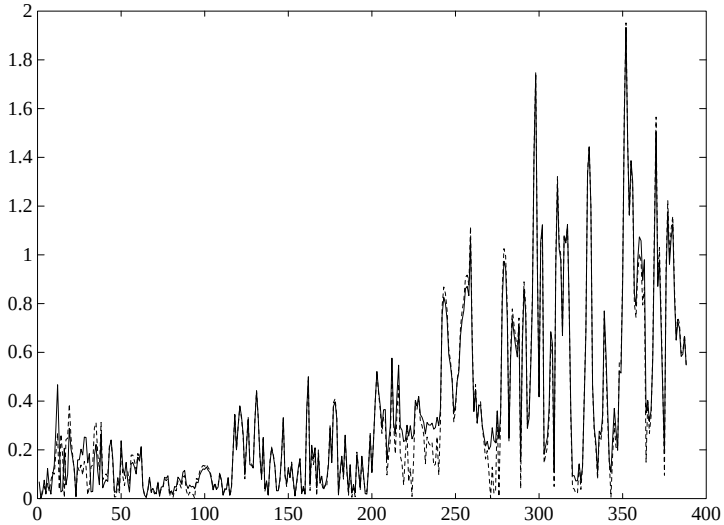
The second order GTLS model for the first 200 observations ($\mathcal{B}^2_{US}$) has relative misfit 0.0259 for the next 60 observations. The second order GTLS model for the data on the validation interval has relative misfit 0.0418 on that interval.

the optimal model for the validation data, we conclude that the system law in (6.2) is indeed confirmed on the validation interval, although the size of deviations has increased.

### 6.1.4 Analysis of Outliers

We investigate whether the misfit of obtained models is due to so-called outliers, i.e., exceptionally high deviations from the model restrictions at only a few time instants. The recursive projection Algorithm 2 in Section 4.4 gives insight in how the misfit is affected by each subsequent observation, given the values in the past. An outlier corresponds to an exceptionally high increase of the misfit at a certain time instant. In Fig. 6.6 we compare the sequence $m(t)$, denoting the square root of the increase of the squared misfit due to the observation $w(t)$, given its past, and $|\tilde{w}(t)|$, which is the misfit at time $t$ given the whole observation, as computed just before.

**Figure 6.6**: Analysis of outliers in the US interest rates.




The solid line is the size of the approximation error of the second order GTLS model $\mathcal{B}^2_{US}$ in (6.2) at each time instant ($|w(t) - \hat{w}_s(t)|$). The dashed line is the square root of the increase of the squared misfit of $\mathcal{B}^2_{US}$ due to observation $w(t)$ given the past observations ($m(t)$).

The increase of misfit due to an new observation given the past, corresponding to $m(t)$, has no exceptionally high peaks, although on average the peaks increase at the end of the interval. We conclude that there is no severe outlier, but a structural increase of volatility.
Notice that the recursive misfit sequence $m(t)$ is close to the size of the misfit $\|\tilde{w}\|$. This means that the misfit at time $t$ is for the largest part determined by past values of the interest rates, and hardly affected by their future. Hence we might say that the future confirms the past-induces approximations.

### 6.1.5 Comparison with Local Methods

As a further illustration of the global character of the GTLS approach we compare the results with those obtained by the local methods ARX and LTLS, which are described in Section 2.5. For $w_s$ the second order ARX model is given by

$$\begin{aligned}\ell(t) - 1.0030\ell(t-1) + 0.0124\ell(t-2) = \\ 0.1422\{s(t) - 0.4899s(t-1) - 0.4168s(t-2)\} + \varepsilon(t),\end{aligned} \tag{6.4}$$

with $\varepsilon$ the first step ahead prediction error. Here we have chosen $s$ as exogenous variable, and $\ell$ endogenous, in accordance with our findings in Section 6.1.2, cf. Fig. 6.3, and also in correspondence with the so-called *expectation hypothesis* on interest rates, see e.g. [37]. The LTLS method yields for $w_s$

$$\begin{aligned}\ell(t) - 2.0108\ell(t-1) + 1.0109\ell(t-2) \approx \\ 0.0960(s(t) - 1.8234s(t-1) + 0.8249s(t-2)),\end{aligned} \tag{6.5}$$

with the interpretation that this equation is locally optimally fitting, i.e., the equation holds for each time *separately* for minimal changes of all involved variables, cf. (2.18). The misfits of these models on the sample and validation interval are given in Table 6.5.

**Table 6.5**: Misfit for local methods ARX and LTLS compared with the misfit of the GTLS model.

| relative misfit | on sample | on validation |
|---|---|---|
| ARX | 0.037 | 0.084 |
| LTLS | 0.132 | 0.171 |
| GTLS | 0.024 | 0.048 |

It takes a change of the data of 3.7% to make the US interest rates satisfy the second order ARX equation (6.4) exactly (i.e. with $\varepsilon(t) = 0$) on the sample interval (6.1). It takes a change of 17.1% of the data to make it satisfy the LTLS equation (6.5) on the validation interval. The GTLS model defeats both models on the sample interval (by definition of GTLS), and also on the validation interval.

This clearly shows that the system laws obtained by the local methods have less quality with respect to the global GTLS criterion.

To do justice to the ARX method, we also compare the first step ahead prediction errors for the long term interest rates, which is the criterion on which the ARX model is based. In Table 6.6 we list the first step ahead prediction errors for the long term interest rates induced by the obtained models. On the sample interval this is by definition the smallest for the ARX model, and also on the validation interval ARX defeats the other methods, albeit less pronounced. This shows that the ARX and GTLS models contain different information on the interest rates. A comparison of their quality will depend on a preference for local or global criteria.

**Table 6.6**: First step ahead prediction error for the long term interest rate for the methods ARX, LTLS and GTLS.

| rel. prediction error | on sample | on validation |
|---|---|---|
| ARX | 1.60 | 4.63 |
| LTLS | 2.21 | 5.54 |
| GTLS | 2.47 | 5.66 |

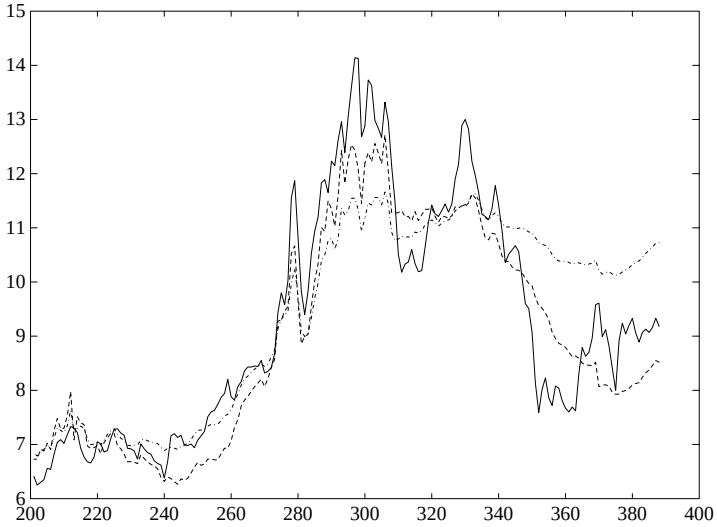
The LTLS model (6.5) induces a first step ahead prediction error for the long term interest rates on $[3, 200]$ with norm 1.60. The GTLS model $\mathcal{B}^2_{US}$ in (6.2) induces prediction errors on $[203 : 388]$ of norm 5.66. The ARX model defeats both models on the sample interval (by definition of ARX), and also on the validation interval, although less pronounced.

For further comparison we also consider the predictions over a long interval induced by these models. We determine the predictions for the long term interest rates on the validation interval (September ′73 - April ′89), from the model based on the sample interval (January ′57 - August ′73), the data on this interval, and the short term interest rates on the validation interval. The results are depicted in Fig. 6.7.
Notice that the computation of the $k$-step ahead prediction involves the $(k+1)$-th order implications of the second order difference equation in (6.2), so for instance the prediction of $\ell(388)$ involves the 189-th order implication. In the GTLS method we take account of these implications for all $k$, whereas the local methods only consider them for $k = 1$. This difference is reflected in the quality of the predictions, as especially at the end of the prediction interval the GTLS method turns out to give better results.

## 6.2 Constants and Trends

As an extension of the GTLS method we consider the handling of constants, also called *intercepts*. Instead of the restriction to linear models, it is sometimes desirable to impose a weaker condition that allows for adding a constant to each component of a time series in the model behaviour. This might be relevant, for example, after taking the logarithm of measurements, as then constants correspond to the scaling of the original series. Then we replace the condition of linearity 2.1.3 by

**Figure 6.7**: Long term predictions of long term interest rates from short term interest rates.




The solid line is the long term interest rates on the validation interval (Sep 73 – April 89). The dashed line is the prediction induced by the GTLS model $\mathcal{B}^2_{US}$ in (6.2), the dahs/dotted line the prediction induced by the ARX model (6.4). Predictions are based on past (long and short term) interest rates on the sample interval (Jan 57 – Aug 73), and the short term interest rates on the validation interval.

**Definition 6.2.1 (Linearity modulo Constants)** *A model $\mathcal{B} \in (\mathsf{R}^q)^{\mathsf{Z}}$ is called linear modulo constants if there exists a constant time series $c$ such that $\{w - c;\ w \in \mathcal{B}\}$ is linear.*

The misfit and size of such models are defined in the obvious way. The corresponding extension of the GTLS problem can be formulated as

**Definition 6.2.2 (GTLS with Constants)**
*Given:*

- *an observation $w : T \to \mathsf{R}^q$, with $T = [1, N] \subset \mathsf{Z}$*
- *tolerated size $(m, n)$,*

*determine:*

- *a decomposition $w = \hat{w}^* + c + \tilde{w}^*$ with $\hat{w}^* \in \mathcal{B}^{q,m,n}_T$, $c$ a constant time series, and $\|\tilde{w}^*\|$ minimal.*

The estimation of the optimal constants in $c$ can be incorporated in a straightforward way in the iterative estimation procedures for the standard GTLS problem. Let $\mathcal{C}$ denote the set of all constant time series $\{w : \mathsf{Z} \to \mathsf{R}^q;\ w(t) = c \text{ for some } c \in \mathsf{R}^q\}$.

**Proposition 6.2.3 (GTLS with Constants)** *The GTLS problem with constants can be solved by the following modifications of the GTLS algorithms:*

- *In Algorithm 1, replace in the second step $w(t)$ by $w(t)-c$, and determine $x(N+1)$ and $c$ such that $\|\tilde{v}\|$ is minimal.*
- *In Algorithm 3, replace $\mathcal{E}$ and $\mathcal{F}$ by $\mathcal{E}' := \mathcal{E} + \mathcal{C}$ and $\mathcal{F}' := \mathcal{F} + \mathcal{C}$.*
- *In Algorithm 4, add to $\mathcal{T}$ also the effect of constant time series on $\bar{v}$ in (5.21).*

*Proof.* See Appendix A.6.

In addition to constants, it is also easy to incorporate trends, i.e., straight lines $c+dt$ with $c, d$ constants in $\mathsf{R}^q$ in the modelling of time series. Many time series in economics show an exponential growth, which becomes a linear trend after taking logarithms. Then one might first remove this trend in order to obtain a time series that looks more 'stationary', so that it can be modelled as the realization of a stationary stochastic process. We remark that in our approach it is not compulsory first to transform the data to a 'stationary' series, as non-stationary behaviour may be absorbed by the degrees of freedom in time-invariant models, cf. our discussion in Section 6.1.2. The estimation of optimal trends can be handled as before, just by replacing $c$ by $c + dt$ in the previous exposition.
In the same way quadratic trends ($c + dt + et^2$) can be handled. In fact the incorporation of a linear set of deterministic components leads to a linear extension of the spaces $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{T}$ in the algorithms. Besides trends to arbitrary order, this also concerns all kinds of dummy variables, but we will not pay attention to this issue.

### 6.2.1 Constants and Trends for US Interest Rates

It is not a prerequisite of the GTLS method that the observation should consist of a 'stationary' time series, i.e. with the same type of behaviour in all components during the observation interval. Although GTLS models are time-invariant, non-stationary behaviour can be incorporated in the model as there is a degree of freedom at each time-instant. This is apparent in the results obtained for the US interest rates in the previous sections.
Nevertheless there may be reasons to incorporate constants and trends in models. If we allow for constants, the dynamic part of optimal models is not affected by the level of the data components. For instance, then it makes no difference whether we analyse the interest rates themselves, or the corresponding multiplication factor given by the interest rate plus 100%, which is also a reasonable starting point.
The effect of constants and trends on the misfit is shown in Table 6.7. The models are estimated on the sample interval consisting of the first 200 observations.
The models are given by the equations

**Table 6.7**: Effect of constants and trends on the relative misfit of second order GTLS models.

| rel. misfit | on [1:200] | on [201:260] | on [201:388] |
|---|---|---|---|
| no const, no trend | 0.0242 | 0.0259 | 0.0483 |
| const., no trend | 0.0232 | 0.0266 | 0.0575 |
| const. and trend | 0.0228 | 0.0225 | 0.1147 |

The first line is also depicted in Table 6.4. Incorporating a constant term (and trend) decreases the relative misfit on the sample interval from 0.0242 to 0.0232 (0.0228), but this leads to an increase of the relative misfit on the validation interval from 0.0483 to 0.0575 (0.1147).

$$\begin{array}{ll} \mathcal{B}^2_{US}: & \ell(t) - 1.9660\ell(t-1) + 0.9665\ell(t-2) = \\ & 0.2787(s(t) - 1.8910s(t-1) + 0.8930s(t-2)) \\ constants: & \ell(t) - 1.9695\ell(t-1) + 0.9713\ell(t-2) = \\ & 0.3029(s(t) - 1.9340s(t-1) + 0.9384s(t-2)) + 0.0027 \\ const\&trend: & \ell(t) - 2.0042\ell(t-1) + 1.0058\ell(t-2) = \\ & 0.2714(s(t) - 1.9723s(t-1) + 0.9732s(t-2)) + \\ & +0.0041 + 2.02\ 10^{-5}t \end{array} \tag{6.6}$$

We conclude that the incorporation of constants and trends is not effective in explaining the data, as it hardly decreases the misfit and, moreover, even leads to a higher misfit on the validation interval.
Notice that the left hand side of the last model equation almost equals the second difference of $\ell(t)$, as in (6.2). However, if we replace it by the second difference, the relative misfit increases to 0.0398. Imposing only one unit root by replacing the left hand side by $\Delta\ell(t) - 1.005\Delta(\ell(t)$ yields relative misfit 0.0354. This shows again that replacing nearly unit roots by exact unit roots may lead to a relatively large increase of misfit, although the effect is less dramatic than in Section 6.1.2.

## 6.3 Periodic GTLS

Many phenomena are known to exhibit periodic behaviour, and then it is appropriate to take account of that in modelling. In such situations the condition of time-invariance (2.1.2) may be too severe, as this does not allow for periodically changing model restrictions.
A straightforward way to weaken shift-invariance to periodicity is as follows.

**Definition 6.3.1 (Periodic Systems)** *A system $\mathcal{P} \in (\mathsf{R}^q)^{\mathsf{Z}}$ is called periodic of period $\pi$ if it satisfies $\sigma^\pi\mathcal{P} = \mathcal{P}$.*

Stated otherwise, models of period $\pi$ do not change under shifting the time axis over $\pi$ time-instants. Of particular practical interest are the cases $\pi = 4$ and $\pi = 12$ for respectively quarterly and monthly data. Quarterly and monthly models have been studied extensively in econometrics, see e.g. [14, 28]. The differences between periodic GTLS and other methods for periodic modelling are the same as for the time-invariant case: its global character

and the symmetric treatment of all data components, cf. Section 2.5, and further the absence of stochastics and its formulation without reference to model parameters.
Most definitions and results for the time-invariant case can be extended to the periodic case in the obvious way. The behaviour of linear periodic models on a finite interval $T$ consists of a linear subspace of $(\mathsf{R}^q)^T$. As for time-invariant models, we define the rank of periodic models in terms of the increase of dimension at each time, and the degree in terms of the number of initial conditions at each time, which now may depend on the period.

**Definition 6.3.2 (Periodic Size, Rank, Degree)** *The size of a linear periodic system $\mathcal{P}$ of period $\pi$ is defined as the pair of $\pi$-tuples $(m_1, \ldots, m_\pi; n_1, \ldots, n_\pi)$ such that with $m_{k\pi+i} := m_i$ and $n_{k\pi+i} := n_i$, $\dim(\mathcal{P}_{[t_0,t_1]}) = n_{t_0} + \Sigma_{t=t_0}^{t_1} m_t$ for $t_1 - t_0 \geq \max\{n_1, \ldots, n_\pi\}$. The tuples $(m_1, \ldots, m_\pi)$ and $(n_1, \ldots, n_\pi)$ are called respectively the periodic rank and degree of $\mathcal{P}$.*

For a proof of correctness of this definition we refer to Appendix A.6.
The sizes of systems are partially ordered by comparing the size at each period, i.e.,

$$(m_1, \ldots, m_p; n_1, \ldots, n_p) \leq (m'_1, \ldots, m'_p; n'_1, \ldots, n'_p) \text{ if } m_k \leq m'_k \text{ and } n_k \leq n'_k \qquad (6.7)$$

for all $k = 1, \ldots, p$.
We remark that for a time invariant system of size $(m, n)$, the periodic size is just given by $(m, \ldots, m; n, \ldots, n)$.
The misfit of linear periodic models is defined exactly as for the time-invariant case. The periodic GTLS problem consists of the construction of a periodic linear model of tolerated size with minimal misfit with respect to an observed time series.
Periodic SR's (PSR's) are state representations in which the system matrices depend on the period. Periodic ISR's (PISR's) are PSR's that satisfy (3.7) at each period. Periodic difference equations have coefficients that depend on the period. We summarize some results on the representation of periodic systems in the following proposition.

**Proposition 6.3.3 (Periodic Systems)** *Let $\mathcal{P}$ denote a periodic linear complete system. Then*

1. *$\mathcal{P}$ is the solution set of a set of periodic linear difference equations of finite lag.*

2. *$\mathcal{P}$ can be described by a periodic state representation, and its periodic rank and degree equal the minimal number of auxiliary variables and states for each period.*

3. *(conjecture) If $\mathcal{P}$ is stabilizable, it admits a periodic isometric state representation.*

*Proof.* See Appendix A.6 The proof of 3. is not complete, so it is a conjecture.

The relation between the periodic rank of a system and its representation in terms of periodic difference equations is straightforward. The periodic rank $m_k$ equals $q$ minus the number of independent equations that apply at period $k$, completely analogous to the time-invariant case. However, the periodic degree does not correspond to the minimal total order of these equations, but is a mixture of the total orders of the equations for all periods. We explain this by the following example.

**Example.** Consider a linear model in two variables $(u, y)$ of period two, that is described by

$$\{(u, y) : \mathsf{Z} \to \mathsf{R}^2;\ y_t = u_{t-1} \text{ if } t \text{ odd and } y_t = u_{t-100} \text{ if } t \text{ is even}\} \tag{6.8}$$

In the long run $\dim(\mathcal{B}_{[1,t]}) = t + 50$, as we can freely choose values for $u$ and for $y_1, y_2, y_4, \ldots, y_{96}, y_{98}$ (this determines $y_{100} = y_1$). Similarly, for large $t$ $\dim(\mathcal{B}_{[2,t+1]}) = t + 50$, as besides $u$ the values of $y$ at the even time instants on $[1, 100]$ are free. This implies that the rank of the model equals $(m_1, m_2) = (1, 1)$, and the degree $(n_1, n_2) = (50, 50)$. Notice that both at odd and even time-instants the system should somehow remember the last 50 values of $u$ at the even time-instances.
This also illustrates that not all combinations of periodic degrees can occur. Consider a linear model of period two consisting of time series of two variables, so its rank is a pair of integers $(m_1, m_2)$, and clearly $0 \leq m_k \leq 2$. Let us assume that $m_1 = m_2 = 1$, as in the previous example, and moreover that now $n_1 = 1$. Then it is impossible that e.g. $n_2 = 100$, as this would mean that at even time instances the number of degrees of freedom due to initial conditions is 100, while at the next odd time instant this is suddenly reduced to one, which clearly is impossible. In fact, it is easily derived that in this example $0 \leq n_2 \leq 2$.
More generally, the periodic degree in two subsequent periods cannot differ more than the number of system variables, i.e.,

$$|n_k - n_{k+1}| \leq q, k = 1, \ldots, \pi, \tag{6.9}$$

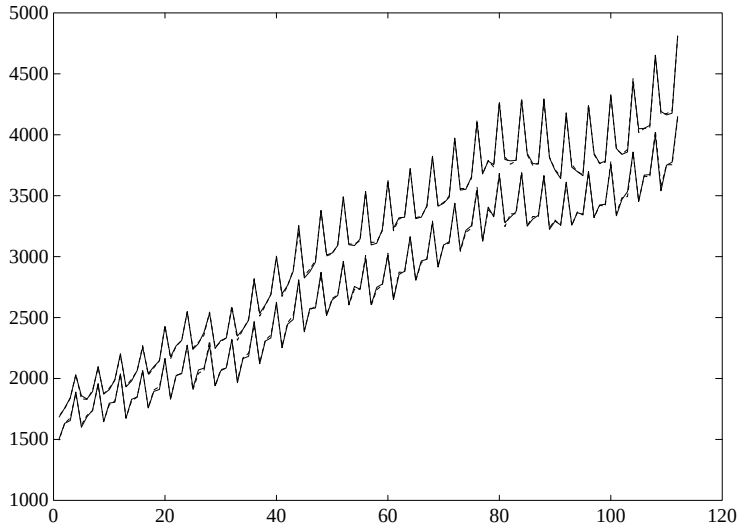
as the state at time $t$ is a linear function of the previous state and the system trajectory at time $t - 1$, which follows from the definition of SR's. $\diamond$

Once it is known how to determine periodic ISR's, the construction of periodic GTLS models requires only some minor modifications of the algorithms. In all algorithms we replace SR's by PSR's and ISR's by PISR's. Further, the projections onto linear spaces is decomposed into $p$ projections for each period, which gives the optimal variations of the periodic system matrices.

### 6.3.1 GFR Seasonal Consumption and Income

We illustrate the periodic GTLS method by applying it on the seasonal consumption and income in the German Federal Republic (GFR) during the years 1960–1987, cf. [25, Table E.4]. In Fig. 6.8 the data is depicted, together with the optimal approximation in the periodic first order GTLS model.

**Figure 6.8**: Seasonal consumption and income in West Germany and their first order seasonal GTLS approximation.




Seasonal West German real per capita personal disposable income (upper solid line) and personal consumptions expenditures (lowed solid line) from 1960 through 1987. The optimal approximation in the first order seasonal GTLS model for the first 100 observations is given by the dashed lines (hardly visible as they are very close).

In Table 6.8 we compare the misfit of periodic models with those of time-invariant models for order zero and one. The models are based on the data over $[1, 100]$, and the misfit is computed over the complete observation interval.

**Table 6.8**: Misfits of time-invariant and periodic GTLS models for GFR seasonal consumption and income.

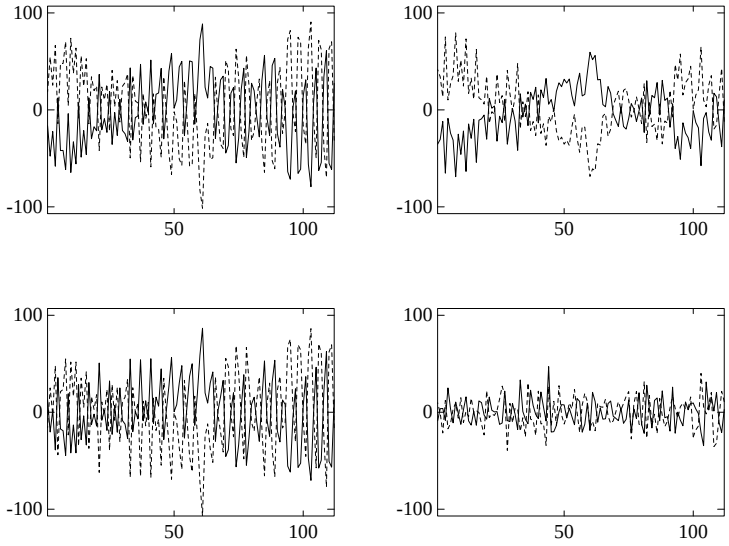
| misfit | static | first order |
|---|---|---|
| time-invariant | 609 | 564 |
| periodic | 417 | 225 |

The optimal first order time-invariant GTLS model for the first 100 observations has misfit 564 (on the complete observation interval $[1, 112]$). For the first order periodic model this is 225, which equals the norm of the difference between the data and its optimal approximation in the system depicted by resp. the solid and dashed lines in Fig. 6.8.

The approximation errors are depicted in Fig. 6.9. Notice that the approximation error for the time-invariant models clearly shows a seasonal pattern,

while this is less apparent for the periodic models.

**Figure 6.9**: Approximation errors of time-invariant and periodic GTLS models for GFR seasonal consumption and income.




Approximation error in income (solid lines) and consumption (dashed lines) for static time-invariant and seasonal model (upper plots) and first order time-invariant and periodic models (lower plots).

The models are given by the following equations, where $c$ denotes consumption and $i$ income.

$$\begin{array}{lrl} n=0 & \text{time-inv.} & c_t = 0.87 i_t \\ & I & c_t = 0.85 i_t \\ & II & c_t = 0.90 i_t \\ & III & c_t = 0.89 i_t \\ & IV & c_t = 0.87 i_t \\ n=1 & \text{time-inv.} & c_t - 0.983 c_{t-1} = 0.81(i_t - 0.981 i_{t-1}) \\ & I & c_t - 0.32 c_{t-1} = 2.06(i_t - 0.66 i_{t-1}) \\ & II & c_t - 1.14 c_{t-1} = 0.94(i_t - 1.08 i_{t-1}) \\ & III & c_t - 1.05 c_{t-1} = 0.55(i_t - 1.08 i_{t-1}) \\ & IV & c_t - 1.40 c_{t-1} = -0.27(i_t - 0.18 i_{t-1}) \end{array} \tag{6.10}$$

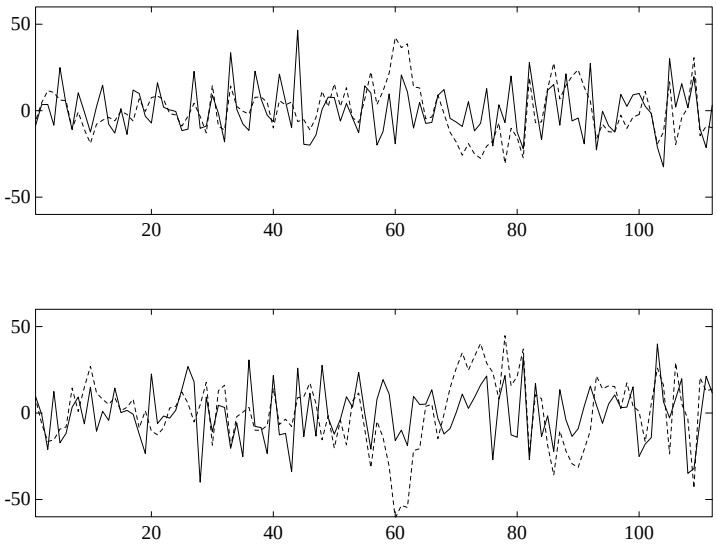
This indicates that the consumption and income are related by a linear equation that depends on the period. However, the periodic model class covers the time-invariant class of the same order as a special case, so the misfit for *every* time series with respect to its periodic GTLS model is lower than with respect to its time-invariant GTLS model of the same order. Therefore we also compare the periodic model with a higher order time-invariant one. The total number of parameters in the time-invariant model class of rank $m$ and degree $n$, $\mathsf{B}^{q,m,n}$, equals $nq + mp$, and each system in this model class has dimension $mN + n$ with $N$ the length of the observation interval, cf. Section 5.1. For seasonal models with all periodic ranks equal to $m$ and periodic degrees equal to $n$ the model class contains $4(nq + mp)$ parameters, and each periodic model in this class has dimension $mN + n$. For $N = 112$ the total number of parameters for the first order periodic case equals $12 + 113 = 125$. For the

time-invariant case we take $n = 4$, which corresponds also to $9 + 116 = 125$ parameters. The resulting system for the first 100 observations is given by

$$\begin{aligned} c_t + 0.046c_{t-1} - 0.011c_{t-2} + 0.013c_{t-3} - 0.932c_{t-4} = \\ 0.71(i_t + 0.038i_{t-1} - 0.006i_{t-2} + 0.049i_{t-3} - 0.933i_{t-4}) \end{aligned} \tag{6.11}$$

and has misfit 254 over the complete observation interval. The approximation errors are compared in Fig. 6.10. We remark that the peak around $t = 60$ (1974) might be explained as the effect of the oil crisis.

**Figure 6.10**: Comparison between first order seasonal and fourth order time-invariant GTLS model for GFR seasonal consumption and income.




In order to make the data satisfy the first order periodic equation in (6.10) we have to substract from the data the approximation error corresponding to the solid lines, with norm 225 (see Table 6.8). In order to make it satisfy the fourth order time-invariant equation (6.11) we have to substract the approximation error given by the dashed lines (with norm 254). (The relatively high deviation around $t = 60$ (1974) might be explained as an effect of the oil crisis.)

It turns out that the periodic first order model is only slighly more accurate than the fourth order time-invariant one. This shows that, although both components of the data clearly show a seasonal behaviour, they nevertheless satisfy a time-invariant relation rather accurately. An explanation could be that the periodicity in consumption and income separately is due to seasonal external effects that hardly affects the dynamic relationship between both quantities. We leave a thorough comparison between the periodic and time-invariant models as a topic for further reasearch.

## 6.4 Multiple Outputs

In this example we consider a system with multiple outputs, so that a single difference equation does not suffice to describe the system. Our aim is to

show how GTLS handles simultaneous equations. For simplicity we consider a system with one input and two outputs. The data are generated as $w = w' + e$, where $w' \in \mathcal{B}^{3,1,2}$ satisfies the equations

$$\begin{aligned} w_2'(t) &= w_2'(t-1) + w_1'(t) \\ w_3'(t) &= w_2'(t) + w_1'(t-1) \end{aligned} \tag{6.12}$$

For $w_1$ we take white noise with unit variance, and for $e$ a three-dimensional white noise process with independent components and variance 0.01. The relative size of the noise is given by $\|e\|/\|w\| = 0.20$. The observation interval has length 50. The GTLS model of rank one and degree two is described by

$$\begin{aligned} w_2(t) &= 0.95w_2(t-1) - 0.04w_2(t-2) + 1.08w_1(t) + 0.07w_1(t-1) + 0.15w_1(t-2) \\ w_3(t) &= 0.95w_3(t-1) - 0.04w_3(t-2) + 1.12w_1(t) + 1.08w_1(t-1) - 0.89w_1(t-2) \end{aligned} \tag{6.13}$$

and has relative misfit 0.15. In fact both GTLS algorithms yield an ISR of this model with two state variables and one auxiliary input. We translated this into difference equations by using Proposition 3.4.3 and transforming the resulting input/state/output representation into difference equations by state elimination (which is a standard procedure in Matlab). We remark that the equations are not uniquely defined by the model, as every pair of independent linear combinations of these equations represent the same system.

The GTLS model consists of all time series that satisfy both equations. So the data satisfies these equation *simultaneously*, throughout the whole observation interval, after a modification of the size of 15% of the data. It takes a larger change to make the data satisfy any other set of two independent equations with total lag 2 that is not logically equivalent to (6.13).

In contrast to for instance Vector Auto Regression (VAR), cf. e.g. [25, 27], in which each equation corresponds to a prediction for a specific component, the equations in (6.13) play a symmetric role, and may be replaced by an arbitrary couple of logically equivalent equations.

The non-uniqueness of parameters is often called an 'identification problem', but in our approach the term 'representation issue' seems more appropriate: we do not estimate parameters in equations, but use equations to represent estimated models which are defined as sets of time series. In fact we use ISR's in our algorithms, for which the multi-equation case (rank $m$ less than $q-1$, with $q$ the number of components) can be treated exactly like the single-equation case ($m = q - 1$).

Another issue in simulation experiments is the relation between the estimated model and the 'data generating process', which we only briefly discuss. For a first comparison we transform the model equations (6.12) to

$$\begin{aligned} w_2'(t) - w_2'(t-1) &= w_1'(t) \\ w_3'(t) - w_3'(t-1) &= w_1'(t) + w_1'(t-1) - w_1'(t-2), \end{aligned} \tag{6.14}$$

which are similar to the estimated equations. However, a comparison at this level has several drawbacks. Firstly, the equations may be replaced by an

arbitrary couple of logically equivalent equations, as this represents exactly the same system. Secondly, it involves arbitrariness in the sense that any other *type* of representation may be chosen as well and probably leads to different results.
It would be more in the spirit of the GTLS approach to compare both systems as sets of time series, independent of representations. Notice that the systems are somehow close together, as the data belongs to both approximately. A more thorough analysis might for example concern the computation of the canonical angles between both models on the observation interval. We leave this issue as a topic for further research.
In order to reconstruct the size of the model from the data, we compare the misfits of GTLS models of various complexity in Table 6.9.

**Table 6.9**: Selection of rank and degree in a simultaneous equations model.

| rel. misfit | $n=0$ | $n=1$ | $n=2$ | $n=3$ | $n=4$ |
|---|---|---|---|---|---|
| $m=1$ | 0.55 | 0.24 | 0.15 | 0.15 | 0.14 |
| $m=2$ | 0.21 | 0.10 | 0.091 | 0.085 | 0.075 |

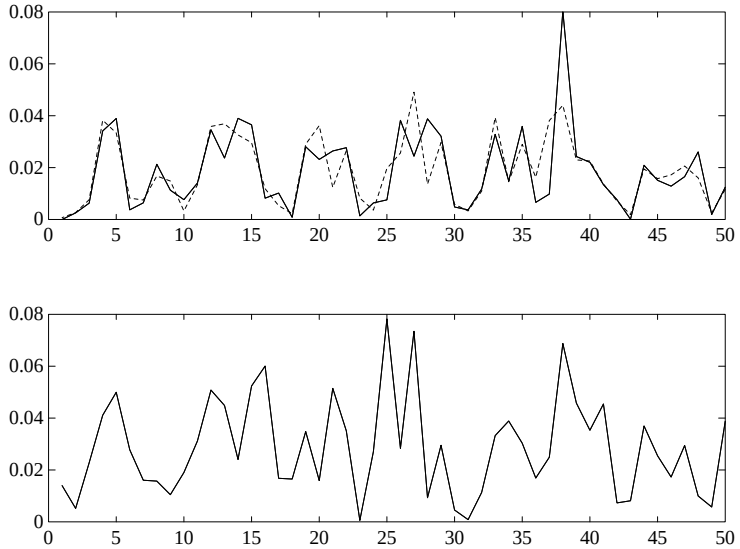
It takes a change of 14% of the data to make it satisfy two independent linear difference equations with constant coefficients and sum of degree at most 4 ($m=1, n=4$). To make it satisfy one such an equation of order one ($m=2, n=1$) requires a change of 10%.

For rank one, the misfit hardly decreases for orders above two. This could be expected, as the regular part of the data belongs to a system of order two. For rank two the results suggest to take the order one. Comparing the complexities $(m,n) = (1,2)$ and $(2,1)$, the first one of course leads to a larger misfit, as it imposes more restrictions. However, the relative misfit is still small, which motivates to choose $(m,n) = (1,2)$.

### 6.4.1 Recursive Evaluation of the Misfit

The recursive projection Algorithm 2 in Section 4.4 gives insight in how the misfit is affected by each subsequent observation, given the values in the past, cf. also Section 6.1.4. In this section we consider its application on the simulation data of the preceding section, consisting of three components. In Fig. 6.11 we compare the sequence $m(t)^2$, denoting the increase of the squared misfit due to the observation $w(t)$, given its past, and $|\tilde{w}(t)|^2$, which is the squared misfit at time $t$ given the whole observation, as computed just before. For comparison we also plot the size of the noise $e$ in the simulated data, $|e(t)|^2$. This illustrates that the recursive procedure can be helpful for locating exceptional deviations in the data. For example, consider the results around $t = 37$. The peak in $m(t)^2$ at $t = 38$ clearly shows a relatively high deviation in the data at that point, while this is not obvious from the misfit sequence $|\tilde{w}(t)|^2$, as this shows a moderate error for both $w(37)$ and $w(38)$. So, although it appears to be optimal to distribute the deviation in $w(38)$ over several time-instances

**Figure 6.11**: Recursive evaluation of the misfit in a simultaneous equations model.




The dashed line in the upper plot is the squared approximation error of the model given by (6.13) at each time instant ($|\tilde{w}(t)|^2$). The solid line is the increase of squared misfit due to observation $w(t)$, given the past observations ($m(t)^2$). This indicates $w(37)$ as an outlier. This is confirmed by the lower plot, which depicts the size of the noise in the data generating process ($|e(t)|^2$). The noise around $t = 25$ is also large in absolute value, but it turns out to be in less sharp conflict with the model laws in (6.13).

afterwards, the recursive procedure clearly points out its exact location. Indeed, it turns out that the noise $e$ in the simulated data has a peak at $t = 38$. However, notice that not all peaks in the noise lead to a high increase of the misfit, as they may be in correspondence with the degree of freedom in the system at each time-instant, due to the fact that there are only two independent equations for three components.

As a further illustration of Algorithm 2 we give an interpretation of the matrix $\tilde{H}_t$, which measures the effect of 'innovations' $\epsilon$ on the misfit, cf. (4.25). We consider the steady state value $\tilde{H} := \lim_{t\to\infty} \tilde{H}_t$. The singular value decomposition (cf. Definition 2.3.2) of $\tilde{H}$ is given by

$$\tilde{H} = \begin{bmatrix} 0.45 & -0.68 & 0.58 \\ -0.81 & -0.06 & 0.58 \\ 0.35 & 0.73 & 0.58 \end{bmatrix} \begin{bmatrix} 0.97 & 0 & 0 \\ 0 & 0.56 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.45 & -0.68 & 0.58 \\ -0.81 & -0.06 & 0.58 \\ 0.35 & 0.73 & 0.58 \end{bmatrix}^\top \tag{6.15}$$

This gives a picture of the effect of an observation on the misfit, given the past. If $w(t)$ happens to be equal to $C\hat{x}_t(t)$, then $\epsilon(t) := w(t) - C\hat{x}_t(t) = 0$, so the misfit does not increase. Moreover, $w(t)$ may vary in the direction of the kernel of $\tilde{H}$ without any cost. This means that given the observations up to time $t$, the model leaves the level of the next observation completely free: an observation $w(t) + [c\ c\ c]^\top$ leads to the same increase of misfit for all $c \in \mathsf{R}$.

This is in accordance with (6.12), as the equations remain valid on $[1, t]$ if the same constant is added to all three components of $w(t)$. If $\epsilon(t)$ is not of the form $[c\ c\ c]^\top$ the misfit is increased. The ratio between the nonzero singular values gives the relative severeness of deviations in the direction of the first and second singular vectors of $H$ with respect to the GTLS criterion.

## 6.5 Optimal Model Reduction

In this section we show how the GTLS approach can be used for model reduction, i.e., the approximation of a system by a smaller one. This amounts to the reduction of the rank and/or degree of a system, cf. Definition 2.1.6, but we only consider the reduction of the degree. The basic idea is to determine a GTLS system of reduced order for a system trajectory in the original system. We concentrate on the case that this system trajectory consists of a systems impulse response. We extend the results in Section 6.5.2 to optimal frequency weighted model reduction. Finally we show that the algorithm can handle noncausal systems without any additional difficulty.

### 6.5.1 Impulse Response Approximation

We consider the single input, single output system $\mathcal{B}$ with poles in $\pm 0.9i$ and $-0.7 \pm 0.6i$ and gain 0.5, so

$$\begin{aligned}\mathcal{B} = \{(u, y) : \mathsf{Z} \to \mathsf{R}^2;\ y(t) = 0.5u(t) - 1.4y(t-1) + \\ -1.66y(t-2) - 1.13y(t-3) - 0.69y(t-4).\}\end{aligned} \tag{6.16}$$

This system has one input, and state dimension $n = 4$, so its complexity is $(1, 4)$. We consider reduction to complexity $(1, 2)$. The data $w \in \mathcal{B}$ consists of two components $u$ and $y$, where $u$ is a unit pulse at time $t = 0$ and $y$ is the corresponding response. The aim is to determine a second order system with impulse response $y'$ as close as possible to $y$, i.e. with $\|y - y'\|$ minimal. As $y$ and $y'$ are infinite sequences, it is conventional to call this the $l_2$-norm instead of the Euclidian norm, and therefore we refer to this as the $l_2$-optimal impulse response approximation problem, cf. [38] and the references therein.
First we apply the GTLS algorithm, starting in a randomly chosen model. When the decrease in the misfit has become sufficiently small, below $10^{-10}$, the iterations are stopped. Notice that this is a naive approach to $l_2$-optimal impulse response approximation, as we also allow for adaptation of the input. The final model $\mathcal{B}_{gtls}$ is compared in Table 6.10 with the balanced reduction $\mathcal{B}_{bal}$ and the Hankel norm reduction $\mathcal{B}_{hank}$, that have been developed especially for model reduction, see [15] and [26].
The error in the impulse response in $\mathcal{B}_{gtls}$ is somewhat larger than that in $\mathcal{B}_{bal}$ and $\mathcal{B}_{hank}$. If one is interested in this response then one should prevent an approximation of the input, so that an optimal approximation of the output becomes the criterion. This is achieved by taking the norm $\|\tilde{w}\|_\alpha^2 := \alpha^2\|\tilde{u}\|^2 + \|\tilde{y}\|^2$ with $\alpha$ sufficiently large. It is easily verified that this corresponds to

**Table 6.10**: GTLS compared with balanced reduction and Hankel norm approximation.

| approximation criteria | $\mathcal{B}_{gtls}$ | $\mathcal{B}_{bal}$ | $\mathcal{B}_{hank}$ |
|---|---|---|---|
| misfit with respect to $w$ | 0.32 | 0.36 | 0.38 |
| error in impulse response | 0.54 | 0.46 | 0.57 |
| Hankel norm distance | 0.93 | 0.80 | 0.74 |

It takes a change with norm 0.32 in the impulse and its response to make it satisfy a second order linear difference equation with constant coefficients. It takes a change with norm 0.57 to make the observed impulse response $y$ equal to the impulse response of $\mathcal{B}_{hank}$.

the unweighted GTLS problem for data $(\alpha u, y)$. The effect of increasing $\alpha$ is given in Table 6.11. This shows that for large $\alpha$ the method determines better approximations of the impulse response.

**Table 6.11**: Effect of scaling in GTLS.

| scaling factor | $\alpha = 1$ | $\alpha = 10$ | $\alpha = 100$ |
|---|---|---|---|
| $\|.\|_\alpha$-misfit with respect to $w$ | 0.3204 | 0.4457 | 0.4476 |
| error in impulse response | 0.5407 | 0.4477 | 0.4476 |

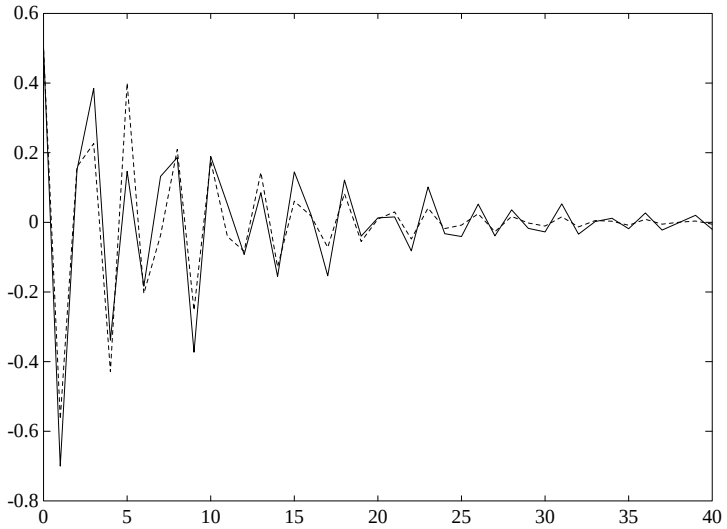
$\alpha = 1$ corresponds to unweighted GTLS, cf. Table 6.10. The minimal value of $\alpha^2\|\tilde{u}\|^2 + \|\tilde{y}\|^2$ for which $(\hat{u}, \hat{y}) := (u - \tilde{u}, y - \tilde{y})$ belongs to a second order linear time-invariant system equals 0.4457 for $\alpha = 10$. The error in impulse response for the corresponding system equals 0.4477.

This also gives bounds for the minimally achievable error in the impulse response, which we denote by $e^*$. Let $\mathcal{B}_\alpha$ be the GTLS model for $\|.\|_\alpha$, and let $y_\alpha$ be the impulse response of $\mathcal{B}_\alpha$, then it is easily checked that $\mathrm{d}(w, \mathcal{B}_\alpha) \leq e^* \leq \|y - y_\alpha\|$. By increasing $\alpha$ we can obtain an arbitrarily accurate estimate of $e^*$. This gives a solution method for the least squares optimal impulse response approximation problem. For $\alpha = 100$ we obtain $e^* = 0.4476$, see Table 6.11. The corresponding model is given by the equation $y(t) = -1.35y(t-1) - 0.78y(t-2) + 0.50u(t) + 0.11u(t-1) - 0.21u(t-2)$. The impulse response is depicted in Fig. 6.12.

### 6.5.2 Frequency Weighting

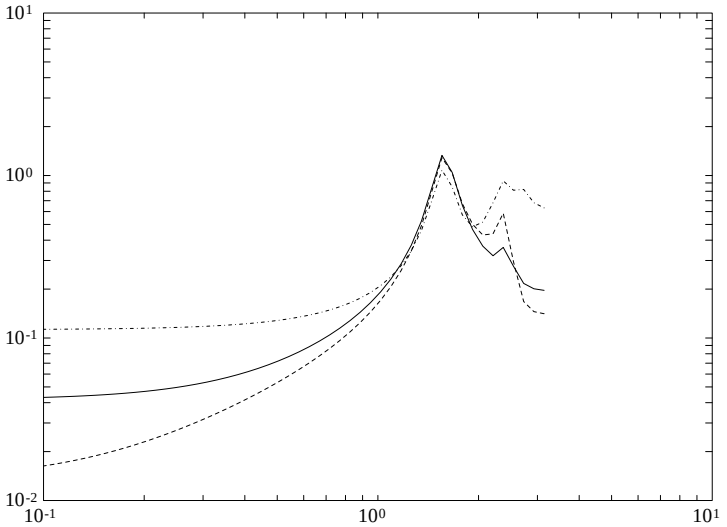
In the previous section we described how to determine the optimal approximation of an impulse response by the impulse response of a lower order system. One could question the quality of the result as an approximation of the whole system, and not only of the impulse response. Therefore we consider the difference in outputs for all frequencies, i.e., we compare the size of the effect of sinusoidal inputs on the outputs in the original system with their effect in the reduced system, for all frequencies. This is shown in Fig. 6.13, which is called the error magnitude response.

**Figure 6.12**: Second order $l_2$-optimal impulse response approximation of a fourth order system.




The solid line is the impulse response of the fourth order system $\mathcal{B}$ in (6.16). The dashed line is its $l_2$-optimal approximation of second order.

**Figure 6.13**: GTLS compared to balanced reduction and Hankel norm approximation in terms of magnitude responses.




Magnitude response of the difference between the system $\mathcal{B}$ in (6.16) and its approximations, obtained by GTLS (solid line), balancing (dashed line), and Hankel approximation (dash/dotted line).

The $l_2$-optimal approximation of the impulse response corresponds to the minimization of the squared error magnitude response, integrated over all frequencies in $[0, \pi]$, i.e.

$$\int_0^\pi |G(e^{i\omega}) - G^{reduced}(e^{i\omega})|^2 d\omega, \tag{6.17}$$

where $G, G^{reduced}$ denote the transfer function of the system of full and reduced order.
This criterion has the property that the error for all frequencies is weighted equally. This is reasonable for instance if the input consists of white noise, which is composed of all frequencies in equal quantities. However, if one expects dominant frequencies in the input, it may be desirable to emphasize those frequencies in the criterion. Then we should not take our starting point in the impulse response of the system, but in a system trajectory in which the input has a typical frequency distribution. This is motivated by the fact that the $l_2$-optimal approximation of the output for an input with frequency content $U(\omega) := \Sigma_{t\in T} u(t)e^{i\omega t}$, minimizes the criterion

$$\int_0^\pi |U(\omega)|^2 |G(e^{i\omega}) - G^{reduced}(e^{i\omega}|)^2 d\omega, \tag{6.18}$$

So the frequency weights in the input are precisely the frequency weights in the criterion that is minimized. This result applies to infinite sequences with finite $l_2$-norm. For finite series there is a minor boundary problem, as a part of the effect of the inputs falls outside the observation interval. Nevertheless, the idea of optimal frequency weighted approximation is still valid for the finite time case. In the GTLS approach we do not assume the inputs and outputs to be zero after the observation interval, but estimate an optimal final state, which decreases the 'finite time effect'.
As an example we consider two frequency weighted approximations. First we try to improve the approximation at low frequencies. This means an emphasis on long term dynamics in contrast to local changes, which correspond to high frequencies. Therefore we take as input $u_{low}$ the effect of white noise in a low-pass filter,

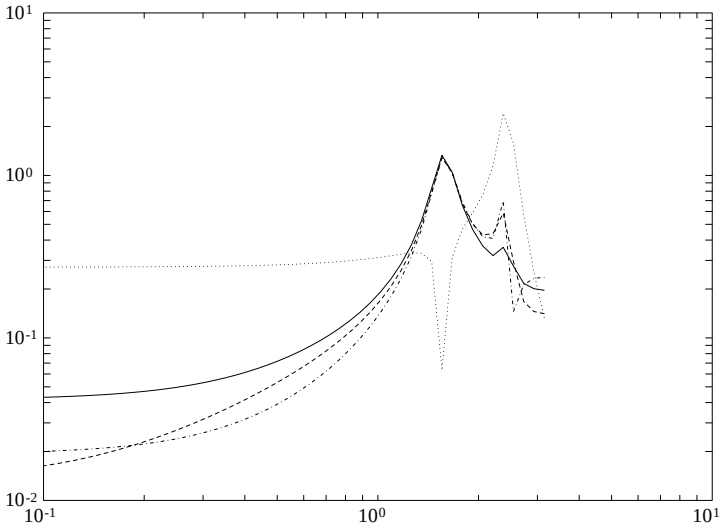
$$u_{low} := \frac{1}{1 - 0.9\sigma}\varepsilon, \tag{6.19}$$

of length 128. We determine the optimal approximation of the corresponding output, as described in Section 6.5.1 with $\alpha = 100$. The error magnitude response, depicted in Fig. 6.14, shows that indeed the accuracy for low frequencies has increased, at the cost of a slightly worse approximation at higher frequencies.
As a second example we consider an input with extreme high energy content for frequency $\pi/2$, in order to improve the results around that frequency. Therefore we take as input

$$u_{\pi/2}(t) = sin(\pi t/2) + \epsilon(t) \tag{6.20}$$

with $\epsilon$ white noise with variance 0.005. The white noise is added to avoid the degenerate case of approximating a system in only one frequency. Fig. 6.14 shows that indeed the approximation around $\pi/2$ is quite accurate, although at high costs for other frequencies.

**Figure 6.14**: Error Magnitude for frequency-weighted $l_2$-optimal approximations.




Magnitude response of the difference between the system $\mathcal{B}$ in (6.16) and its approximations with emphasis on low frequency, by taking input (6.19) (dash/dotted line), and on the frequency $\pi/2$, by taking input (6.20) (dotted line). For comparison the results of unweighted $l_2$-optimal impulse response approximation (solid line) and balanced reduction (dashed line) are shown again.

### 6.5.3 Noncausal Systems

Suppose we are interested in detecting rapid changes in an observed signal. In order to mitigate the effects of small irregularities, a smoothed version of the signal is obtained by convolution with the filter[1] $\phi(x) = (5/\sqrt{2\pi})\, e^{-x^2}$. Rapid changes of a signal $w_1$ will be detected by the second derivative of this smoothed version, $w_2$, which is related to $w_1$ by the 'Mexican hat',

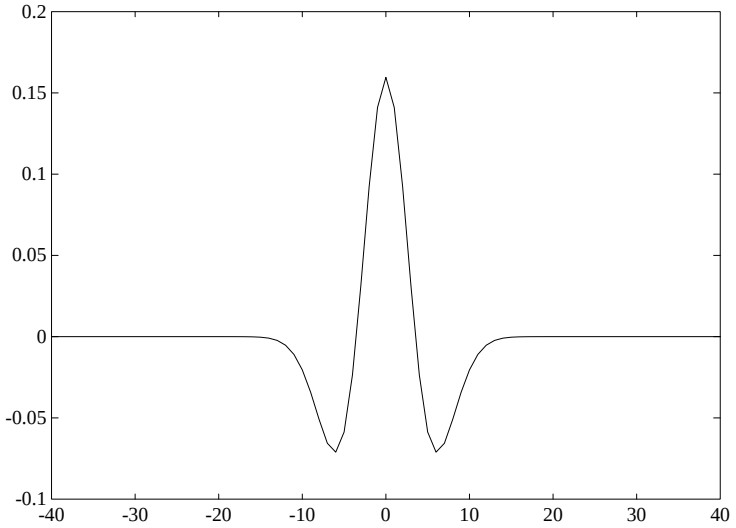
$$w_2(t) = -\frac{d^2}{dt^2}\left\{\int_{-\infty}^{\infty} \varphi(x) w_1(t-x)dx\right\}. \tag{6.21}$$

In the simulations we consider a discrete time version $w_2(t) = \Sigma_{j=-N}^{N} G_j w_1(t-j)$ with $N = 40$ and time steps of size 0.2, so $G_j$ is the steplength 0.2 times the second derivative of $\varphi$ in $0.2j$. Note that $w_2$ is not a causal output, as the value of $w_2(t)$ depends on the future values of $w_1$.
First we apply Algorithm 4 to the impulse response observation, i.e., $w_1$ is a unit pulse at time $t = 0$ and $w_2$ is the corresponding response, see Fig. 6.15.

[1]For comparison we mention that the normal density is given by $(2\pi)^{-\frac{1}{2}}\, e^{-\frac{1}{2}x^2}$.

Initial models for the iterations are determined by Algorithm 5, as described extensively in [33].

**Figure 6.15**: The Mexican hat.




This is the effect of a pulse $w_1$ on $w_2$ in the system (6.21).

The misfits of the optimal models of orders 2, 4 and 6 are given in Table 6.12. They are compared with the optimal Hankel norm approximations of orders 2, 4, and 6. These are obtained by approximations of orders 1,2 and 3 of the causal part of the impulse response, and using the symmetry of the Mexican hat to estimate the anticausal part. Analogously we determined approximations by balanced reduction. In Table 6.12 we also list the error in the impulse response of these models, i.e., the Euclidian distance between the systems impulse response and the Mexican hat $w_2$, with norm $\|w_2\| = 0.35$.

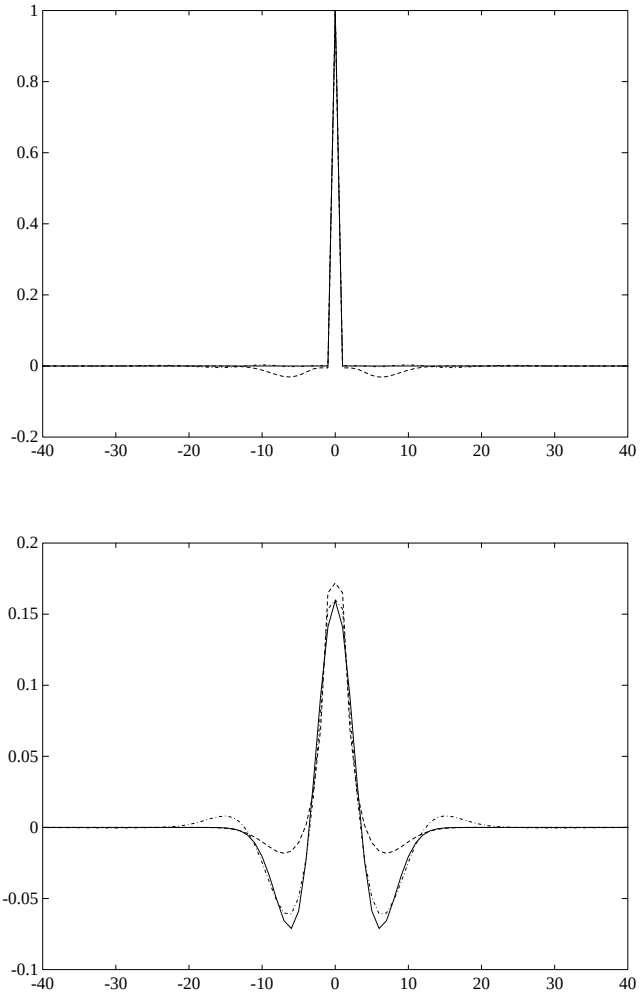
**Table 6.12**: GTLS compared with balanced reduction and Hankel norm approximation in reducing the Mexican hat.

| model | | $\mathcal{B}_{gtls}$ | $\mathcal{B}_{bal}$ | $\mathcal{B}_{hank}$ |
|---|---|---|---|---|
| n=2 | misfit | 0.17 | 0.18 | 0.32 |
| | error in impulse response | 0.21 | 0.20 | 0.34 |
| n=4 | misfit | 0.046 | 0.047 | 0.052 |
| | error in impulse response | 0.051 | 0.050 | 0.055 |
| n=6 | misfit | 0.0070 | 0.0071 | 0.0078 |
| | error in impulse response | 0.0076 | 0.0075 | 0.0081 |

It requires a change of norm 0.17 to make the data $(w_1, w_2)$ belong to the second order GTLS model, while the impulse response of this system differs 0.21 from the Mexican hat. Balanced reduction of the right half of the Mexican hat to a third order system, and using the symmetry in the Mexican hat, yields a sixth order model with misfit 0.0071.

Hankel norm reduction and especially balancing give rather good results. Ho-

**Figure 6.16**: GTLS approximations of the Mexican hat.




The solid lines give the impulse (upper plot) and the Mexican hat (lower plot). The dashed lines depicts the optimal approximation of the data of second order, and the dash/dotted lines of fourth order (hardly visible).

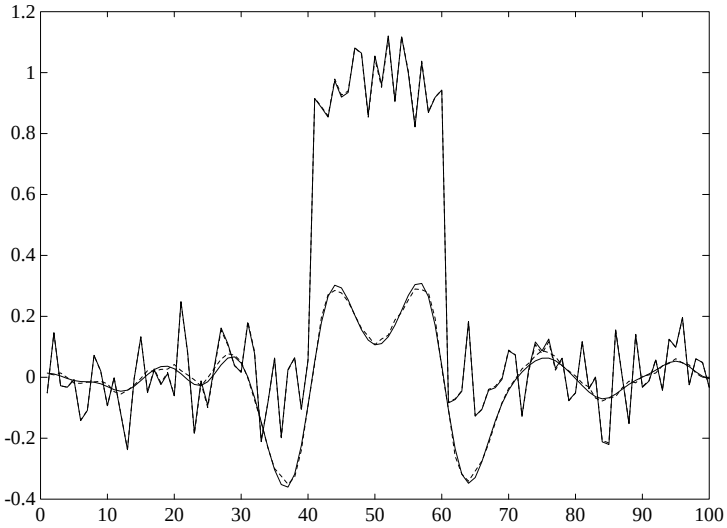
wever, they can only be used when a causal impulse response is available. The GTLS method makes no use of the symmetry of the observed signals, but this property is preserved well in the identified models. This is illustrated in Fig. 6.16, which contains the optimal approximations of orders 2 and 4.

We also apply the GTLS method to data $w_e$ consisting of two noisy steps for the input and the corresponding system output. These data and the optimal approximation of order 4 are given in Fig. 6.17. We should mention that the approximation error in the input is so small that it is nearly invisible in this figure.

The misfit of this model and the error in its impulse response are listed in the first column of Table 6.13. In view of the results for $n = 4$ in Table 6.12, this shows that the identified model is an accurate approximation of the Mexican hat.

In order to illustrate the difference between global and local methods, we finally

**Figure 6.17**: Estimating the Mexican hat from noisy step measurements.




The solid lines depict the data: a noisy step as input and its (smooth) effect in the Mexican hat filter (6.21). The dashed lines depict the optimal fourth order GTLS approximation.

consider once more the local total least squares method described in Section 2.5. The results for the fourth order model are given in the second column in Table 6.13.

**Table 6.13**: Local versus global method for the Mexican hat.

| method | Global TLS | Local TLS |
|---|---|---|
| misfit with respect to $w_e$ | 0.11 | 1.17 |
| error in impulse response | 0.066 | 0.522 |

The fourth order GTLS model for the noisy step measurement $w_n$ (solid lines in Fig. 6.17) has misfit 0.11. The difference between the impulse response of the fourth order Local TLS model for this data and the Mexican hat equals 0.522, which is even larger than the size of the Mexican hat itself (0.35).

This clearly shows that, perhaps not surprisingly, the local method gives poor results with respect to the global total least squares criterion. For example, the error in the impulse response of the local model is even larger than the Mexican hat itself, which has norm $\|w_2\| = 0.35$. This illustrates that local methods may completely miss global relations in the data.

# Chapter 7

# Conclusions

Global Total Least Squares (GTLS) is a method for determining approximate deterministic models for an observed time series. Models are defined as sets of time series that are linear and shift-invariant, and the aim is to determine a model of restricted dimension with minimal distance to the data. For further explanation we gave concrete descriptions in terms of difference equations and linear systems.
In the solution method, state representations played a central role. In particular isometric state representations turned out to be a useful tool in the GTLS algorithms. The reason is that they induce an orthonormal basis of a model and give insight in the relation between a model and its orthogonal complement. This was exploited in Algorithm 1, which determines the time series in a given model that is closest to the data. It also played a basic role in the recursive version Algorithm 2, which can be considered as a deterministic version of Kalman filtering. We derived optimality conditions that characterize stationary points of the GTLS criterion, i.e., models for which the misfit has zero derivative (as a function of model parameters). These conditions were derived from three model improvement constructions, that were worked into an algorithm for determining GTLS models, Algorithm 3. Considerably faster is the second GTLS Algorithm 4 in terms of Gauss-Newton iterations. Also in this algorithm isometric state representations plays a crucial role. Initial estimates were determined on basis of a modified canonical correlation analysis in Algorithm 5.
We concluded by several applications, concerning both economic time series and system theoretic issues.

We come to the following conclusions:

> *The behavioural approach to systems theory gives a useful framework for time series analysis.*

In our opinion it is enlightening in identification to concentrate on the behaviour of systems, i.e., the set of events that are compatible with it, instead of taking a starting point in a particular type of description of systems, as difference equations or state representations. The idea is that it is important *what*

a model says, and not *how*. The behavioural definition of models as sets of time series enables a description of the GTLS identification problem without any reference to model parameters. In this way we can make a clear distinction between identification problems, concerning the relation between data and model, and parametrization issues, concerning the numerical representation of models. In particular the uniqueness of representations plays a minor role in the algorithms, as we concentrate on system behaviour. In our approach it is not necessary to achieve uniqueness of parameters by choosing canonical forms, which is a complex issue especially for multi-input multi-output systems.

*The open approximate deterministic modelling methodology is suitable for economic time series analysis.*

Our modelling approach has some attractive properties that are relevant in econometrics. Firstly, we treat all variables symmetrically, without a priori distinction in explanatory (exogenous, inputs) variables and explained (endogenous, output) variables. In many cases in econometrics it is not obvious how to make such a decomposition, and then this is a desirable property of GTLS. Secondly, no stochastic assumptions are made, neither on external effects, as we use free auxiliary inputs to represent the effect of the environment, nor on deviations in the data with respect to the model, as we only minimize the size of these deviations. This is well-suited for many applications in econometrics, in which representing the environment or model deviations as random variables may be hard to defend.

*GTLS is new.*

The method has quite conventional aspects: models are linear and time-invariant, and the accuracy of models is measured by the sum of squared residuals. The distinctive feature of GTLS is the way in which the residuals are defined: as the difference between the observed time series and the closest time series that is in exact correspondence with the model laws. For the static case this amounts to (static) total least squares or orthogonal regression, cf. Section 2.3, but for the dynamic case the method is new. The method is called 'global' because we compare the data with time series that satisfy the (deterministic) model equations *throughout the entire observation interval*. This is in contrast with 'local' methods that take into account approximation errors on small intervals separately, e.g. the first step ahead prediction errors. The term 'total' refers to the fact that we allow for approximation errors in all data components, in contrast to methods that explain some (endogenous) components in terms of unexplained (exogenous) components that are not approximated. The combination of its global character and the symmetric treatment of all variables makes the difference with other least squares methods. GTLS models are also different from dynamic 'errors in variables' and 'factor' models.

*GTLS is feasible.*

The iterative GTLS algorithms, together with the heuristic algorithm 5 for determining initial models, turns out to be sufficiently powerful to determine small GTLS models at high accuracy in short time. The estimation of simultaneous equations does not cause additional problems. With some effort the method can be extended to the periodic case. On the other hand, resulting models might be only locally optimal, and we did not analyse the performance of the algorithms for the construction of large models.

*GTLS gives promising results.*

The applications illustrate the use of GTLS in various situations. The example on the US interest rates shows that its scope is not restricted to 'stationary looking' time series, as the non-stationary aspects may be incorporated in the degree of freedom in the system represented by the auxiliary input. GTLS is not only new from a theoretical point of view, it also gives different results in applications on real data. This was illustrated by the results on 'unit roots'. As the GTLS criterion takes into account the global implications of difference equations, it makes a sharp distinction between permanent effects (unit roots) and long lasting but decaying effects (nearly unit roots). The GTLS method is also of use in optimal model reduction, and the modelling of input/output data.

We have already proposed some technical topics for further research in the previous chapter (see also the index on 'further research') We conclude by some general ideas that may deserve further investigation.

*Analytic approach to GTLS.*

We have derived optimality conditions for stationary points of the GTLS criterion. One could try to solve these equations analytically. The advantage of such a method would be that it might give the exact global optimal GTLS model. This might also give insight in the number of local optima. Interesting results in this direction are obtained in [19].

*Stochastic analysis of GTLS.*

In order to obtain additional insight in the GTLS method, a stochastic analysis of the method would be interesting. For which data generating processes does the GTLS method give consistent estimates? How can the GTLS problem be formulated on the level of random variables? What are the statistical properties of the misfit in simulation experiments? Some of these questions are discussed in [22], but this is still a large open field.

*Analysis of isometric state representations.*

Isometric state representations form a highly structured representation of linear time-invariant systems in which all system variables are treated symmetrically. These representations may be exploited to gain further insight in the

structure of linear systems. For instance, both the eigenvalues and the singular values of the $A$ matrix are uniquely determined for a system, so they reflect properties of the system that are independent of input/output decompositions. It would be interesting to relate these 'system-invariants' to the memory length of systems.

*Use of GTLS for robust control.*

The aim of *robust control* is to develop controllers that do not only stabilize a given system, but also all systems in a certain neighbourhood. In this way one takes account of the approximate character of models. One of the popular approaches is based on a definition of the neighbourhood of systems in terms of so-called coprime factor perturbations, cf. [43]. We briefly indicated a relationship between this type of model uncertainty and ISR's in [35], that might serve as a starting point for further analysis.

# Appendix A

# Proofs

## A.2 The Global Total Least Squares Method

**Correctness of Definition 2.1.6.** We refer to [44], but in order to give some intuition behind the definition we give the proof below.
As $\mathcal{B}_{[1,t]}$ is a linear space, its dimension is well-defined. Consider the sequence $d_t := \dim(\mathcal{B}_{[1,t]})$. Clearly $d_t$ is increasing, and $0 \leq d_t \leq q$. So the increase of dimension, $\Delta d_t := d_t - d_{t-1}$, is a sequence of numbers between zero and $q$. Moreover, this sequence is non-increasing, which can be seen as follows. Suppose that $\Delta d_k > \Delta d_{k-1}$ for some $k \in \mathsf{N}$. This would mean that there is a larger degree of freedom in $\mathcal{B}$ for $w(k)$ given $w(1), \ldots, w(k-1)$ than for $w(k-1)$ given $w(1), \ldots, w(k-2)$. By time-invariance the latter is equal to the degree of freedom for $w(k)$ given $w(2), \ldots, w(k-1)$, which by definition cannot be smaller than for the degree of freedom given $w(1), \ldots, w(k-1)$. So the condition $\Delta_k > \Delta_{k-1}$ leads to a contradiction, from which it follows that $\Delta d_t$ is non-increasing.
As $0 \leq \Delta d_t \leq q$ this means that $\Delta d_t$ reaches its limit value within finite time. Let $m$ denote this limit value, and $T_{lim}$ the smallest time instant for which this is reached, so $\Delta d_{t'} = m$ for all $t' \geq T_{lim}$. Then there exists an $n \in \mathsf{N}$ such that $d_t = mt + n$ for $t \geq T_{lim}$. By definition, $\Delta d_{t'} > m$ for $1 \leq t' < T_{lim}$, from which it follows that $d_t > mt + T_{lim}$. So $T_{lim} \leq n$, from which the result follows.
We remark that models that coincide on finite intervals have the same size, so completeness plays no role in this definition. This should not be surprising, as completeness only concerns infinite-lag properties of $\mathcal{B}$ , while the definition of rank and degree only concern finite time intervals. ♣

**Proof of Lemma 2.4.2.** Clearly $\bar{\mathcal{B}}$ is complete and $\bar{\mathcal{B}}_T = \mathcal{B}_T$ for all finite intervals $t \subset \mathsf{Z}$. Further, $\bar{\mathcal{B}} \in \mathsf{B}^q$, as it is linear and shift-invariant. Concerning the uniqueness of $\bar{\mathcal{B}}$, let $\mathcal{B}'$ denote a complete model that coincides with $\mathcal{B}$ on finite intervals. Then $\hat{w} \in \mathcal{B}' \Rightarrow \hat{w}_T \in \mathcal{B}_T$ for all finite $T \Rightarrow \hat{w} \in \bar{\mathcal{B}}$, and conversely, $\hat{w} \in \bar{\mathcal{B}} \Rightarrow \hat{w}_T \in \mathcal{B}_T$ for all finite $T \Rightarrow \hat{w}_T \in \mathcal{B}'_T$ for all finite $T \Rightarrow \hat{w} \in \mathcal{B}'$. So $\mathcal{B}' = \bar{\mathcal{B}}$, which proves uniqueness. ♣

**Proof of Proposition 2.4.4.** Clearly the solution set (2.12) is linear, shift-invariant and complete. The fact that every system in $\bar{\mathsf{B}}^q$ admits such a description is proved by construction in [44, Theorem 10], in which also the other claims of

the proposition are proved. As the results are well-established, we give a somewhat informal proof on an intuitive level.

Let $\mathcal{B}$ denote a system in $\bar{\mathsf{B}}^q$ with rank $m$ and degree $n$, and suppose that it equals the solution set of $p$ independent difference equations of total lag $n'$, cf. (2.12). We first show that $p = q - m$, then that $n' \geq n$, and finally we indicate how to obtain a description with total lag $n' = n$.

As the equations are time-invariant, they can be organized in such a way that the leading matrix coefficient $R_0$ in (2.11) does not contain zero rows. Moreover, we may even assume that $R_0$ is of full row rank $p$, otherwise we could obtain an equivalent set of $p$ equations with a zero row in $R_0$ by taking linear combinations, then shift the equations such that this zero row vanishes, and repeat this procedure until $R_0$ has rank $p$. By assumption $\mathcal{B}$ is given by (2.12), so for $t > d$ there must hold $p$ independent linear equations for $w(t)$ given the past $\{w(t');\ 1 \leq t' < t\}$. This means that the number of degrees of freedom at each time instant equal $q - p$, so $m = q - p$.

Next we prove that the total lag $n' \geq n$. Therefore we compare for $t \geq n$ the dimension of $\dim(\mathcal{B}_{[1,t]})$ with the dimension on this interval if the past $\{w(t');\ t' \leq 0\}$ would have been specified. By definition of rank and degree, the first equals $mt + n$ while the latter equals $mt$, which is $n$ smaller. This must be due to the fact that for $t \geq n$ there hold $n$ linear relations that involve both time instants in $[1, t]$ and time instants $t' < 1$. Clearly a set of linear time-invariant difference equations cannot impose more independent linear relations of this type than their total lag, from which it follows that $n' \geq n$.

Finally we indicate how to obtain a description with minimal total lag $n' = n$. Consider the first time instant $t_1$ for which $\dim(\mathcal{B}_{[1,t_1]})$ is rank deficient, i.e. has rank below $qt_1$. Determine a corresponding linear relation, which is of lag $t_1 - 1$. By shift-invariance this relation must hold at every time, which implies that for $t \geq t_1$ it holds that $\dim(\mathcal{B}_{[1,t]}) \leq (q-1)t$. If $\mathcal{B}$ would be equal to the solution set of this equation, equality would hold. Now repeat the procedure by considering the first time-instant $t_2 \geq t_1$ for which equality does not hold, determine the corresponding second relation, and compute the induced restrictions on the model dimension. This should be repeated until the actual dimension of $\mathcal{B}$ is exactly in correspondence with the restrictions induced by the constructed relations. By completeness it then follows that $\mathcal{B}$ equals their solution set. By careful bookkeeping it can be shown that the number of independent equations equals $q - m$, and that their total lag equals $n$, the degree of $\mathcal{B}$ . ♣

## A.3 Isometric State Representations

**Proof of Proposition 3.1.2.** *1.* This is proved by construction in [45, Theorem 9]. There also the construction of SR's from a description in terms of difference equations is described. We give some intuition by describing how to translate a system described by one equation to state space form, so we consider

$$\mathcal{B} = \{\hat{w} : \mathsf{Z} \to \mathsf{R}^q;\ r_0\hat{w}(t) + \ldots + r_n\hat{w}(t-n) = 0,\ \text{for all } t \in \mathsf{Z}\}, \tag{A.1}$$

with $r_i \in \mathsf{R}^{1\times q}$ and $r_0$ and $r_n$ nonzero. We assume that the equation is scaled such that $r_0 r_0^\top = 1$. Now the definition of the state at time $t = 1$ can be derived from those restrictions that both involve the past $\{t' < 1\}$ and the future $\{t' \geq 1\}$, as follows. We define the $n$ components of the state at $t = 1$ by

$$\begin{array}{r|l} \textit{past} & \textit{future} \\ \hat{x}_1(1) := r_d\hat{w}(t-d-1) + \ldots + r_1\hat{w}(0) & = -r_0\hat{w}(1) \\ \vdots & \vdots \\ \hat{x}_n(1) := r_d\hat{w}(0) & = -r_{d-1}\hat{w}(1) - \ldots - r_0\hat{w}(d) \end{array}$$

Notice that the state is both a linear function of the past and of the future. This is a characterizing property of the state that is further discussed in cf. Section 5.6.1. It is easily derived that $\hat{w}(1)$ can be expressed as

$$\hat{w}(1) = [-r_0^\top; 0 \ldots 0]\hat{x}(1) + D\hat{v}(1) \tag{A.2}$$

with $D \in \mathsf{R}^{q\times q-1}$ such that $r_0 D = 0$. Now the state evolution equation follows from substituting this in the equations

$$\begin{array}{rcl} \hat{x}_1(2) & = & \hat{x}_2(1) + r_1\hat{w}(1) \\ \vdots & & \vdots \\ \hat{x}_{n-1}(2) & = & \hat{x}_n(1) + r_{n-2}\hat{w}(1) \\ \hat{x}_n(2) & = & r_{n-1}\hat{w}(1). \end{array}$$

*2.* Trivial
*3.* For a system $\mathcal{B}$ of size $(m, n)$ there exists an SR with $m$ auxiliary inputs and $n$ states. This is proved by construction in [45, Theorem 9,10], and we mention that this also follows from the proof of the sufficiency of the minimality conditions that will be given in Proposition 3.1.5. For systems that are smaller than $(m, n)$ an SR with $m$ auxiliary inputs and $n$ states is obtained by adding ineffective auxiliary variables. This proves the existence of SR's for systems $\mathcal{B} \in \bar{\mathsf{B}}^{q,m,n}$.
The proof of the fact that there does not exist SR's with less auxiliary inputs than the rank of a system or with smaller state dimension than its degree is left to the reader.

♣

**Proof of Proposition 3.1.4.** By definition, $\hat{w}$ is contained in $\mathcal{B}(A, B, C, D)$ if and only if there exist an $\hat{v}$ and $\hat{x}$ such that $\sigma\hat{x} = A\hat{x} + B\hat{v}$ and $\hat{w} = C\hat{x} + D\hat{v}$. These equations are equivalent to $\sigma(Sx) = S(A + BF)S^{-1}(Sx) + SBR(R^{-1}(v - Fx))$ and $w = (C + DF)S^{-1}(Sx) + DR(R^{-1}(v - Fx))$, provided that $S$ and $R$ are invertible.

This shows the equivalence of the representations. For a proof of the fact that for minimal representations all equivalent representations are obtained in this way we refer to [20, Corollary II.3-25]. This proof relies on the fact that the state in a minimal SR for a given system trajectory is unique modulo a basis transformation, which is derived in [45, 2.4.3, 3.2.5, 4.7.5]. ♣

**Proof of Proposition 3.1.5.** *Necessity of 1.* Clearly $\hat{x}(t) \in \mathrm{im}[A\ B]$ for all $t \in \mathsf{Z}$, as it satisfies $x(t) = Ax(t-1) + Bv(t-1)$ for some $x(t-1) \in \mathsf{R}^n$ and $v(t-1) \in \mathsf{R}^n$. Now if $[A\ B]$ has not full row rank, there is a basis transformation of the state $\hat{x} \to S\hat{x}$ such that $S\hat{x}$ contains zero components. Removing these components yields an equivalent SR with less states.
*Necessity of 2.*
The matrix pair $(A, C)$ is called *observable* if the matrix $\mathrm{col}(\mathrm{C}, \mathrm{CA}, \ldots, CA^{n-1})$ has full column rank $n$. Suppose there exists an $F$ such that $(A + BF, C + DF)$ is not observable. Then the unobservable components of the state can be removed, as in the previous part.
*Necessity of 3.* An obvious condition is that $\begin{bmatrix} B \\ D \end{bmatrix}$ has full column rank, so it suffices to prove that $\ker D \subset \ker B$. Suppose $\ker D \backslash \ker B \neq \{0\}$, then there exists an invertible $R \in \mathsf{R}^{m\times m}$ such that $\begin{bmatrix} BR \\ DR \end{bmatrix} = \begin{bmatrix} B' & b \\ D' & 0 \end{bmatrix}$ with $0 \neq b \in \mathsf{R}^n$. Then $\mathcal{B}$ is represented by $\sigma x = Ax + B'v' + bz$ and $w = Cx + D'v'$. As $z$ influences $w$ only with a delay, we remove its direct influence on $\sigma x$ by defining $\sigma x' := \sigma x - bz$. This gives $\sigma x' = Ax' + B'v' + Ab\sigma^{-1}z$ and $w = Cx' + D'v' + Cb\sigma^{-1}z$. Hence $(A, B, C, D)$ is equivalent to $(A, [B'\ Ab], C, [D'\ Cb])$. From Proposition 3.1.4 it follows that this is equivalent to $(A + Abf, [B'\ Ab], C + Cbf, [D'\ Cb])$, for all $f \in \mathsf{R}^{1\times n}$. As $b \neq 0$, $f$ can be chosen such that $I_n + bf$ is singular, by taking $f = -b^\top/\|b\|^2$. It is easily verified that then $(A + Abf, C + Cbf) = (A(I_n + bf), C(I_n + bf))$ is not observable. From the necessity of *2* it follows that the state dimension can be reduced.

*Sufficiency of the conditions.*
Let $(A, B, C, D)$ be an SR with $m$ auxiliary inputs and $n$ states that satisfies conditions *1*,*2* and *3*, and consider $\mathcal{B}(A, B, C, D)$. We show that its rank equals $m$ and its degree $n$, or, equivalently, that $\dim(\mathcal{B}_{[1,t]}) = mt + n$ for $t \geq n$.
As $[A\ B]$ has full rank, every state can occur at $t = 1$, so there are no restrictions on $\hat{x}(1)$. Then

$$\begin{aligned} \mathcal{B}_T = \hat{w} : T \to \mathsf{R}^q;\ &\exists \hat{x}(1) \in \mathsf{R}^n, \hat{v} : T \to \mathsf{R}^m \text{ such that} \\ &\hat{x}(t+1) = A\hat{x}(t) + B\hat{v}(t) \text{ and } \hat{w}(t) = C\hat{x}(t) + D\hat{v}(t). \end{aligned} \tag{A.3}$$

As $\hat{v}$ is a free variable, and $\mathrm{rank} D = m$, $\dim(\mathcal{B}_{[1,t]}) \geq mt$, from which it is easily seen that $\mathrm{rank}(\mathcal{B}) = m$.
In order to prove that $\mathrm{degree}(\mathcal{B}) = n$ we will derive a contradiction from $\dim(\mathcal{B}_{[1,t]}) < mt + n$ for $t > n$. Namely, then there must exists a nonzero initial state $\hat{x}(1)$ and auxiliary inputs $\hat{v} : T \to \mathsf{R}^m$ such that the corresponding system trajectory on $T$ in (A.3) satisfies $\hat{w} = C\hat{x} + D\hat{v} = 0$. Let $D^\sharp \in \mathsf{R}^{m\times q}$ denote the left-inverse of $D$, so that $D^\sharp D = I_m$. The existence is guaranteed by the injectivity of $D$. From $D^\sharp C\hat{v} + D^\sharp D = D^\sharp \hat{w} = 0$ it follows that $\hat{v} = -D^\sharp C\hat{x}$. Define $F := -D^\sharp C$, then $\hat{x}(t+1) = (A + BF)\hat{x}(t)$ and $\hat{w}(t) = (C + DF)\hat{x}(t) = 0$, so $(A + BF, C + DF)$ is unobservable, which contradicts the second minimality condition. ♣

**Proof of Proposition 3.2.5.**
Write $(A', B', C', D') := (S(A+BF)S^{-1}, SBR, (C+DF)S^{-1}, DR)$. Then

$$\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} S^{-1} & 0 \\ FS^{-1} & R \end{pmatrix}.$$

Equation (3.7) for $(A', B', C', D')$ gives, with $K = S^\top S$,

$$\begin{pmatrix} A^\top & C^\top \\ B^\top & D^\top \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} S^{-1} & 0 \\ FS^{-1} & R \end{pmatrix}^{-T} \begin{pmatrix} S^{-1} & 0 \\ FS^{-1} & R \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} S^\top & -F^\top R^{-\top} \\ 0 & R^{-\top} \end{pmatrix} \begin{pmatrix} S & 0 \\ -R^{-1}F & R^{-1} \end{pmatrix} =$$

$$\begin{pmatrix} K + F^\top (RR^\top)^{-1} F & -F^\top (RR^\top)^{-1} \\ -(RR^\top)^{-1} F & (RR^\top)^{-1} \end{pmatrix}.$$

Now verification of the equations (3.13), (3.14) and (3.11) is straightforward. For a proof of the uniqueness of $K$ we refer to [29, Theorem 2.1]. We remark that the criterion for the existence and uniqueness of a positive definite solution for $K$ formulated there is that $(A, B)$ is stabilizable and $(A, B, C, D)$ strongly observable (which is the same as the second condition in Proposition 3.1.5.2), so the injectivity of $D$ is not essential. ♣

**Proof of Proposition 3.2.6.** The equivalence of the representations is obvious. The fact that all minimal ISR's are obtained this way can be derived as follows. From the proof of Proposition 3.2.5 it follows that the equations (3.12) - (3.14) are necessary conditions. Equation (3.12) determines $S$ modulo a left unitary factor, corresponding to $U$. Equation (3.13) determines $R$ modulo a right unitary factor, corresponding to $V$. ♣

**Proof of Proposition 3.2.7.** *1.* Let $m$ denote the number of auxiliary inputs in an ISR. If $\mathrm{rank}(\mathcal{B}) < m$ then $\dim(\mathcal{B}_{[1,t]}) < (m-1)t + n'$ for all $t$ larger than some $n' \in \mathsf{N}$. This implies the existence of a $t - n'$ dimensional linear space of auxiliary inputs on $T = [1, t]$ that yield a zero system trajectory for zero initial state. For $N > n' + n$ this space has dimension larger than $n$, so there must be a non-zero auxiliary input $\hat{v} : [1, N] \to \mathsf{R}^m$ that corresponds to a zero system trajectory $\hat{w} = 0$ with both the initial state $\hat{x}(1)$ and final state $\hat{x}(N+1)$ equal to zero. This contradicts (3.6).
*2.* From (3.7) it follows that $A^\top A = I_n - C^\top C$, hence $|Ax|^2 = |x|^2 - |Cx|^2 \leq |x|^2$. So $A$ is stable. We prove that the representation is not minimal if $A$ is not asymptotically stable. In that case $A$ has an eigenvalue $\lambda$ with $|\lambda| = 1$. Let $x$ denote a corresponding eigenvector, and $x^*$ its complex conjugate. Then $|Cx|^2 = x^* C^\top C x = x^* x - x^* A^\top A x = |x|^2 - |Ax|^2 = 0$. This implies that $CA^k x = 0$, $k \geq 0$, so that $(A, C)$ is not observable and hence not minimal.
*3.* Let $m$ denote the number of auxiliary inputs, and $n$ the number of state variables in $(A, B, C, D)$. This ISR is minimal if and only if $(m, n)$ equals the size of the system $\mathcal{B} := \mathcal{B}(A, B, C, D)$, cf. Corollary 3.1.3, and this is equivalent to the condition that

$$\dim(\mathcal{B}_{[1,N]}) = mN + n \tag{A.4}$$

for all $N \geq n$, cf. Definition 2.1.6. By definition of SR's, if $\hat{w} \in \mathcal{B}_{[1,N]}$, then there exist an initial state $\hat{x}(1)$ and auxiliary inputs $\hat{v} : [1, N] \to \mathsf{R}^m$ such that

$$\hat{w}(t) = C\hat{x}(t) + D\hat{v}(t) \text{ for } \hat{x}(t+1) = A\hat{x}(t) + B\hat{v}(t). \tag{A.5}$$

Notice that $\hat{w}$ is a linear function of the initial state and auxiliary inputs, and that the number of parameters for $\mathcal{B}_{[1,N]}$ equals $mN + n$. So (A.4) is equivalent to the following two conditions.

*a:* Every initial state $\hat{x}(1)$ can occur.

*b:* The parametrization is injective for $N \geq n$.

We prove that *a* is equivalent to '$W$ non singular', and that *b* is equivalent to '$WW^\top$ asymptotically stable'. Then the result follows immediately.
The condition *a* is equivalent to the condition that $[A\ B]$ has full row rank $n$, cf. Proposition 3.1.5.1. Namely, for all $k \in \mathsf{N}$ it holds that

$$\text{rank}([A\ B]) = n \Leftrightarrow \text{rank}([A^k \vdots A^{k-1}B \vdots \ldots \vdots B]) = n.$$

This is easily derived from an inductive argument, based on the fact that $\text{rank}([A\ B]) = n$ if and only if $\text{rank}([A.[A\ B]\ B]) = n$. Taking $k = n$ yields $a \Leftrightarrow \text{rank}[A\ B] = n$.
It remains to prove that condition *b* is equivalent to '$WW^\top$ asymptotically stable'. Condition *b* means that for $N \geq n$ all nonzero parameters $\hat{x}(1) \in \mathsf{R}^n$ and $\hat{v} : [1, N] \to \mathsf{R}^m$ correspond to a nonzero system trajectory $\hat{w}$ on $[1, N]$. Clearly this is true for all $N \geq n$ if and only if it is true for $N = n$. From (3.6) it follows that this is equivalent to the condition that

$$|\hat{x}(n+1)|^2 < |\hat{x}(1)|^2 + |\hat{v}(1)|^2 + \ldots + |\hat{v}(n)|^2,$$

for all $\hat{x}$ and $\hat{v}$ that can occur in (A.5). Now this is exactly the same condition as '$WW^\top$ asymptotically stable', as $W$ denotes the mapping from $\hat{x}(1)$ and $\hat{v}$ to $\hat{x}(n+1)$. ♣

**Proof of Proposition 3.3.2.** First we prove that

$$\mathcal{B}_{[1,n+1]} = \mathcal{B}_{[1,n+1]}(A, B, C, D) \Rightarrow \mathcal{B}_{[1,n+2]} = \mathcal{B}_{[1,n+2]}(A, B, C, D). \tag{A.6}$$

$\supset$: Suppose $\hat{w} \in \mathcal{B}_{[1,n+2]}(A, B, C, D)$. Then clearly both $\hat{w}_{[1,n+1]}$ and $\hat{w}_{[2,n+2]}$ in $\mathcal{B}_{[1,n+1]}(A, B, C, D) = \mathcal{B}_{[1,n+1]}$. Hence $\hat{w}_{[1,n+1]}, \hat{w}_{[2,n+2]} \in \mathcal{V} := \{\hat{w} : [1, n+2] \to \mathsf{R}^q;\ \hat{w}_{[1,n+1]} \in \mathcal{B}_{[1,n+1]} \text{ and } \hat{w}_{[2,n+2]} \in \mathcal{B}_{[2,n+2]}\}$. We show that $\mathcal{V} = \mathcal{B}_{[1,n+2]}$, from which it follows that $\hat{w} \in \mathcal{B}_{[1,n+2]}$. Clearly $\mathcal{B}_{[1,n+2]} \subset \mathcal{V}$, so if the dimension of both sets are equal, then $\mathcal{B} = \mathcal{V}$. Indeed, $\dim\mathcal{V} = \dim(\mathcal{B}_{[1,n+1]}) + \dim(\mathcal{B}_{[2,n+2]}) - dimb[1, n+1] = (n+2)m + n = \dim(\mathcal{B}_{[1,n+2]})$, with $m$ the rank of $\mathcal{B}$.
$\subset$: Suppose $\hat{w} \in \mathcal{B}_{[1,n+2]}$. From time-invariance of $\mathcal{B}$ it follows that both $\hat{w}_{[1,n+1]}$ and $\hat{w}_{[2,n+2]}$ are in $\mathcal{B}_{[1,n+1]} = \mathcal{B}_{[1,n+1]}(A, B, C, D)$. In order to show that $\hat{w} \in \mathcal{B}_{[1,n+2]}(A, B, C, D)$ we introduce the following notation. Let $\hat{x}, \hat{v}$ denote the state and auxiliary input for $\hat{w}_{[1,n+1]}$ in $(A, B, C, D)$, and $\hat{x}', \hat{v}'$ for $\hat{w}_{[2,n+2]}$, and consider the difference between the state at $t = 2$ denoted as $\bar{x}(2) := \hat{x}(2) - \hat{x}'(2)$. If $\bar{x}(2) = 0$, then $\hat{w}_{[1,n+2]}$ corresponds to initial state $\hat{x}(1)$ and auxiliary inputs $\{\hat{v}(1), \hat{v}'(2), \ldots, \hat{v}'(n+2)\}$, so $\hat{w}_{[1,n+2]} \in \mathcal{B}_{[1,n+2]}(A, B, C, D)$. If $\bar{x}(2) \neq 0$, we proceed as follows. Notice that for $\bar{x}(2)$ as initial state and auxiliary inputs $\{\hat{v}(2) -$

$\hat{v}'(2), \ldots, \hat{v}(n+1) - \hat{v}'(n+1)\}$ the corresponding system trajectory in $(A, B, C, D)$ is zero. We remark that this indicates that the SR is not minimal, cf. Proposition 3.2.7. It is easily verified that if $\bar{x}(2)$ need not have effect on $[2, n+1]$, then it also need not have effect on $[2, n+2]$. This means that there exist auxiliary inputs $\{\bar{v}(2), \ldots, \bar{v}(n+1)\}$ such that for initial state $\bar{x}(2)$ the system trajectory is zero. Then $\hat{w}_{[1,n+2]}$ corresponds to initial state $\hat{x}(1)$ and auxiliary inputs $\{\hat{v}(1), \hat{v}'(2) + \bar{v}(2), \ldots, \hat{v}'(n+1) + \bar{v}(n+2)\}$, so $\hat{w} \in \mathcal{B}_{[1,n+2]}(A, B, C, D)$. This concludes the proof of (A.6).

*Proof of (3.19).* Suppose that the lefthand side in (A.6) holds true. By an obvious induction argument it follows that then the righthand side is true not only for $[1, n+2]$ but also for $[1, k]$ with $k \geq n+1$. From shift-invariance it follows that $\mathcal{B}_T = \mathcal{B}_T(A, B, C, D)$ for all finite intervals on $\mathsf{Z}$. Now the result follows from the completeness of $\mathcal{B}$ , cf. Lemma 2.4.2.

*Proof of (3.20).* Suppose that $\mathcal{B} = \mathcal{B}(A, B, C, D)$. Clearly $\mathcal{B}_T \subset \mathcal{B}_T(A, B, C, D)$ for all finite $T \subset \mathsf{Z}$. To prove the converse inclusion, suppose that $\hat{w} \in \mathcal{B}_T(A, B, C, D)$. Now let $\hat{x}_T, \hat{v}_T$ denote the state and auxiliary input for $\hat{w}$ on $T$ in $(A, B, C, D)$. As $[A\ B]$ has rank $n$, there exists an extension of $\hat{x}_T$ and $\hat{v}_T$ to $\mathsf{Z}$ for which $\hat{x}(t+1) = A\hat{x}(t) + B\hat{v}(t)$ keeps valid. This proves that $\hat{w} \in \mathcal{B}_T$. ♣

## A.4 Evaluation of the Misfit

**Proof of Lemma 4.1.1.** This is a well-known result. Clearly $\hat{w}$ is the orthogonal projection of $w$, as $\hat{w} \in \mathcal{M}$ and $\langle \hat{w}, w - \hat{w} \rangle = 0$. Further, $\hat{w}$ minimizes $|w - Gv|$ over $v \in \mathsf{R}^k$, which can be seen as follows. Let $v' = v - G^*w$, then $|w - Gv|^2 = |w - GG^*w - Gv'|^2 = |w - GG^*w|^2 + |Gv'|^2$, as $\langle w - GG^*w, Gv' \rangle = \langle G^*w - G^*w, v' \rangle = 0$ where we use that $G$ is isometric so that $G^*G = I_m$. So the minimum is achieved by taking $v' = 0$, and $\hat{w} = GG^*w$. ♣

**Proof of Proposition 4.1.2.** Let $(\hat{w}^*, \hat{x}^*)$ denote the optimal approximation of $w$ in $\mathcal{B}_T$ together with its corresponding end state in the ISR corresponding to $G$. Then $GG^*(w, \hat{x}_{N+1}) = (\hat{w}^*, \hat{x}^{N+1})$, which can be seen as follows. From Proposition 4.1.2 it follows that $GG^*(w, \hat{x}_{N+1})$ is the element $(w', x')$ in the image of $G$ that is the closest to $(w, \hat{x}_{N+1}$, i.e., with $\|w - w'\|^2 + \|\hat{x}_{N+1} - x'\|^2$ minimal. Clearly the minimum is attained for $(\hat{w}^*, \hat{x}^*_{N+1})$, as this corresponds to the optimal approximation of $w$ in $\mathcal{B}_T$ and no error in the end state. This proves the result, as the minimum for (4.5) is attained for $\hat{x}_{N+1}$ and corresponds to the optimal approximation $\hat{w}^*$. ♣

**Proof of Proposition 4.2.1.** The proof is based on the following result. Let $(\hat{x}, \tilde{v})$ and $(\tilde{x}, \tilde{v})$ denote the state and auxiliary input for respectively $\hat{w} \in \mathcal{B}_T(A, B, C, D)$ and $\tilde{w} \in \mathcal{B}_T(A, \tilde{B}, C, \tilde{D})$, with $T = [1, N]$. Then

$$\langle \hat{w}, \tilde{w} \rangle = \hat{x}(1)^\top \tilde{x}(1) - \hat{x}(N+1)^\top \tilde{x}(N+1) \tag{A.7}$$

This can be proved as follows. Define $\hat{z} := \begin{pmatrix} \hat{x} \\ \hat{v} \\ 0 \end{pmatrix}$ and $\tilde{z} := \begin{pmatrix} \tilde{x} \\ 0 \\ \tilde{v} \end{pmatrix}$, so that $\hat{w} = [CD\tilde{D}]\hat{z}$ and $\tilde{w} = [CD\tilde{D}]\tilde{z}$. Then it follows that $\langle \tilde{w}, \hat{w} \rangle = \left\langle [CD\tilde{D}]\hat{z}, [CD\tilde{D}]\tilde{z} \right\rangle = \left\langle [CD\tilde{D}]^\top [CD\tilde{D}]\hat{z}, \tilde{z} \right\rangle = \left\langle I - [AB\tilde{B}]^\top [AB\tilde{B}]\hat{z}, \tilde{z} \right\rangle = \langle \hat{z}, \tilde{z} \rangle - \left\langle [AB\tilde{B}]\hat{z}, [AB\tilde{B}]\tilde{z} \right\rangle = \Sigma_{t=1}^N \hat{x}(t)^\top \tilde{x}(t) - \hat{x}(t+1)^\top \tilde{x}(t+1)$, from which (A.7) follows.
Now let $\mathcal{V}$ denote the space $\{\tilde{w} : T \to \mathsf{R}^q;\ \{\ldots, 0, 0, \tilde{w}, 0, 0, \ldots\} \in \mathcal{B}(A, \tilde{B}, C, \tilde{D})\}$, so elements of $\mathcal{V}$ correspond to system trajectories in $\mathcal{B}(A, \tilde{B}, C, \tilde{D})$ with boundary states $\hat{x}(1) = 0$ and $\hat{x}(N+1) = 0$. From (A.7) it follows that $\mathcal{V} \subset (\mathcal{B}_T)^\perp$. Finally, $\mathcal{V} = (\mathcal{B}_T)^\perp$ follows from $\dim(\mathcal{B}_T)^\perp = \dim(\mathsf{R}^q)^\top - \dim(\mathcal{B}_T) = q(N+1) - m(N+1) - n = (q-m)(N+1) - n = \dim(\mathcal{V})$. ♣

**Proof of Proposition 4.2.2.** In Proposition 3.2.7.3. we have characterized minimality of ISR's with degree $n$ by the following conditions, formulated in terms of $W := [A^n \vdots A^{n-1}B \vdots \ldots \vdots B]$:

1. $W$ non-singular
2. $WW^\top$ asymptotically stable.

The first condition is equivalent to '$[A\ B]$ of full row rank', as has already been proved in Proposition 3.2.7.3. The second condition is equivalent to '$(A, \tilde{B})$ controllable', i.e. such that $[\tilde{B}\ A\tilde{B}\ \ldots\ A^{n-1}\tilde{B}]$ has full row rank $n$. Namely, for $\tilde{W} := W := [A^{n-1}\tilde{B} \vdots \ldots \vdots \tilde{B}]$ it holds that $WW^\top + \tilde{W}\tilde{W}^\top = I_n$, which is easily obtained from

$[A\ B\ \tilde{B}][A\ B\ \tilde{B}]^\top = I_n$. So $WW^\top$ is asymptotically stable if and only if $\tilde{W}$ is non-singular, which is equivalent to the controllability of $(A, \tilde{B})$. ♣

**Correctness of Algorithm 1.** First we describe how to determine the optimal value for $x_{N+1}$ in step 2. Let $\tilde{v}'$ be given by (4.15), now for $x_{N+1} = 0$. Let $\mathcal{X}$ be the $n$-dimensional linear space corresponding to the effect of $x_{N+1}$ in (4.15) on $\tilde{v}$ , now for $w(t) = 0$, so $\mathcal{X} = \{z : T \to \mathsf{R}^{q-m}$ for which there exist an $x'(N+1) \in \mathsf{R}^n$ such that $z(t) = \tilde{B}^\top x'(t+1)$ for $x'(t) = A^\top x(t+1)\}$. Then determine the orthogonal projection of $\tilde{v}'$ onto $\mathcal{X}$, and let $x^*$ denote the corresponding value of $x'(N+1)$. Then $x_{N+1} := -x^*$ is the optimal value in step 2.
*1:* $\hat{w} + \tilde{w} = w$, *2:* $\|\tilde{w}\| = \|\tilde{v}\|$. *3:* $\hat{w}$ is the optimal approximation of $w$ in $\mathcal{B}_T$.
*1.* By premultiplying the equations (4.15) by $\begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix}$ it is easily shown that $\hat{x} + \tilde{x} = x$ and $\hat{w} + \tilde{w} = w$.
*2.* This follows from Proposition 3.2.2 and the fact that $(A, \tilde{B}, C, \tilde{D})$ is isometric.
*3.* This follows from Proposition 4.1.2. Namely, in step 2 we determine $(x(1), \hat{v}) = G^*(w, x(N+1))$, and in step 3 $(\hat{w}, \hat{x}(N+1)) = G(x(1), \hat{v})$, with $x(N+1)$ determined such that $\|\tilde{v}\| = \|w - \hat{w}\|$ is minimal. ♣

**Correctness of Algorithm 2.** The main part of this proof is devoted to the state update equation (4.23). Once this has been established, the formula for the past induced approximation $\hat{w}_t(t)$ in (4.24) and the formulas for the misfit follow from straightforward calculations.
We derive (4.23) by induction. In principle we could start in $t = 0$, but notice that $\hat{x}_0(1)$ is a meaningless variable, as it denotes the optimal end state of an empty observation. In fact, it can be shown that the value for $\hat{x}_0(1)$ is irrelevant, as it has no effect on the evolution of $\hat{x}_t(t+1)$ in (4.23). However, it might be more clarifying to start the inductive proof at $t = 1$.
From Algorithm 1 it follows that $x_1(2)$ should minimize $\|\tilde{v}(1)\|^2 = \|\tilde{B}^\top x_1(2) + \tilde{D}^\top w(1)\|^2$. Taking the derivative at $x_1(2)$ gives $\tilde{B}\tilde{B}^\top x_1(2) + \tilde{B}\tilde{D}^\top w(1) = 0 = \tilde{W}_1 x_1(2) + \tilde{F}_1 w(1)$. If $\tilde{W}_2$ is invertible, it follows that $x_1(2) = \tilde{W}_2^{-1}\tilde{F}_1 w(t)$ is indeed optimal. For the case that $\tilde{W}_2$ is singular, we need the following lemma.

**Lemma A.4.1 (Pseudo-inverse)** *Consider the equation $Wx + z = 0$ with $W$ singular, and $z \in \mathrm{im}W$. Let $W^\sharp$ denote a pseudo-inverse of $W$, i.e., $W^\sharp W W^\sharp = W^\sharp$ and $WW^\sharp W = W$. Then $x = -W^\sharp z$ is a solution.*

*Proof of the lemma.* As $z \in \mathrm{im}W$, there exists a $z'$ such that $z = Wz'$. Sustituting $x = -W^\sharp z$ gives $Wx + z = -WW^\sharp Wz' + Wz' = 0$. □

This lemma can be applied to the equation $\tilde{W}_1 x_1(2) + \tilde{F}_1 w(1) = 0$, as $\tilde{F}_1 w(1) = \tilde{B}\tilde{D}^\top w(1) \subset \mathrm{im}\tilde{W}_2$, which gives $x_1(2) = \tilde{W}_2^\sharp \tilde{F}_1 w(t)$.
So we proved (4.23) for $t = 1$. The induction step is proved with the aid of the following lemma.

**Lemma A.4.2** *Suppose $x_{t-1}(t)$ minimizes $\|\tilde{v}_{t-1}\|$ in (4.15) with $N = t-1$. Then $x_t(t+1)$ minimizes $\|\tilde{v}_t\|$ in (4.15) with $N = t$ if and only if*

$$\tilde{W}_{t+1} x_t(t+1) = \tilde{W}_{t+1} A x_{t-1}(t) - \tilde{F}_t \epsilon(t) \tag{A.8}$$

.

*Proof of the lemma.* Define $\bar{x}(t) := x_t(t) - x_{t-1}(t)$, denoting the adaption of $x$ at time $t$ due to the observation $w(t)$. Similarly, define $\bar{v}(t) := \tilde{v}_t(t) - \tilde{v}_{t-1}(t)$. From the definition of $\tilde{v}_{t-1}$ and $\tilde{v}_t$, cf. (4.15), it follows that $\bar{v}(t)$ equals the backward effect of $\bar{x}(t)$ in (4.15), i.e.,

$$\bar{v}(t) := \tilde{v}_t(t') - \tilde{v}_{t-1}(t') = \tilde{B}^\top A^{\top t-t'-1}\bar{x}(t) \text{ for } 1 \leq t' \leq t-1, \tag{A.9}$$

From the definition of $\tilde{W}_t$ it follows that $\|\bar{v}\| = (\bar{x}^\top \tilde{W}_t \bar{x}(t))^{1/2}$. Moreover, $\bar{v}$ is orthogonal to $\tilde{v}_{t-1}$, cf. the proof of Algorithm 1, part 3. This gives

$$\|\tilde{v}_t\|^2 = \|\tilde{v}_{t-1}\|^2 + \bar{x}^\top \tilde{W}_t \bar{x}(t) + |\tilde{v}_t(t)|^2. \tag{A.10}$$

So $x_t(t+1)$ minimizes $\|\tilde{v}_t\|$ if and only if it minimizes $\bar{x}^\top \tilde{W}_t \bar{x}(t) + |\tilde{v}_t(t)|^2$. Substituting $\tilde{v}_t = \tilde{B}^\top x_t(t+1) + \tilde{D}^\top w(t)$ and $\bar{x}(t) = A^\top x_t(t+1) + C^\top w(t) - x_{t-1}(t)$, and then taking the derivative after $x_t(t+1)$ gives the result. □

If $\tilde{W}_{t+1}$ is invertible, (4.23) follows immediately. If $\tilde{W}_{t+1}$ is singular, we have to replace the inverse $\tilde{W}_{t+1}$ by a pseudo-inverse $\tilde{W}_{t+1}^\sharp$ . Notice that $\mathrm{im}\tilde{F}_t \subset \mathrm{im}\tilde{W}_{t+1}$, as $\tilde{W}_{t+1} = [A\tilde{W}_t^{1/2}\ \tilde{B}][A\tilde{W}_t^{1/2}\ \tilde{B}]^\top$ and $\tilde{F}_t = [A\tilde{W}_t^{1/2}\ \tilde{B}][C\tilde{W}_t^{1/2}\ \tilde{D}]^\top$. Now the result follows from applying Lemma A.4.1. This concludes the inductive proof of (4.23).

Next we prove (4.24), by showing that

$$\tilde{w}_t(t) = \tilde{H}_t \epsilon(t). \tag{A.11}$$

Then the result follows from $\hat{w}_t(t) = w(t) - \tilde{w}_t(t) = C\hat{x}_t(t) + (I - \tilde{H}_t)\epsilon(t)$, where we used the definition of $\epsilon$ (4.22). By definition, $\tilde{w}_t(t) = C\tilde{x}_t(t) + \tilde{D}\tilde{v}_t(t)$, where $\tilde{x}_t(t) = \Sigma_{k=1}^t A^{t-1}\tilde{B}\tilde{v}_t(t-k) = \Sigma_{k=1}^t A^k \tilde{B}\bar{v}_t(t-k) + \Sigma_{k=1}^t A^k \tilde{B}\tilde{v}_{t-1}(t-k) = \tilde{W}_t\bar{x}_{t-1} + \tilde{x}t-1(t) = \tilde{W}_t\bar{x}_{t-1}$, where we used (A.9) to obtain the third equation, while $\tilde{x}_{t-1}(t) = 0$ follows from Proposition 4.2.1.2. So $\tilde{w}(t) = C\tilde{W}_t\bar{x}_t(t) + \tilde{D}\tilde{v}_t(t)$. Now substitute $\bar{x}(t) = x_t(t) - x_{t-1}(t)$ and

$$\begin{aligned} x_t(t) &= A^\top x_t(t+1) + C^\top w(t) && \text{(A.12)} \\ v_t(t) &= \tilde{B}^\top x_t(t+1) + \tilde{D}^\top w(t) && \text{(A.13)} \end{aligned}$$

gives $\tilde{w}_t(t) = \tilde{F}_t^\top x_t(t+1) + \tilde{G}_t w(t) - C\tilde{W}x_{t-1}(t)$. Using (4.23) gives $\tilde{w}_t(t) = \tilde{H}_t\epsilon(t) + (\tilde{F}_t A + \tilde{G}_t C - C\tilde{W}_t)x_{t-1}(t)$. It is easily derived from (3.7) that the last term cancels, from which (A.11) follows.

Finally we prove the equations for the misfit. From (A.10) it follows that

$$m(t)^2 = |\tilde{v}_t(t)|^2 + \bar{x}(t)^\top \tilde{W}\bar{x}(t) \tag{A.14}$$

Rewrite $\bar{x}(t)$ as

$$\begin{aligned} \bar{x}(t) &= A^\top x_t(t+1) + C^\top w(t) - x_{t-1}(t) \text{ (from A.12)} \\ &= A^\top x_t(t+1) + C^\top \epsilon(t) + C^\top C x_{t-1}(t) \text{ (from 4.22)} \\ &= (A^\top A + C^\top C - I)x_t(t+1) + (C^\top - A^\top \tilde{W}_{t+1}^{-1}\tilde{F}_t)\epsilon(t) \text{ (from 4.23)} \\ &= (C^\top - A^\top \tilde{W}_{t+1}^{-1}\tilde{F}_t)\epsilon(t) \text{ (from 3.7).} \end{aligned}$$

In a similar way we obtain $\tilde{v}_t(t) = (\tilde{D}^\top - \tilde{B}^\top \tilde{W}_{t+1}^{-1}\tilde{F}_t)\epsilon_t$. Substituting these formulas for $\bar{x}(t)$ and $\tilde{v}_t(t)$ in (A.14) gives the result for $m(t)$.
The formula for $m_0(t)$ is directly obtained from (A.11). Finally, the equation for $m_-(t)$ follows from $m(t)^2 = m_-(t)^2 + m_1(t)^2$. ♣

## A.5 GTLS Models

**Motivation for Conjecture 5.1.1.**
*1.* Concerning the existence of solutions, the example (5.2) indicates that the absence of solutions is indeed a non-generic case. The set of data for which the optimal approximation is of the form $[0, 0, c]$ is at most a two-dimensional manifold, hence non-generic in the neighbourhood of $[0, 0, 1]$. A further motivation is given in Section 5.5.
Concerning the uniqueness of solutions, we remark the following. Suppose there exist several GTLS models for data $w$. Then it requires only an arbitrarily small change of the data to make one of them the unique solution. Namely, if $\hat{w}$ is one of the solutions for $w$, then it is the only solution for $w - \delta(w - \hat{w})$ for all $\delta > 0$. This shows that the set of data for which there is at most one solution is a dense set in $(\mathsf{R}^q)^\top$. In order to show that this set is generic in the topological sense, it remains to prove that this set is also open. This is a weaker condition than 'having zero measure', so additional arguments are needed to prove algebraic genericity.
*2.* It is not hard to prove that the set of data for which the GTLS model is smaller than the tolerated size is non-generic. Suppose the GTLS model has less auxiliary inputs than tolerated, with approximation error $\tilde{w} = w - \hat{w}$. Then an arbitrary non-zero component can be absorbed in the approximation $\hat{w}$ as the effect of an additional auxiliary input, which would decrease the misfit. Hence $\tilde{w} = 0$, which means that the data itself is exact, an of less than tolerated rank. Next suppose that the GTLS model is of degree less than tolerated. This implies that all components of the approximation error $\tilde{w} = w - \hat{w}$ should be orthogonal to the set of all exponentials $z(t) = \lambda^t$, $\lambda \in \mathsf{R}$. Consequently, $\tilde{w}(1) = 0$, as otherwise there is a correlation with $\lambda^t$ for $\lambda$ sufficiently close to zero, and by induction it follows that $\tilde{w}(t) = 0$ for all $t \in T$.
As non-stabilizable systems of rank $m$ and degree $n$ form a non-generic set in $\mathsf{B}^{q,m,n}$ if $m > 0$, we conjecture that the GTLS model for generic data is also stabilizable.
*3.* This is based on the following considerations. Firstly, we conjecture that there is a finite number of locally optimal approximations, i.e., models that are optimal in a neighbourhood of the approximation, which will be made precise in Definition 5.1.2. It is likely that for generic data the GTLS criterion, which is polynomial in the parameters of an SR and ISR, cf. (5.6) and (5.17), has a finite number of locally optimal approximations that are continuous in the data. Now let $\Delta_\epsilon$ denote the minimal difference between the globally optimal misfit and the misfit in the non-global local optima. Then for all variations of the data of size less than $\Delta_\epsilon$ the global optimum is continuous in the data. ♣

**Proof of Theorem 5.3.1.** It suffices to prove the theorem for a minimal isometric SR. This can be seen as follows. If condition *2* holds for an arbitrary minimal SR, then it holds for all equivalent minimal SR's, as in the transformation $(S(A+BF)S^{-1}, SBR, (C+DF)S^{-1}, DR)$, the auxiliary input and state are linearly transformed to $R^{-1}(\hat{v} - F\hat{x})$ and $S\hat{x}$, see the proof of Proposition 3.1.4.
So let $(A, B, C, D)$ be a minimal ISR of $\mathcal{B}$, and let $\tilde{B}$, $\tilde{D}$ be defined as in Proposition 4.2.1. In the proof we will make use of the following relations holding on $T$, which follow from the projection algorithm in Section 4.3.

$$\textit{for } \hat{v}\textit{:} \quad x = A^\top \sigma x + C^\top w, \quad \hat{v} = B^\top \sigma x + D^\top w \tag{A.15}$$

$$\textit{for } \tilde{v}\textit{:} \quad x = A^\top \sigma x + C^\top w, \quad \tilde{v} = \tilde{B}^\top \sigma x + \tilde{D}^\top w \tag{A.16}$$

$$\textit{for } \hat{w}: \quad \sigma\hat{x} = A\hat{x} + B\hat{v}, \quad \hat{w} = C\hat{x} + D\hat{v} \tag{A.17}$$
$$\textit{for } \tilde{w}: \quad \sigma\tilde{x} = A\tilde{x} + \tilde{B}\tilde{v}, \quad \tilde{w} = C\tilde{x} + \tilde{D}\tilde{v} \tag{A.18}$$

Further, from equation (3.7) we obtain

$$\hat{x} = A^\top\sigma\hat{x} + C^\top\hat{w}, \quad \hat{v} = B^\top\sigma\hat{x} + D^\top\hat{w}, \quad 0 = \tilde{B}^\top\sigma\hat{x} + \tilde{D}^\top\hat{w} \tag{A.19}$$
$$\tilde{x} = A^\top\sigma\tilde{x} + C^\top\tilde{w}, \quad 0 = B^\top\sigma\tilde{x} + D^\top\tilde{w}, \quad \tilde{v} = \tilde{B}^\top\sigma\tilde{x} + \tilde{D}^\top\tilde{w} \tag{A.20}$$

In addition, we know that $\hat{x}(1) = x(1)$, $\hat{x}(N+1) = x(N+1)$ and $\tilde{x}(1) = \tilde{x}(N+1) = 0$. We first describe those optimality conditions that can be derived straightforwardly from the model improvement constructions in Theorem 5.2.1.

**Lemma A.5.1**

1. *Construction 1 gives no improvement if and only if* $\mathrm{cov}([\hat{x}, \hat{v}], \tilde{w}) = 0$.
2. *Construction 2 gives no improvement if and only if* $\mathrm{cov}(\hat{v}, [\sigma\tilde{x}, \tilde{w}]) = 0$.
3. *If Construction 3 gives no improvement, then* $\mathrm{cov}(\hat{v}, \tilde{v}) = 0$.

*Proof of the lemma.* We use the notation of Theorem 5.2.1.
*1.* As $\hat{w} = C\hat{x} + D\hat{v}$, clearly $\hat{w} \in \mathcal{E}$. Construction 1 does not yield an improvement iff the projection of $\tilde{w}$ onto $\mathcal{E}$ is zero, from which the result follows.
*2.* Construction 2 gives no improvement iff the projection of $\tilde{w}$ onto $\mathcal{F}$ is zero, cf. part *1.* This is equivalent to $\langle \tilde{w}, \hat{w}' \rangle = 0$ with $\hat{w}'(t) := D'\hat{v}(t) + \Sigma_{k=1}^{t-1} CA^{k-1}B'\hat{v}(t-k) + CA^{t-1}\hat{x}'_1$ for all $B'$, $D'$ and $\hat{x}'_1$. Using $\langle a, Mb \rangle = \left\langle M^\top a, b \right\rangle$, it follows that this is true if and only if $\langle \tilde{w}', \hat{v} \rangle = 0$ with $\tilde{w}'(t) := {D'}^\top\tilde{w}(t) + \Sigma_{k=1}^{N-t} {B'}^\top (A^\top)^{k-1} C^\top \tilde{w}(t+k) = {B'}^\top\sigma\tilde{x} + {D'}^\top\tilde{w}$, from which the result follows.
*3.* If $\mathrm{cov}(\hat{v}, \tilde{v}) \neq 0$, then construction 3 decreases $\|\tilde{v}\| = \|\tilde{w}\|$. □

Returning to Theorem 5.3.1, we first prove part *2.* From the lemma it follows that $\mathrm{cov}(\hat{v}, \tilde{v}) = 0$, and $\mathrm{cov}(\hat{v}, \tilde{x}) = \mathrm{cov}(\hat{v}, A^\top\sigma\tilde{x} + C^\top\tilde{w}) = 0$. As $\tilde{D}$ is injective, $\mathrm{cov}(\hat{x}, \tilde{v}) = 0$ if and only if $\mathrm{cov}(\hat{x}, \tilde{D}\tilde{v}) = 0$, which is equivalent to $\mathrm{cov}(\hat{x}, \tilde{w} - C\tilde{x}) = -\mathrm{cov}(\hat{x}, C\tilde{x}) = 0$. So in order to prove *2* it remains to show that $\mathrm{cov}(\hat{x}, \tilde{x}) = 0$. By assumption $\mathcal{B}$ is controllable, which implies that $(A, B)$ is a controllable matrix pair (cf. [46, Prop. IX.4]). So the condition $\mathrm{cov}(\hat{x}, \tilde{x}) = 0$ is equivalent to $\mathrm{cov}(\sigma\hat{x}, {B^\top A^\top}^k \sigma\tilde{x}) = 0$ for all $k \geq 0$. We prove this by induction. For $k = 0$, the lemma shows that $\mathrm{cov}(\sigma\hat{x}, B^\top\sigma\tilde{x}) = \mathrm{cov}(\sigma\hat{x}, -D^\top\tilde{w}) = 0$. Now suppose that $\mathrm{cov}(\sigma\hat{x}, {B^\top A^\top}^k \sigma\tilde{x}) = 0$ for $k \leq N$. Then $\mathrm{cov}(\sigma\hat{x}, {B^\top A^\top}^{N+1} \sigma\tilde{x}) = \mathrm{cov}(A\hat{x} + B\hat{v}, {B^\top A^\top}^{N+1} \sigma\tilde{x}) = \mathrm{cov}(A\hat{x}, {B^\top A^\top}^{N+1} \sigma\tilde{x}) = \mathrm{cov}(A\hat{x}, {B^\top A^\top}^N (\tilde{x} - C^\top\tilde{w})) = A\mathrm{cov}(\hat{x}, {B^\top A^\top}^N \tilde{x}) = 0$.
Concerning the equivalence of *1*, *2* and *3*, the implications *3* ⇒ *2* ⇒ *1* are trivial. Further, if *1* holds then $\mathrm{cov}(\hat{x}, \tilde{x}) = \mathrm{cov}(\sigma\hat{x}, \sigma\tilde{x}) = \mathrm{cov}(A\hat{x} + B\hat{v}, A\tilde{x} + \tilde{B}\tilde{v}) = \mathrm{cov}(A\hat{x}, A\tilde{x}) = A\mathrm{cov}(\hat{x}, \tilde{x})A^\top$. As $A$ is asymptotically stable, it follows that $\mathrm{cov}(\hat{x}, \tilde{x}) = 0$, so *1* implies *2.* Finally, *3* is easily derived from *2* by using equations (A.17) and (A.18). ♣

**Proof of Theorem 5.3.2.** Let $w$ be a given observation and let $\hat{w}$ be its optimal approximation in $\mathcal{B}$. Further let $(A, B, C, D)$ be a minimal SR of $\mathcal{B}$ with $A$ asymptotically stable, and write the approximation error as $\tilde{w}(t) := w(t) - \hat{w}(t) = w(t) - D\hat{w}(t) - \Sigma_{k=1}^{t-1} CA^{k-1}B'\hat{v}(t-k) + CA^{t-1}\hat{x}_1'$. We have to prove that $\mathcal{B}$ satisfies the optimality conditions if and only if the derivative of $\|\tilde{w}\|$ with respect to the parameters in $A, B, C, D$, $\hat{v}$ and $\hat{x}_1$ is zero. First we analyse the tangent space of $\tilde{w}$ with respect to these parameters. Let $\mathcal{E}$ and $\mathcal{F}$ be defined as in Theorem 5.2.1, and let $\mathcal{G} := \{\bar{w}T \to \mathsf{R}^q;\ \exists H \in \mathsf{R}^{n\times n}$ such that $\bar{w}(t) = \Sigma_{k=1}^{t-1} CA^{k-1}H\hat{x}(t-k)\}$ with $\hat{x}$ the state corresponding to $\hat{w}$.

**Lemma A.5.2** *The tangent space $\mathcal{T}$ of $\tilde{w} : T \to \mathsf{R}^q$ with $\tilde{w}(t) = \tilde{w}(t) := w(t) - \hat{w}(t) = w(t) - D\hat{w}(t) - \Sigma_{k=1}^{t-1} CA^{k-1}B'\hat{v}(t-k) + CA^{t-1}\hat{x}_1'$ is given by $\mathcal{T} = \mathcal{B}_T + \mathcal{E} + \mathcal{F} + \mathcal{G}$.*

*Proof of the lemma.* The tangent space is defined as the linear space that contains all partial derivatives. Note that $\tilde{w}$ is linear in $\hat{v}$, in $B$, in $C$ and in $D$. A change of $\hat{v}$ and $\hat{x}_1$ corresponds to adding $\bar{w} \in \mathcal{B}_T$ to $\tilde{w}$, a change of $C$ and $D$ to adding $\bar{w} \in \mathcal{E}$ to $\tilde{w}$ and a change of $B$, $D$ and $\hat{x}_1$ to adding $\bar{w} \in \mathcal{F}$ to $\tilde{w}$. It remains to prove that the derivatives of $\tilde{w}$ with respect to the parameters in $A$ span the space $\mathcal{G}$. For $H \in \mathsf{R}^{n\times n}$ let $x'$ be defined by $\sigma x' = (A - H)x' + B\hat{v}$ and let $w' := Cx' + D\hat{v}$. The corresponding error is $\tilde{w}' := w - w'$, so that the change in $\tilde{w}$ is given by and $\bar{w} = \tilde{w}' - \tilde{w} = \hat{w} - w' = C(\hat{x} - x') = C\bar{x}$ for $\bar{x} := \hat{x} - x'$. As $\sigma\bar{x} = A\hat{x} - (A - H)x' = A\bar{x} + Hx' = A\bar{x} + H\hat{x} - H\bar{x}$, ignoring the second order term $H\bar{x}$ for small $H$ gives the result. □

We will next prove the theorem by showing that both the optimality conditions and the stationarity condition are both equivalent to $\tilde{w} \perp \mathcal{T}$.
Stationarity is equivalent to the condition that $\lim_{\delta\to 0} \delta^{-1}\{\|\tilde{w} + \delta\bar{w}\| - \|\tilde{w}\|\} = 0$ for all $\bar{w} \in \mathcal{T}$. It is easily verified that this limit equals $\langle \tilde{w}, \bar{w}\rangle / \|\tilde{w}\|$, so stationarity is equivalent to $\tilde{w} \perp \mathcal{T}$.
Finally we show that $\tilde{w} \perp \mathcal{T}$ is equivalent to the optimality conditions. First suppose that the optimality conditions hold. As $\hat{w}$ is an optimal approximation within $\mathcal{B}_T$ there holds that $\tilde{w} \perp \mathcal{B}_T$. Further, Theorem 5.3.1.3 states that $\mathrm{cov}([\hat{v}, \hat{x}], \tilde{w}) = 0$, so that $\tilde{w} \perp \mathcal{E}$, and $\mathrm{cov}(\hat{v}, [\sigma\tilde{x}, \tilde{w}]) = 0$, so that the proof of Lemma A.5.1.2 shows that $\tilde{w} \perp \mathcal{F}$. Finally, for $\bar{w} \in \mathcal{G}$ given by $\bar{w}(t) := \Sigma_{k=1}^{t-1} CA^{k-1}H\hat{x}(t-k)$ we obtain by using (A.20) that $\langle \bar{w}, \tilde{w}\rangle = \left\langle \hat{x}, H^\top \sigma\tilde{x} \right\rangle = 0$, so $\tilde{w} \perp \mathcal{G}$. From Lemma A.5.2 it follows that $\tilde{w} \perp \mathcal{T}$.
Second, supposing that $\tilde{w} \perp \mathcal{T}$ we prove the optimality conditions. The fact that $\tilde{w} \perp \mathcal{E} + \mathcal{F}$ implies the conditions in Lemma A.5.1.1 and A.5.1.2, cf. the proof of that lemma. The condition in Lemma A.5.1.3 follows from that in A.5.1.2 by using (A.20). Further, the optimality conditions were derived from these conditions in the proof of Theorem 5.3.1. ♣

**Proof of Proposition 5.4.2.** First we prove that $\mathcal{Z}$ is a linear space. Observe that for every $z \in (\mathcal{B}_T)^\perp$ the optimal approximation of $\hat{w} + z$ within $\mathcal{B}_T$ is given by $\hat{w}$. Let $(\hat{v}, \hat{x})$ be the auxiliary input and state for $\hat{w}$ in a minimal SR of $\mathcal{B}$ , and let $(\tilde{v}_z, \tilde{x}_z)$ be defined analogously for $z$ in $(\mathcal{B}_T)^\perp$. As $(\hat{v}, \hat{x})$ is fixed, the condition in Theorem 5.3.1.2 for stationarity of $\mathcal{B}$ with respect to $\hat{w} + z$ consists of linear restrictions on $(\tilde{v}_z, \tilde{x}_z)$. As $\tilde{x}_z$ is a linear function of $\tilde{v}_z$, these conditions can be expressed as linear restrictions on $\tilde{v}_z$ alone. This shows that $\mathcal{B}$ is a stationary point for $\hat{w} + z$ if and only

if the auxiliary input $\tilde{v}_z$ is restricted to a linear subspace of time series $\mathsf{Z} \to \mathsf{R}^{q-m}$, from which the linearity of $\mathcal{Z}$ follows.
Next we show that the minimum in (5.12) is achieved by taking $\bar{w}_0 = \tilde{w} - \tilde{w}'$. As $w - \bar{w}_0 = \hat{w} + \tilde{w} - (\tilde{w} - \tilde{w}') = \hat{w} + \tilde{w}'$ and $\tilde{w}' \in \mathcal{Z}$, it follows that $\mathcal{B}$ is stationary for $w - \bar{w}_0$, by definition of $\mathcal{Z}$. Further, $\bar{w} \in (\mathcal{B}_T)^\perp$ is such that $\mathcal{B}$ is stationary for $w - \bar{w} = \hat{w} + \tilde{w} - \bar{w}$ if and only if $z := \tilde{w} - \bar{w} \in \mathcal{Z}$. Now the norm of $\bar{w} = \tilde{w} - z$ is minimized by taking $z = \tilde{w}'$, the orthogonal projection of $\tilde{w}$ on $\mathcal{Z}$, hence $\bar{w}_0 = \tilde{w} - \tilde{w}'$.
□ ♣

**Proof of Lemma 5.5.1.** The space $\mathcal{U}$ of unitary $(k \times k)$ matrices is a differentiable manifold in $\mathsf{R}^{k\times k}$ of dimension $\frac{1}{2}k(k-1)$, so the tangent space $T_M\mathcal{U}$ of $\mathcal{U}$ in a point $M$ a linear space of this dimension, contained in $\{MK;\ K \in \mathsf{R}^{k\times k}\}$. Now it is easily verified that $MK \in T_M\mathcal{U}$ if and only if $K + K^\top = 0$, by considering the fact that $(M + MK)^\top(M + MK) = I + K^\top + K + K^\top K$. ♣

**Correctness of Algorithm 4.** We will prove that

1. the formula (5.21) for the derivative of $H$ is correct
2. $\|\tilde{v} + H'\|$ is minimized for the variation $(-\Delta_M, -\bar{x}_{N+1})$ with $\Delta_M := (\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta\tilde{D})$ (hence step 3 is indeed a Gauss-Newton step).
3. if $\Delta_M = 0$ in Step 4, then the model is a stationary point
4. different variations of the form (5.19) are not equivalent if $(A, B, C, D)$ is minimal.

Then the convergence to stationary points follows from the general considerations underlying Gauss-Newton algorithms. The main issue is to prove that the second point holds true despite the fact that we have restricted the variations of ISR's to the form (5.19). Further, the fourth point implies that generically we cannot further reduce the number of variations.
*1.* From straightforward calculations it follows that for $\Delta_M := (\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}})$ the derivative [1] $H'(\Delta_M, \bar{x}_{N+1}) =: \bar{v}$ of $H$ is given by

$$\begin{aligned} \bar{x}(t) &= A^\top \bar{x}(t+1) + \Delta_A^\top x(t+1) + \Delta_C^\top w(t);\ \bar{x}(N+1) = \bar{x}_{N+1} \\ \bar{v}(t) &= \tilde{B}^\top \bar{x}(t+1) + \Delta_{\tilde{B}}^\top x(t+1) + \Delta_{\tilde{D}}^\top w(t) \end{aligned} \tag{A.21}$$

Substituting (5.19) and using (4.15) gives the result.

*2.* Define $\tilde{V} := \mathrm{im}H$, then $\mathrm{im}H' = T_{\tilde{v}}\tilde{V}$. Clearly, $\|\tilde{v} + H'(\Delta_M, \bar{x}_{N+1})\|$ is minimized iff $H'(\Delta_M, \bar{x}_{N+1}) = -\tilde{v}^*$, with $\tilde{v}^*$ the orthogonal projection of $\tilde{v}$ onto $T_{\tilde{v}}\tilde{V}$. Notice that $H'(\Delta_M, \bar{x}_{N+1}) = \tilde{v}'$, which is the orthogonal projection of $\tilde{v}$ onto $\mathcal{T}$. We will prove that

$$T_{\tilde{v}}\tilde{V} = \mathcal{T} + \mathcal{R}, \text{ with } \mathcal{R} \perp \tilde{v}. \tag{A.22}$$

Then the orthogonal projection $\bar{v}$ onto $T_{\tilde{v}}\tilde{V}$ and $\mathcal{T}$ coincide, i.e., $\tilde{v}' = \tilde{v}^*$, so that subtracting the variation $\Delta_M$ corresponding to $\tilde{v}'$ is indeed optimal. So it remains

[1] By definition, the derivative of $H$ is given by $H' : T_{(M,x)}\mathcal{I} \to \mathsf{R}^{p\times N}$ that assigns to every direction $(\Delta_M, \bar{x}(N+1))$ the directional derivative of $H$, i.e., $H'_{(M,x)}(\Delta_M, \bar{x}(N+1)) := \frac{\partial H(c(t))}{\partial t}|_{t=0}$ with $c : (-\epsilon, \epsilon) \to \mathcal{U}_{n+q} \times \mathsf{R}^n$ a differentiable curve with $c(0) = (M, x)$ and $c'(0) = (\Delta_M, \bar{x}(N+1))$.

to prove (A.22). Therefore we compare the definitions of $T_{\tilde{v}}\tilde{V}$ and $\mathcal{T}$. By definition, $T_{\tilde{v}}\tilde{V} = \mathrm{im}H' = \{H'(\Delta_M, \bar{x}_{N+1});\ \Delta_M \in T_M\mathcal{I}, \bar{x}_{N+1} \in \mathsf{R}^n\}$. According to Lemma 5.5.1, $T_M\mathcal{I}$ consists of variations $(\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}})$ of the following form. In order to preserve the symmetry in the formulas, we also describe the corresponding variations of $B$ and $D$.

$$\begin{bmatrix} \Delta_A & \Delta_B & \Delta_{\tilde{B}} \\ \Delta_C & \Delta_D & \Delta_{\tilde{D}} \end{bmatrix} = \begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix} \begin{bmatrix} X & L & P \\ -L^\top & Y & Q \\ -P^\top & -Q^\top & R \end{bmatrix}, \tag{A.23}$$

with $X$, $Y$ and $R$ anti-symmetric, i.e., $X = -X^\top$ etc. Notice that $\mathcal{T}$ is defined similarly to $T_{\tilde{v}}\tilde{V}$, but with the variations restricted to (5.19). So it remains to show that by taking the diagonal blocks in (A.23) equal to zero we only remove some directions $\mathcal{R}$ in $T_{\tilde{v}}\tilde{V}$ that are orthogonal to $\tilde{v}$. Such directions are irrelevant in the Gauss-Newton algorithm, as explained before. In fact, we will show that, roughly speaking, $R$ corresponds to a unitary basis transformation of the auxiliary input $\tilde{v}$ , leaving $\|\tilde{v}\|$ unchanged, while the effect of $X$ in (A.23) can always be compensated by an appropriate unitary basis transformation of the state space.
First we analyse the effect of $R$ in (A.23), i.e., we consider variations of the form $\Delta_M = (\Delta_A, \Delta_{\tilde{B}}, \Delta_C, \Delta_{\tilde{D}}) = (0, \tilde{B}R, 0, \tilde{D}R)$ with $R$ anti-symmetric. This is nothing else than a variation corresponding to a unitary basis transformation of the auxiliary input $\tilde{v}$, cf. Proposition 3.2.6.2 and Lemma 5.5.1. For this variation (A.21) gives $H'(\Delta_M, 0) = \bar{v}$ with $\bar{v}(t) = R^\top(\tilde{B}x(t+1) + \tilde{D}w(t)) = R^\top\tilde{v}(t)$. Now define

$$\mathcal{R} := \{R^\top\tilde{v}; R + R^\top = 0\}. \tag{A.24}$$

Then $\mathcal{R} \perp \tilde{v}$, as $\left\langle R^\top\tilde{v}, \tilde{v} \right\rangle = \langle \tilde{v}, R\tilde{v} \rangle = -\left\langle \tilde{v}, R^\top\tilde{v} \right\rangle = 0$. This reflects the fact that a unitary basis transformation of $\tilde{v}$ has no effect on $\|\tilde{v}\|$.
Finally we will prove that without loss of generality we can take $X = 0$ in (A.23), by showing that all elements in $T_{\tilde{v}}\tilde{V}$ can be generated with $X = 0$. Notice that a basis transformation of the state space does not change the value of $H$, hence $H'$ is a singular mapping as $H'(\Delta_M, \bar{x}_{N+1}) = 0$ for $\Delta_M$ a variation corresponding to a unitary state space transformation and $\bar{x}_{N+1}$ the corresponding change in the end state. The structure of these variations is determined as follows. To preserve symmetry in the formulas, again we also consider variations of $B$ and $D$. Define $\bar{B} := [B\ \tilde{B}]$ and $\bar{D} := [D\ \tilde{D}]$, and define $\Delta_{\bar{B}}$ and $\Delta_{\bar{D}}$ accordingly. The linearization of the mapping

$$U \to (UAU^\top, UB, CU^\top, D)$$

around $U = I_p$ is given by

$$I + \Delta_U \to (\Delta_U A + A\Delta_U^\top, \Delta_U B, C\Delta_U^\top, 0),$$

so it follows from Lemma 5.5.1 that unitary state space transformations correspond to variations $(\Delta_A, \Delta_{\bar{B}}, \Delta_C, \Delta_{\bar{D}}) = (\Delta_U A - A\Delta_U, \Delta_U B, -C\Delta_U, 0)$ with $\Delta_U + \Delta_U^\top = 0$. Rewriting this according to (A.23) gives

$$\begin{bmatrix} \Delta_A & \Delta_{\bar{B}} \\ \Delta_C & \Delta_{\bar{D}} \end{bmatrix} = \begin{bmatrix} A & \bar{B} \\ C & \bar{D} \end{bmatrix} \begin{bmatrix} A^\top\Delta_U A - \Delta_U & A^\top\Delta_U\tilde{B} \\ A^\top\Delta_U\tilde{B} & \bar{B}^\top\Delta_U\tilde{B} \end{bmatrix}. \tag{A.25}$$

As transformations of the state space do not affect $H$, it follows that these variations belong to the kernel of $H'$, taking $\bar{x}_{N+1} = \Delta_U x(N+1)$. This can also be proved

directly, by substituting (A.25) in (A.21), which gives $\bar{x}(t) = \Delta_U x(t)$ and $\bar{v}(t) = 0$ for $t \in [1, N]$.
Now define for a variation of the form (A.23), $\Delta_u$ such that

$$X - A^\top \Delta_u A - \Delta_u = 0;\ \Delta_u + \Delta_u^\top = 0 \tag{A.26}$$

and subtract the corresponding variation (A.25). This results in an equivalent variation with zero value for the $X$-block in (A.23). Notice that for minimal ISR's, $A$ is asymptotically stable, cf. Proposition 3.2.6.1. Then the solution is given by $\Delta_U := -\Sigma_{k=1}^{\infty} A^k K_1 A^{\top k}$.

So there is only one detail left, concerning the case that $A$ is not asymptotically stable. Notice that there might be no solution for (A.26) in this case. However, then $(A, B, C, D)$ is not minimal, and there is an additional source of non-uniqueness beside the choice of the unitary transformations as described in Proposition 3.2.6.2, which can be seen as follows. As $A$ is stable, cf. Proposition 3.2.6.1, it must have some poles on the unit circle if it is not also asymptotically stable. Together with the isometry property this implies that the matrix in (5.18) can be transformed by a unitary state space transformation to the form

$$\begin{bmatrix} A_0 & 0 & 0 \\ 0 & A_- & \tilde{B}' \\ 0 & C' & \tilde{D}' \end{bmatrix}$$

with $A_0$ unitary and $A_-$ asymptotically stable. As the state components corresponding to $A_0$ are completely ineffective, we can replace $A_0$ by an arbitrary unitary matrix, or, equivalently, we can apply the transformation $A_0 L$ with $L$ unitary. Consequently, in addition to (A.25), also the variations corresponding to this transformation are contained in the kernel of $H'$. Using this extra freedom, we can to obtain an equivalent variation with $X = 0$, even if $A$ is not asymptotically stable. As this concerns only a non-generic case, we leave the details to the reader.

*3.* If $\Delta_m = 0$, it follows that also $\bar{x}_{N+1} = 0$, as $x_{N+1}$ is optimal for $(A, \tilde{B}, C, \tilde{D})$, cf. Step 1. This implies that $\|\tilde{v} + H'\|$ is minimized for $(\Delta_M, \bar{x}_{N+1}) = (0, 0)$, cf. part 1, which means that the derivative of $\|H\|$ is zero in $\tilde{v}$ , hence $((A, \tilde{B}, C, \tilde{D}), x_{N+1})$ is a stationary point of $H$.

*4.* It is easily verified that the reduction of the number of parameters in the variations (A.23) by taking the anti-symmetric diagonal blocks $X = 0$ and $R = 0$ equals $\frac{1}{2}n(n-1)+\frac{1}{2}p(p-1)$. This is exactly the number of free parameters in the equivalent transformation of $(A, \tilde{B}, C, \tilde{D})$ as described in Proposition 3.2.6.2. As for minimal ISR's this is the only source of non-uniqueness, it follows that different variations in (5.19) can not be equivalent. ♣

**Proof of Lemma 5.6.2.** A minimal state is both past and future induced, which follows from [45, 2.4.3]. From the existence of minimal SR's it is easily verified that an $n$-dimensional state it suffices to take $k = n$.
Another way to obtain the result is to consider the description of systems by means of difference equations, cf. Section 2.4.2. A construction of the state in terms of linear functions of past and future is given in the proof of Proposition 3.1.2. This shows that the state is both a linear function of the finite past and future. ♣

**Correctness of Algorithm 5.** First we prove that $(z_i^-, z_i^+) := (U_i^\top V_-^\top, V_i^\top V_+^\top)$ denotes the $i$-th pair of canonical variables of $(\mathcal{H}_\theta^-, \mathcal{H}_\theta^+)$ as defined in Section 5.6.3, Step 1. Notice that $V_-^\top$ and $V_+^\top$ form an orthonormal basis for ($\mathcal{H}_\theta^-$ and $\mathcal{H}_\theta^+$) respectively, hence by definition the canonical variables take the form $(z^-, z^+) = (u_-^\top V_-^\top, u_+^\top V_+^\top)$ with $u_-, u_+$ vectors of unit norm. Now the result follows from the properties of the SVD.
Further, it is easily verified that $E_i$ denotes the total energy of the $i$-th pair of canonical variables, so that $\bar{x}_i$ is defined accordingly to Section 5.6.3, Step 1.
The verification of the second step is an easy exercise.

♣

## A.6 Applications and Extensions

**Proof of Proposition 6.2.3.** *1.* This follows from the fact that the misfit of a model for $w - c$ is given by $\|\tilde{v}\|$ with $\tilde{v}$ defined as the effect of $w - c$ in (4.15).
*2.* Trivial
*3.* Redefine $H$, defined in (5.16), as a function from ISR's, final states and constant time series $c$ to $\tilde{v}$, with $\tilde{v}$ the effect of $w - c$ in (4.15), so that $\|\tilde{v}\|$ denotes the misfit of $w - c$. Now the extension of $\mathcal{T}$ corresponds to the effect of varying the constants, from which the result follows. ♣

**Correctness of Definition 6.3.2.** Consider the sequence $d_t := \dim(\mathcal{B}_{[t_0,t]})$ for $t_0 \in \mathsf{Z}$ and $t \geq t_0$, and its difference $\Delta d_t := d_t - d_{t-1}$ denoting the increase of dimension at time $t$ for initial time $t_0$. We first show that $\Delta d_t \leq \Delta d_{t-p}$, analogous to the proof of correctness of Definition 2.1.6. As $\Delta d_{t-p}$ corresponds to the degree of freedom in $w(t-p)$ given $\{w(t_0), \ldots, w(t-p-1)\}$, or equivalently of $w(t)$ given $\{w(t_0+p), \ldots, w(t-1)\}$, it is larger than $\Delta d_t$, which is the degree of freedom of $w(t)$ given $\{w(t_0), \ldots, w(t-1)\}$. Clearly $0 \leq \Delta d_t \leq q$, so the $p$ sequences $\{\Delta d_{k+pj}\}_{j\in\mathsf{N}}$ for $k = 1, \ldots, p$ must reach their limit values $m_k$ within finite time. These limit values do not depend on the initial time $t_0$. Namely, let $T_{lim}$ denote the smallest time instant for which $\Delta d_t$ has reached its limit value for each period, so $\Delta d_{k+jp} = m_k$. Then for $t \geq T_{lim}$, $\dim(\mathcal{B}_{[1,t]}) - \dim(\mathcal{B}_{[1,t-1]}) = \dim(\mathcal{B}_{[-\ell,t]}) - \dim(\mathcal{B}_{[-\ell,t-1]})$ for all $\ell \in \mathsf{N}$. Now the correctness of the definition follows in the same way as for Definition 2.1.6. ♣

**Proof of Proposition 6.3.3.**
*1.* Define the *period behaviour* of $\mathcal{P}$ as

$$\begin{aligned}\mathcal{B}_{\mathcal{P}} := \{&\hat{w}_p : \mathsf{Z} \to \mathsf{R}^{\pi q};\\ &w_p(t) = \mathrm{col}(w(\pi t + 1), \ldots, w(\pi t + \pi)) \text{ for some } \hat{w} \in \mathcal{P}.\end{aligned} \tag{A.27}$$

$\mathcal{B}_{\mathcal{P}}$ is a linear, shift-invariant system, so it can be described by a set of linear, time-invariant difference equations of finite lag, cf. Proposition 2.4.4. These equations translate to a set of periodic difference equations for $\mathcal{P}$, i.e.,

$$\mathcal{P} = \{w : \mathsf{Z} \to \mathsf{R}^q;\ R_d^k w(t) + \ldots + R_0^k w(t-d) = 0,\ \text{for } t = k + jp \text{ for all } j \in \mathsf{Z}\}, \tag{A.28}$$

with $R_l^k \in \mathsf{R}^{p_k \times q}$ the $l$-th matrix coefficient of the difference equation that applies at period $k$, and $p_k \in \mathsf{N}$.
*2.* An obvious extension of the construction of SR's from difference equations in the proof of Proposition 3.1.2 to the periodic case gives the result. We remark that the SR in period $k$ does not only depend on the difference equations that apply at that period, so the translation can not be performed period-wise. Further, the class of systems described by an PSR with $m_k$ auxiliary inputs and $n_k$ states in period $k$ coincides with the class of periodic systems with rank and degree in period $k$ at most $m_k$ and $n_k$ (This can be derived analogously to Proposition 3.1.2.3, but a precise proof falls outside the scope of this monograph). Hence in a minimal PSR, the number of auxiliary inputs and states in each period are equal to the period rank and degree respectively.
*3.* A precise proof falls outside our scope, we only sketch the idea. First we remark that stabilizability is defined exactly as for the time-invariant case, see Definition

3.2.3. Let $(A_t, B_t, C_t, D_t)$ denote a minimal PSR of a periodic system $\mathcal{P}$ of period $\pi$, so $A_t = A_{t+\pi}$ etc. Similar to the time-invariant case, equivalent PSR's are given by the transformations

$$\begin{bmatrix} A_t & B_t \\ C_t & D_t \end{bmatrix} \rightarrow \begin{bmatrix} S_{t+1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_t & B_t \\ C_t & D_t \end{bmatrix} \begin{bmatrix} S_t^{-1} & 0 \\ F_t S_t^{-1} R_i & \end{bmatrix}, \tag{A.29}$$

with $S_t = S_{t+\pi}$ a basis transformation of the state, $R_t = R_{t+\pi}$ a basis transformation of the auxiliary inputs, and $F_t = F_{t+\pi}$ a feedback transformation as in the time-invariant case.
An PISR is obtained by determining a transformation such that (3.7) holds in each period. This leads to following the periodic algebraic Riccati equation for $K_t := S_t^\top S_t$:

$$\begin{aligned} &K_t = A_t^\top K_{t+1} A_t \\ &-(B_t^\top K_{t+1} A_t + D_t^\top C_t)^\top (B_t^\top K_{t+1} B_t + D_t^\top D_t)^{-1} (B_t^\top K_{t+1} A_t + D_t^\top C_t) + C_t^\top C_t. \end{aligned} \tag{A.30}$$

This is the dual form of the periodic Riccati equation discussed in [6]. This paper contains some results that might be helpful in proving the solution method given below, but we leave this as a topic for further research. We obtain a positive definite periodic solution for $K$ as follows. Define $K'_0 := I_n$, and determine $K'_t$ for $t < 0$ according to (A.30). Then $K'_t$ converges to a periodic sequence, which give the periodic values for $K$. The equations for $S_t, R_t, F_t$ are given by

$$S_t^\top S_t = K_t \tag{A.31}$$

$$R_t R_t^\top = (B_t^\top K_{t+1} B_t + D_t^\top D_t)^{-1} \tag{A.32}$$

$$F_t = -(B_t^\top K_{t+1} B_t + D_t^\top D_t)^{-1} (B_t^\top K_{t+1} A_t + D_t^\top C_t). \tag{A.33}$$

Applying the periodic transformations corresponding to $S_t$, $R_t$ and $F_t$ yields an PISR. ♣

# Appendix B

# Matlab Programs

This appendix contains an implementation of Algorithms 1–5 in Matlab[1]. In the programs we follow the following conventions:

- Time series are represented by *tall* matrices, which is standard in Matlab. So the data $w$ is an $N \times q$ matrix, etc.
- Systems are represented by one square unitary matrix

$$M := \begin{bmatrix} A & B & \tilde{B} \\ C & D & \tilde{D} \end{bmatrix}, \tag{B.1}$$

  with $(A, B, C, D)$ representing the system $\mathcal{B}(A, B, C, D)$, and with $\tilde{B}$ and $\tilde{D}$ corresponding to its orthogonal complement, cf. Proposition 4.2.1.
- Suffixes -s and -e are used instead of hats and tildes, so $\tilde{D}$ is in the programs `De`, and so on. Further, $\hat{v}$ and $\tilde{v}$ are combined into one time series `dv` = [`vs`, `ve`].

This appendix is organized as follows. First we give some routines for the transformations between SR's, ISR's and i/s/o-representations, and between ISR's and $M$ in (B.1) (we remark that transformations between i/s/o-representations and descriptions in terms of difference equations can be carried out by using the standard Matlab programs `ss2tf` and `tf2ss`). Then we give the implementation of Algorithms 1–5. Algorithms 4 and 5 are combined into one GTLS algorithm, followed by subroutines for the model improvement constructions and a Gauss-Newton step. We conclude by a list of small subroutines that are not standard in Matlab or in the Matlab Control Toolbox, in alphabetical order.

## B.1 Representation Transformations

**SR2ISR**: transforms an SR to an equivalent ISR, cf. Proposition 3.2.5.

```
function [A,B,C,D,Be,De]=SR2ISR(A,B,C,D)
%partly derived from programs of C. Heij, 1989
H=-(D'*D)\(D'*C); A=A+B*H; C=C+D*H;
[F,K]=dlqr(A,B,C'*C+10*eps*eye(length(A)),D'*D);
R=sqrtm(inv(B'*K*B+D'*D)); S=sqrtm(K); iS=inv(S);
```

[1] Matlab is a program from The Mathworks, Inc., USA

```
A=real(S*(A-B*F)*iS); B=real(S*B*R);
C=real((C-D*F)*iS);   D=real(D*R);
F=F-H; F=real(F); K=real(K); R=real(R); S=real(S);
if nargout>4, [u,s,v]=svd([A,B;C D]);
[q,n]=size(C); [n,m]=size(B);
Be=u(1:n,n+m+1:n+q); De=u(n+1:n+q,n+m+1:n+q); end
```

**ISO2SR**: transforms an i/s/o-representation to an SR, cf. Proposition 3.4.2.

```
function [A,B,C,D]=ISO2SR(a,b,c,d)
[n,m]=size(b); A=a; B=b; C=[zeros(m,n);c]; D=[eye(m);d];
```

**SR2ISO**: transforms an SR to an i/s/o-representation, cf. Proposition 3.4.3.

```
function [a,b,c,d]=SR2ISO(A,B,C,D)
[n,m]=size(B); [q,m]=size(D);  p=q-m;
Cu=C(1:m,:); Du=D(1:m,:); invDu=inv(Du);
Cy=C(m+1:q,:); Dy=D(m+1:q,:);
a=A-B*invDu*Cu; b=B*invDu; c=Cy-Dy*invDu*Cu; d=Dy*invDu;
```

**Join**: determines $M$ in (B.1) for a given ISR $(A, B, C, D)$.

```
function M=join(A,B,C,D);
[q,n]=size(C); [n,m]=size(B);
M=[A B;C D]; [U,S,V]=svd(M); M=[M U(:,n+m+1:n+q)];
```

**Split**: splits $M$ in (B.1) into system matrices.

```
function [A,B,C,D,Be,De]=split(M,q,m);
[k,z]=size(M); n=k-q;
A=M(1:n,1:n); B=M(1:n,n+1:n+m); Be=M(1:n,n+m+1:n+q);
C=M(n+1:n+q,1:n); D=M(n+1:n+q,n+1:n+m); De=M(n+1:n+q,n+m+1:n+q);
```

## B.2 Algorithms 1–5 in Matlab

**Alg1**: Optimal Approximation in Given System, Section 4.3.

```
function [misfit,ws,we,dv,xs,xe,xf]=Alg1(w,M,m)
%initialization
  [N,q]=size(w); [A,B,C,D,Be,De]=split(M,q,m); [n,n]=size(A);
%Step 1: use 'join' to obtain parameter M as in (B.1)
%Step 2
  %determine optimal final state xf:=x(N+1)
    dv=dlsim(A',C',Be',De',rev(w));
    M=obsmat(A',Be',N); [h,h,xf]=orthproj(M,dv(:));  xf=-xf;
  %evaluate (4.15) (dv=:[vs ve])
    [dv,x]=dlsim(A',C',[B Be]',[D De]',rev(w),xf);
    dv=rev(dv); x=rev(x);
    xi=A'*x(1,:)'+C'*w(1,:)'; x=[xi';x(1:N-1,:)];
%Step 3
  [ws,xs]=dlsim(A,B,C,D,dv(:,1:m),xi);
  we=w-ws; xe=x-xs; misfit=eucln(we);
  %then [we,xe]=dlsim(A,Be,C,De,dv(:,m+1:q));
```

**Alg2**: Recursive Approximation within a Given System, Section 4.4.

```
function [x,ws,recm,recm0]=Alg2(w,M,m)
%x is the end state of the optimal approximation, x^_t+1(t)
%ws is the past induced optimal approximation, w^_t(t)
%recm is the increase of misfit, m(t), and recm0 is m_0(t)
%initialization
[N,q]=size(w); [A,B,C,D,Be,De]=split(M,q,m); [n,n]=size(A); w=w';
%Step 1
  We=0; x(:,1)=zeros(n,1); %We=W~0, x(:,1)=x^ 0|-1
%Step 2
for k=1:N,
  Wn=A*We*A'+Be*Be'; %W~k (where We denotes W(k-1))
  Fe=A*We*C'+Be*De'; %F~(k-1)
  Ge=C*We*C'+De*De'; %G~(k-1)
  He=Ge-Fe'*pinv(Wn)*Fe; %Hek
%Step 3
  epsil=w(:,k)-C*x(:,k);
  x(:,k+1)=A*x(:,k)-pinv(Wn)*Fe*epsil; %x^(k+1)|k
  ws(:,k)=C*x(:,k)+(eye(q)-He)*epsil; %w^k|k
  recm(:,k)=eucln(sqrtm(He)*epsil);
  recm0(:,k)=eucln(He*epsil); We=Wn;
end
  w=w'; ws=ws'; x=x'; x(1,:)=[]; %then x starts at $t=1$.
```

**GTLS**: GTLS algorithm, by Gauss-Newton (Algorithm 4, Section 5.5) or by the model improvement constructions (Algorithm 3,Section 5.3).

```
function [M,misfits]=GTLS(w,M0,m,tol,maxstep,option);
%tol: stop criterion; maxstep: bound on number of iterations
%option: if zero: Gauss Newton (default), else Model Improvement
%initialization
  [N,q]=size(w); [h,h]=size(M0); n=h-q; M=M0;
  if nargin==5, option=0; end, impr=tol+1;cntr=1; %counter
  [misfit,ws,we,dv,xs]=Alg1(w,M0,m); misfits=misfit;
%Iterations
  while impr>tol,
    if option==0, [M,misfit]=GNstep(w,M,m); %Algorithm 4
    else          M=mic1(w,M,m);            %Algorithm 3, Step 2
                  M=mic2(w,M,m);            %Algorithm 3, Step 3
                  M=mic3(w,M,m);            %Algorithm 3, Step 4
                  misfit=Alg1(w,M,m); end
    misfits=[misfits;misfit];%store sequence of misfits
    impr=misfits(cntr)-misfit %decrease of misfit
    cntr=cntr+1, if cntr>maxstep, 'maxstep reached', return; end
  end % of loop
```

**mic1**: Model Improvement Construction 1, Theorem 5.2.1.1.

```
function M=mic1(w,M,m);
[misfit,ws,we,dv,xs]=Alg1(w,M,m); [N,q]=size(w); [N,n]=size(xs);
[A,B,C,D]=split(M,q,m);
[we,ws,Q]=orthproj([xs dv(:,1:m)],w); C=Q(1:n,:)'; D=Q(n+1:n+m,:)';
[A,B,C,D]=SR2ISR(A,B,C,D); M=join(A,B,C,D);
```

**mic2**: Model Improvement Construction 2, Theorem 5.2.1.2.

```
function M=mic2(w,M,m);
[misfit,ws,we,dv,xs]=Alg1(w,M,m); [N,q]=size(w); [N,n]=size(xs);
[A,B,C,D,Be,De]=split(M,q,m); RB=[]; RD=[];
for i=1:m
  for k=1:n
    ek=[zeros(k-1,1);1;zeros(n-k,1)];
    rb=dlsim(A,ek,C,zeros(q,1),dv(:,i)); RB=[RB rb(:)];
  end
  for k=1:q
    rd=[zeros(N,k-1) dv(:,i) zeros(N,q-k)]; RD=[RD rd(:)];
  end
end
RX0=obsmat(A,C,N); [wel,ws,coef]=orthproj([RB RD RX0],w(:));
we=zeros(w); we(:)=wel;
B=zeros(n,m); B(:)=coef(1:n*m);
D=zeros(q,m); D(:)=coef(n*m+1:n*m+m*q);
[A,B,C,D,Be,De]=SR2ISR(A,B,C,D); M=join(A,B,C,D);
```

**mic3**: Model Improvement Construction 3, Theorem 5.2.1.3.

```
function M=mic3(w,M,m);
[misfit,ws,we,dv]=Alg1(w,M,m); [N,q]=size(w); [h,h]=size(M);
n=h-q; [A,B,C,D,Be,De]=split(M,q,m);
[U,S,V]=svd(dv,0); M=M*[eye(n) zeros(n,q);zeros(q,n) V];
```

**GNstep**: Gauss-Newton step, one iteration of Algorithm 4.

```
function [M,misfit,RR]=gnstep(w,M,m);
%Initialization
[N,q]=size(w);[A,B,C,D,Be,De]=split(M,q,m);
p=q-m; [n,n]=size(A); Mold=M; RL=[]; RP=[]; RQ=[];
%Step1
  [oldmf,ws,we,dv,xs,xe]=Alg1(w,M,m); x=xs+xe; %dv=[vs ve]
%Step 2
  %regressors for L coefficient (p columns are put below each other)
    for j=1:m, for k=1:n,
      ek=[zeros(k-1,1);1;zeros(n-k,1)];
      r=dlsim(A',ek,Be',zeros(p,1),rev(dv(:,j))); r=rev(r);
      RL=[RL r(:)];
    end, end
  %regressors for P coefficient
    for j=m+1:q, for k=1:n,
      ek=[zeros(k-1,1);1;zeros(n-k,1)];
      r=dlsim(A',ek,Be',zeros(p,1),rev(dv(:,j))); r=rev(r);
      ej=[zeros(j-m-1,1);1;zeros(q-j,1)];r=r-x(:,k)*ej';RP=[RP r(:)];
    end, end
  %regressors for Q coefficient
    for j=1:m, for k=1:p,
      r=[zeros(N,k-1) -dv(:,j) zeros(N,p-k)]; RQ=[RQ r(:)];
    end, end
  %regressors corresponding to varying the end state x(N+1)
```

```
    RX0=obsmat(A',Be',N,0);
  RR=[RL RP RQ RX0]; %spans the relevant part of the tangent space
%Step 3 and 4
  ve=dv(:,m+1:q); [dvo,dvp,R]=orthproj(RR,ve(:));
  L=zeros(n,m); L(:)=-R(1:n*m); P=zeros(n,p); P(:)=-R(n*m+1:n*q);
  Q=zeros(m,p); Q(:)=-R(n*q+1:n*q+m*p);
  DeltaA=-B*L'-Be*P'; DeltaB=A*L-Be*Q';
  DeltaC=-D*L'-De*P'; DeltaD=C*L-De*Q';
  f=2; misfit=oldmf+1; %initialization for loop
  while (misfit>oldmf) & f>2^(-10), %min. step length
    f=f/2; %first value is 1;
    Ai=A-f*DeltaA; Bi=B-f*DeltaB; Ci=C-f*DeltaC; Di=D-f*DeltaD;
    [Ai,Bi,Ci,Di]=SR2ISR(Ai,Bi,Ci,Di); M=join(Ai,Bi,Ci,Di);
    misfit=Alg1(w,M,m)
  end
  if misfit>oldmf, 'NO IMPROVEMENT!', M=Mold; misfit=oldmf; end
```

**Alg5**: Initial models by canonical correlation analysis, Section 5.6.

```
function [A,B,C,D,cc]=Alg5(w,m,n,k,theta)
%initialization
  [N,q]=size(w);
%Step 1
  %construction of Hmin and Hplus defined by (5.30)
    Hmin=w(1:N-2*k+1,:)'; Hplus=w(k+1:N-k+1,:)';
    for i=2:k, Hmin=[Hmin;w(i:N-2*k+i,:)'];
               Hplus=[Hplus;w(k+i:N-k+i,:)']; end
  %computation of SVD
    [Vmin,Smin,Umin]=svd(Hmin',0);     %then Hmin=Umin*Smin*Vmin
    [Vplus,Splus,Uplus]=svd(Hplus',0); %then Hplus=Uplus*Splus*Vplus
  %removal of small energy directions
    passmin=[diag(Smin)>sqrt(theta)];
    passplus=[diag(Splus)>sqrt(theta)];
    Vmin=Vmin(:,passmin); Vplus=Vplus(:,passplus);
    Umin=Umin(:,passmin); Uplus=Uplus(:,passplus);
    Smin=Smin(passmin,passmin), Splus=Splus(passplus,passplus)
  %canonical variables and energies
    [U,S,V]=svd(Vmin'*Vplus); cc=diag(S);
  %rows of U'*Vmin' and V'*Vplus' are canonical variables
  %with canonical correlations given by diagonal of S
    iSmin=inv(Smin); iSplus=inv(Splus); xbar=[];
alpha=U'*iSmin*Umin'; beta=V'*iSplus*Uplus';
    for i=1:n,
      Ei=1/(eucln(U(:,i)'*iSmin)^2+eucln(V(:,i)'*iSplus)^2);
        xbar=[xbar (Vmin*U(:,i)+Vplus*V(:,i))*sqrt(Ei)];
    end
    %then xbar(:,i) is appr. state at t=k+i for t=k+1,...,N-k+1
%Step 2
  %Construction of M defined by (5.31)
  M=[xbar(1:N-2*k,:) xbar(2:N-2*k+1,:) w(k+1:N-k,:)]';
  %Approximation by rank n+m matrix
```

```
  [V,S,U]=svd(M',0); %then M=U*S*V^T
  Ur=U(:,1:n+m); Vr=V(:,1:n+m); Sr=S(1:n+m,1:n+m);
  %Computation of SR
  K=Ur(1:n,:); [Ubar,Sbar,Vbar]=svd(K); Sbar=Sbar(:,1:n);
  Z=Vbar*[inv(Sbar)*Ubar' zeros(n,m);zeros(m,n) eye(m)]
  UrZ=Ur*Z; A=UrZ(n+1:2*n,1:n);     B=UrZ(n+1:2*n,n+1:n+m);
            C=UrZ(2*n+1:2*n+q,1:n); D=UrZ(2*n+1:2*n+q,n+1:n+m);
```

## B.3 Auxiliary Programs

**eucln**: Euclidian norm.

```
function n=eucln(w)
n=sqrt(sum(diag(w'*w)));
```

**orthproj**: Orthogonal projection.

```
function [yo,yp,alpha]=orthproj(x,y)
%yp=x*alpha is the orthogonal projection of y onto x, yo=y-yp;
%It is assumed that x is tall.
[rx,cx]=size(x); [u,s,v]=svd(x,0); alpha=v*inv(s)*u'*y;
yp=u*(u'*y); yo=y-yp;
```

**rev**: Time reversion.

```
function wr=rev(w)
[N,q]=size(w); wr=w(N:-1:1,:);
```

**obsmat**: Observability matrix.

```
function M=obsmat(A,C,t,bw)
%effect of initial state on system trajectories (bw:backwards)
%i-th column of M contains the effect of x(0)=e_i,
%in a 'long vector' (with components below each other)
[q,n]=size(C);
for i=1:n,
  xinit=[zeros(i-1,1);1;zeros(n-i,1)];
  h=dlsim(A,zeros(n,1),C,zeros(q,1),zeros(t,1),xinit);
  if nargin==4, h=rev(h); end, M(:,i)=h(:);
end
```

# Samenvatting (Summary in Dutch)

Dit boek handelt over het modelleren van de dynamica van verschijnselen op grond van meetgegevens. We veronderstellen dat de waarnemingen het gedrag weergeven op een reeks tijdstippen met regelmatige tussenpozen. Dit is weer te geven door middel van een tijdreeks

$$w = \begin{bmatrix} w_1(1) & w_1(2) & \dots & w_1(N) \\ \vdots & \vdots & \vdots & \vdots \\ w_q(1) & w_q(2) & \dots & w_q(N) \end{bmatrix}. \tag{1}$$

Elk van de $q$ rijen geeft het gedrag van één kwantitatief aspect van het te modelleren verschijsel weer, waarbij de tijdstippen voor het gemak genummerd zijn van 1 tot $N$. Een typisch voorbeeld is de tijdreeks van rentestanden, uitgebeeld in Figuur 6.1. Het betreft twee variabelen ($q = 2$), de lange en korte termijn rente in de Verenigde Staten, in de maanden Januari 1957 tot en met April 1984 (N=388).
Het thema van dit boek is het beschrijven van het modelleringsproces in termen van (uitwendig) gedrag, in scherp onderscheid met eventuele vermoedens over de inwendige struktuur van verschijnselen. In de inleiding werken we dit principe uit op een algemeen niveau, waarna we het toespitsen op het analyseren van tijdreeksen. Dit mondt uit in de volgende beschrijving van het doel van het modelleren, de definitie van modellen en hun kwaliteit ten op zichte van waarnemingen:
We stellen ons ten doel op grond van *waargenomen* gedrag, weergegeven door de tijdreeks, het realistisch gehalte in te schatten van *denkbaar* gedrag, zoals b.v. het toekomstig verloop. Omdat denkbaar gedrag ook tijdstippen betreft buiten het waarnemingsinterval, geven we dit weer met (tweezijdig) oneindige tijdreeksen. Modellen kunnen nu gedefinieerd worden als een verzameling $\mathcal{B}$ van zulke tijdreeksen, met de interpretatie dat realistisch gedrag (grotendeels) bevat is in deze verzameling, en de rest dus als onrealistisch (of als minder waarschijnlijk of dominant) beschouwd kan worden.
Bij de in dit boek beschreven methode, Globale Totale Kleinste Kwadraten (GTLS), leggen we de volgende beperkingen op aan het type modellen dat we in ogenschouw nemen.

1. Tijdsinvariantie: met een tijdreeks moeten ook al haar verschuivingen

bevat zijn in een model. Modelwetten mogen dus niet expliciet van de tijd afhangen.

2. Lineariteit: van twee tijdreeksen in een model moeten ook alle lineaire combinaties bevat zijn in dat model.

Dit betekent dat we als modelklasse alle lineaire, verschuivingsinvariante verzamelingen van tijdreeksen nemen. Deze modellen kunnen beschreven worden door middel van lineaire differentie vergelijkingen met constante coëfficiënten. De kwaliteit van een model hangt af van twee factoren: grootte en nauwkeurigheid. De grootte van een model wordt gedefinieerd in termen van dimensies: de *rang* $m$ (het aantal vrijheidsgraden op elk tijdstip) en de *graad* $n$ (het aantal begincondities), vgl. Definitie 2.1.6. Modellen met grootte $(m, n)$ komen overeen met lineaire tijdsinvariante input/output systemen met $m$ inputs, $q - m$ outputs en *McMillan* graad $n$. Ze kunnen beschreven worden door $q - m$ vergelijkingen waarvan de som van de ordes gelijk is aan $n$.
Als maat voor de nauwkeurigheid wordt de *misfit* genomen. Dit is de (Euclidische) afstand van een model $\mathcal{B}$ tot een waargenomen tijdreeks $w$,

$$\mathrm{d}(w, \mathcal{B}) := \min_{\hat{w} \in \mathcal{B}} \|w - \hat{w}\|. \tag{2}$$

Dit is equivalent met het minimaliseren van de som van de kwadraten van de residuen $\tilde{w} := w - \hat{w}$, hetgeen de laatste twee letters in GTLS verklaart (Least Squares).
In een definitie van de kwaliteit moeten deze aspekten tegen elkaar afgewogen worden. In de GTLS methode kiezen we voor het optimaliseren van de nauwkeurigheid onder een beperking op de grootte van een model. Dit leidt tot de volgende definitie van het GTLS probleem:

> Gegeven: een tijdreeks als in (1), en twee getallen $m$ en $n$.
> Bepaal: een model met ten hoogste rang $m$ en ten hoogste graad $n$ dat minimale misfit (2) heeft.

Bijvoorbeeld, het GTLS model van tweede orde (en rang één) voor de reeds genoemde rentestanden komt overeen met een tweede orde differentie vergelijking, weergegeven in (6.2). Deze vergelijking is optimaal in de zin dat de vereiste aanpassing van de data om de vergelijking precies kloppend te maken minimaal is.
In hoofdstuk 2 gaan we uitgebreid in op het motiveren van de algemene probleemstelling. Het verschil met andere 'kleinste kwadraten' methoden is gelegen in de combinatie van twee eigenschappen van het misfit criterium (2). Ten eerste laat dit aanpassing toe in alle componenten van de data. Voor het statische geval (graad 0) betekent dit dat GTLS samenvalt met de bekende 'totale kleinste kwadraten' methode, hetgeen het 'totale' in GTLS verklaart. Ten tweede wordt de data $w$ benaderd door een tijdreeks $\hat{w}$ die als geheel aan de modelwetten voldoet, dus ook aan alle *globale* implicaties van deze wetten over grote tijdsintervallen.

Voor het daadwerkelijk bepalen van modellen splitsen we het GTLS probleem in twee stappen:

(a) Het bepalen van de optimale benadering van de data $w$ binnen een gegeven model, d.w.z. het bepalen van een oplossing $\hat{w}$ in (2).

(b) Het bepalen van modellen met minimale misfit.

Ter voorbereiding van het oplossen van deze problemen introduceren we in Hoofdstuk 3 *isometrische toestandsrepresentaties* (ISR's). Deze spelen een cruciale rol in de oplossingsmethode, vanwege hun bijzondere eigenschappen. Ze brengen een belangrijk verband aan het licht tussen een systeem en zijn *orthogonaal complement*, weergegeven in Propositie 4.2.1. Het orthogonaal complement van een model komt overeen met dynamische vergelijkingen voor de benaderingsfout $\tilde{w}$. Dit feit legt de basis voor Algoritme 1 in Hoofdstuk 4, dat een oplossing berekent van het deelprobleem (a). Als een alternatief geven we ook een recursieve versie, hetgeen vergelijkbaar is met Kalman filteren, maar geheel gebaseerd op het (niet stochastische) GTLS criterium (2).
In hoofdstuk 5 wordt het tweede gedeelte (b) behandeld. Eerst worden enkele simpele voorbeelden genoemd waaruit blijkt dat er soms geen en soms meer dan één oplossing kan bestaan voor dit niet-lineaire optimaliseringsprobleem. Daarna beschrijven we twee algoritmen voor het bepalen van stationaire punten (lokaal optimale oplossingen). Algoritme 3 berekent iteratief de oplossingen van eenvoudige deeloptimalisaties, deels door orthogonale projectie en deels met behulp van de 'singuliere waarden decompositie' (SVD). Op grond van deze constructies wordt een karakterisering van (lokale) optimaliteit afgeleid. Een model is dan en slechts dan een stationair punt wanneer er geen correlatie is tussen enerzijds de benadering $\hat{w}$ met bijbehorende toestandsvariabele $\hat{x}$, en anderzijds de benaderingsfout $\tilde{w}$ en toestand $\tilde{x}$. Deze karakterisering kan weer gebruikt worden voor het bepalen van de *optimaliteitsmarge* van een model, die de vereiste aanpassing van de data weergeeft om een gegeven model tot een stationair punt te maken. Het tweede, aanzienlijk snellere GTLS algoritme is gebaseerd op de methode van Gauss-Newton. Hierin spelen de ISR's wederom een essentiële rol.
Algoritme 5 bepaalt modellen op grond van canonieke correlatie analyse van verleden en toekomst van de waargenomen tijdreeks. Deze modellen kunnen gebruikt worden als startwaarde voor de iteratieve GTLS algoritmen. Van de genoemde algoritmen zijn implementaties in Matlab beschreven, in Appendix B.
In hoofdstuk 6 illustreren we de GTLS methode met verscheidene toepassingen op zowel econometrisch als systeemtheoretisch gebied. Aan de hand van de reeds genoemde rentestanden behandelen we onderwerpen als modelvalidatie, sterk afwijkende waarnemingen (outliers), het opnemen van trends en constanten en (co-)integratie. Als een uitbreiding op de GTLS methode bepalen we *periodieke* modellen voor de kwartaalcijfers van consumptie en inkomen in West Duitsland. Ook laten we zien hoe 'simultane vergelijkingen' geschat kunnen

worden. De systeemtheoretische toepassingen betreffen (frequentie gewogen) optimale modelreductie en het modelleren van in- en uitgangs gegevens.

# Bibliography


[1] H. Akaike. Markovian representation of stochastic processes by canonical variables. *SIAM Journal on Control and Optimization*, 13:162–173, 1975.

[2] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.

[3] K. Åström and P. Eykhoff. System identification - a survey. *Automatica*, 16:123–167, 1971.

[4] Y. Baram and B. Porat. Identification of minimal order state space models from stochastic input-output data. *SIAM Journal on Control and Optimization*, 26:56–65, 1988.

[5] S. Beghelli, R. P. Guidorzi, and U. Soverni. The Frisch scheme in dynamic system identification. *Automatica*, 26 (1):171–176, 1990.

[6] S. Bittanti, P. Colaneri, and G. de Nicalao. The difference periodic Riccati equation for the periodic prediction problem. *IEEE Transactions on Automatic Control*, AC-33(8):706–712, 1988.

[7] P. E. Caines. *Linear Stochastic Systems*. Wiley, New York, 1988.

[8] R. Davidson and J. G. MacKinnon. *Estimation and Inference in Econometrics*. Oxford University Press, 1993.

[9] J. G. de Gooijer. Dynamic factor analysis of nonstationary multivariate time series. *Psychometrika*, 57:333–349, 1992.

[10] B. de Moor. *Mathematical Concepts and Techniques for Modelling of Static and Dynamic Systems*. PhD thesis, Catholic University Leuven, 1988.

[11] B. de Moor. Numerical algorithms for state space subspace system identification. *Academia Analecta*, Klasse der Wetenschappen, Koninklijke Akademie voor Wetenschappen 55 (5), België, 1994. To be published.

[12] B. de Moor and B. Roorda. $l_2$-Optimal linear system identification: Structured total least squares for siso systems, 1994. Submitted to the 33th Conference on Decision and Control.

[13] M. Deistler. Symmetric modelling in system identification. In H. Nijmeijer and J. M. Schumacher, editors, *Three Decades of Mathematical System Theory*, pages 128–147. Springer, 1989.

[14] P. H. Franses. *Model Selection and Seasonality in Time Series*. PhD thesis, Tinbergen Institute Series No. 18, 1991.

[15] K. Glover. All optimal Hankel norm approximations of linear multivariable systems and their $L^\infty$ error bounds. *International Journal of Control*, 39 (6):1115–1193, 1984.

[16] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 1983.

[17] E. J. Hannan and M. Deistler. *The Statistical Theory of Linear Systems*. Wiley, New York, 1988.

[18] B. Hanzon. *Identifiability, Recursive Identification and Spaces of Linear Dynamical Systems*, volume 63,64 of *CWI Tracts*. CWI, Amsterdam, 1994.

[19] B. Hanzon and J. M. Maciejowski. Constructive algebra methods for the $l_2$-problem for stable linear systems. In *Proceedings of the IFAC Identification SYSID'94 Conference*, pages 1.191–196, Copenhagen, July 1994.

[20] C. Heij. *Deterministic Identification of Dynamical Systems*, volume 127 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, 1989.

[21] C. Heij and B. Roorda. A behavioural approach to $l_2$-approximation. In *Systems and Networks: Mathematical Theory and Applications, Volume 2*, pages 207–210. Akademie Verlag, 1993.

[22] C. Heij and B. Roorda. Noise minimization in factor models. In *Proceedings of the IFAC Identification SYSID'94 Conference*, pages 3:203–206, Copenhagen, July 1994.

[23] W. Immink. *Parameter Estimation in Markov Models and Dynamic Factor Analysis*. PhD thesis, Rijksuniversiteit Utrecht, Utrecht, 1991.

[24] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[25] H. Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer Verlag, Berlin, Germany, 1991.

[26] B. C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, AC-26 (1):17–32, 1981.

[27] M. Ooms. *Empirical Vector Autoregressive Modeling*, volume 407 of *Lecture Notes in Economics and Mathematical Systems*. Springer, Berlin, 1994.

[28] D. R. Osborn. The implications of periodically varying coefficients for seasonal time-series processes. *Journal of Econometrics*, 3:373–84, 1991.

[29] H. J. Payne and L. M. Silverman. On the discrete time algebraic riccati equation. *IEEE Transactions on Automatic Control*, AC-18 (3):226–234, 1973.

[30] R. Peeters. *System Identification Based on Riemannian Geometry - Theory and Algorithms*. PhD thesis, Tinbergen Institute Series No. 18, 1994.

[31] C. R. Rao. *Linear Statistical Inference and Its Applications*. Wiley, New York, 1973.

[32] B. Roorda. Algorithms for global total least squares modelling of finite multivariable time series. *Discussion Paper* 130, Tinbergen Institute, Erasmus University Rotterdam, 1994. To appear in Automatica, April 1995.

[33] B. Roorda and C. Heij. A modified canonical correlation approach to approximate state space modelling. In *Proceedings of the 30th Conference on Decision and Control*, pages 1343–1348, Brighton, England, Dec. 1991.

[34] B. Roorda and C. Heij. Global total least squares modelling of multivariable time series. *Discussion Paper* 93-177, Tinbergen Institute, Erasmus University Rotterdam, The Netherlands, 1993. To appear in IEEE Transactions on Automatic Control, Jan. 1995.

[35] B. Roorda and C. Heij. The $l_2$-optimal approximate modelling problem. In *Proceedings of the 32nd Conference on Decision and Control*, pages 3649–3651, San Antonio, Texas, 1993.

[36] B. Roorda and C. Heij. On $l_2$-optimal approximate modelling of vector time series. In *Proceedings of the 2nd European Control Conference, Groningen, The Netherlands*, pages 1219–1224, Dec 1993.

[37] R. J. Shiller. The volatility of long-term interest rates and expectation models of the term structure. *Journal of Political Economy*, 87:1190–1219, 1979.

[38] J. T. Spanos, M. H. Milman, and D. L. Mingori. A new algorithm for $L_2$ optimal model reduction. *Automatica*, 28 (5):897–909, 1992.

[39] A. J. van der Veen. SVD-based estimation of low-rank system parameters. In F. Deprettere and A. J. van der Veen, editors, *Algorithms and Parallel VLSI Architectures*, pages A.203–228. Elsevier, 1991.

[40] P. van Overschee and B. de Moor. Subspace algorithms for the stochastic identification problem. *Automatica*, 29 (3):649–660, 1993.

[41] P. van Overschee and B. de Moor. N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30 (1):75–93, 1994.

[42] J. H. van Schuppen. Stochastic realization problems. In H. Nijmeijer and J. M. Schumacher, editors, *Three Decades of Mathematical System Theory*, pages 480–523. Springer, 1989.

[43] M. Vidyasagar. *Control System Synthesis: A Factorization Approach.* MIT Press, Cambridge, 1985.

[44] J. C. Willems. From time series to linear system, part I: Finite dimensional linear time invariant systems; part II: Exact modelling; part III: Approximate modelling. *Automatica*, 22/23:561–580,675–694,87–115, 1986,1987.

[45] J. C. Willems. Models for dynamics. *Dynamics Reported*, 2, 1989.

[46] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Transactions on Automatic Control*, AC-36 (3):259–294, 1991.

[47] J. C. Willems and C. Heij. Scattering theory and approximation of linear systems. In C. Byrnes and A. Lindquist, editors, *Modelling, Identification and Robust Control*, pages 397–411, Amsterdam, 1986. North-Holland.

# Basic Notions

**Time Series** (Section 2.1.1)
Sequence of real valued vectors $w : T \to \mathsf{R}^q$, with $T$ a finite or infinite interval in $\mathsf{Z}$, and $q$ the number of components. Most often they are multivariable, i.e., with $q > 1$. Time series on $[1, N]$ are identified with matrices with $q$ rows and $N$ columns.

**Data, Observation, Measurements** (Section 2.1.1)
An observed time series over a finite time-interval.

**Shift-invariance** (Definition 2.1.2)
Time-invariant, not changing under shifting time.

**Model, System** (Definition 2.1.4)
A set of time series on $\mathsf{Z}$ that is linear and shift-invariant. Corresponds to (simultaneous) linear difference equations with constant coefficients. The term system is used if the relation with data is not relevant.

**Misfit** (Definition 2.1.5)
Euclidian distance between data and system. The squared misfit is comparable to the 'sum of squared residuals'.

**Rank, Degree, Size** (Definition 2.1.6)
The rank denotes the number of degrees of freedom in a system at each time instant. Corresponds to the difference between the number of components in time series and the number of imposed independent equations. The degree of a system denotes the number of degrees of freedom due to unspecified initial conditions, or briefly 'the number of initial conditions'. Corresponds to the sum of the orders of difference equations, and to state dimension in a state representation. The size of a system is simply the pair of rank and degree. The size of a system determines the minimal number of model parameters in a representation.

**Auxiliary Input** (Section 3.1, 2.4.3)
A time series that represents the effect of the unobserved, unmodelled environment, with number of components equal to the rank of a system. On the one hand it is comparable to exogenous variables or inputs, as its value is not prescribed by the model. On the other hand it is not part of the data (as inputs and exogenous variables), but has to be reconstructed in the identification procedure. So it is also comparable to disturbances or noise, but without stochastic specification.

**State Representation (SR)** (Section 3.1)
Description of models by first order equations in terms of state variables, which represent the memory of a system. We reserve the term SR for representations that use both state variables and auxiliary inputs. SR's are easily translated to the better-known input/state/output representations, as is described in Section 3.4.

**Isometric State Representation (ISR)** (Section 3.2)
State representations in which both states and auxiliary inputs have a normalized effect on the system behaviour. They induce an isometry from initial states and auxiliary inputs, to system trajectories and corresponding final states.

**Stable, Asymptotically Stable** (used in Proposition 3.2.7)
A matrix is called stable if all its eigenvalues have absolute value at most one, and it is called asymptotically stable if the eigenvalues have absolute value less than one. An input/output mapping is called stable if bounded inputs correspond to bounded outputs. State representations (and input/state/output representations) $(A, B, C, D)$ are called (asymptotically) stable if $A$ is.

**Stabilizable, Controllable** (Definition 3.2.3, in Theorem 5.3.1)
A system is called *stabilizable* if all trajectories on finite time admit a propagation within the system that converges to zero, and it is called *controllable* if they all admit a propagation that becomes zero in finite time.

# Notation

**Mathematical symbols**

| | |
|---|---|
| $\perp$ | is orthogonal to |
| $\|.\|$ | the Euclidian norm of vectors |
| $\|\|.\|\|$ | the Euclidian (or Frobenius) norm of matrices |
| $\langle ., . \rangle$ | inner product |
| | |
| $\{z \in V;\ C\}$ | the set of elements in $V$ that satisfy condition $C$ |
| $\mathsf{Z}$ | the set of integers $\{\ldots, -1, 0, 1, \ldots\}$ |
| $\mathsf{R}$ | the set of real numbers |
| $R^q$ | the set of $q$-dimensional real vectors |
| $\mathsf{R}^{q \times N}$ | the set of real matrices with $q$ rows and $N$ columns |
| $(\mathsf{R}^q)^{\mathsf{Z}}$ | the set of time series on $\mathsf{Z}$ with $q$ components, $\{w : Z \to \mathsf{R}^q\}$ |
| | |
| $\sigma$ | the shift operator defined by $(\sigma f)(t) = f(t+1)$ |
| $\sigma x$ | sometimes used to denote $x(2), \ldots, x(N+1)$ |
| $\dim V$ | the dimension of a linear space $V$ |
| $\mathrm{im} M$ | image of a mapping $M$ |
| $M^\top$ | the transpose of matrix $M$ |
| $G^*$ | the adjoint of an operator $G$ |
| $[t_0, t_1]$ | the discrete time interval $\{t_0, \ldots, t_1\}$ in $\mathsf{Z}$ |
| $\mathcal{B}_T$ | the restriction of a system to time interval $T$ |
| $\mathcal{B}^\perp$ | the orthogonal complement of system $\mathcal{B}$ |
| $\mathrm{d}(w, \mathcal{B})$ | the misfit of system $\mathcal{B}$ with respect to time series $w$ |
| $\mathrm{col}(a_1, \ldots, a_k)$ | the column consisting of $a_1, \ldots, a_k$ |
| cov | covariance |
| arccos | the inverse of the cosine on $[0, \pi]$ |
| $\frac{df}{dx}\|_{x_0}$ | the Jacobian of $f$ in the point $x_0$ |

**Abbreviations**

| | |
|---|---|
| ARX | Auto Regressive with eXogenous variables |
| GFR | German Federal Republic |
| GTLS | Global Total Least Squares |
| ISR | Isometric State Representation |
| LTLS | Local Total Least Squares |
| OE | Output Error |
| PISR | Periodic Isometric State Representation |
| PSR | Periodic State Representation |
| SR | State Representation |
| SVD | Singular Value Decomposition |
| US | United States |

**Special symbols**

| | |
|---|---|
| $q$ | number of components of time series |
| $m$ | rank of a system, or number of auxiliary inputs in an SR |
| $n$ | degree of a system, or number of state components in an SR |
| $p$ | number of independent restrictions, equal to $q - m$ |
| $N$ | number of data points |
| $T$ | observation interval $[1, N]$ |
| $w$ | the data, or an arbitrary (multivariable) time series |
| $\hat{w}$ | an approximation of the data, or a system trajectory |
| $\tilde{w}$ | the approximation error, $\tilde{w} = w - \hat{w}$ |
| $\hat{w}^*$,$\tilde{w}^*$ | optimal approximation of $w$ and corresponding error |
| $\hat{x}$ | a state trajectory of an approximation $\hat{w}$ |
| $\hat{v}$ | an auxiliary input of an approximation $\hat{w}$ |
| $\tilde{x}$,$\tilde{v}$ | as $\hat{x}, \hat{v}$, but now for the approximation error |
| $\mathsf{B}$, $(\bar{\mathsf{B}})$ | the class of linear, shift-invariant (complete) systems |
| $\mathsf{B}^q$ | the class of systems in $\mathsf{B}$ with $q$ components |
| $\mathsf{B}^{q,m,n}$ | the class of systems in $\mathsf{B}^q$ with rank $m$ and degree $n$ |
| $\mathcal{B}$ | a system (a linear, shift-invariant set of time series on $\mathsf{Z}$) |
| $\mathcal{B}^q$ | the set of time series that belong to a system in $\mathsf{B}^q$ |
| $\mathcal{B}^{q,m,n}$ | the set of time series that belong to a system in $\mathsf{B}^{q,m,n}$ |
| $A, B, C, D$ | system matrices in a state representation of a system |
| $\mathcal{B}(A, B, C, D)$ | the system corresponding to SR $(A, B, C, D)$ |
| $A, \tilde{B}, C, \tilde{D}$ | system matrices in an SR of the orthogonal complement of a system |

# Index

The Tinbergen Institute is the Netherlands Research Institute and Graduate School for General and Business Economics founded by the Faculties of Economics (and Econometrics) of the Erasmus University in Rotterdam, the University of Amsterdam and the Free University in Amsterdam. The Tinbergen Institute, named after the Nobel prize laureate professor Jan Tinbergen, is responsible for the PhD-program of the three faculties mentioned. Since January 1991 also the Economic Institute of the University of Leiden participates in the Tinbergen Institute.

Copies of the books which are published in the Tinbergen Institute Research Series can be ordered through Thesis Publishers, P.O. Box 14791, 1001 LG Amsterdam, The Netherlands, phone: +3120 6255429; fax: +3120 6203395.

The following books already appeared in this series:

Subseries A. General Economics

no. 1 Otto H. Swank, "Policy Makers, Voters and Optimal Control, Estimation of the Preferences behind Monetary and Fiscal Policy in the United States".

no. 2 Jan van der Borg, "Tourism and Urban Development. The impact of tourism on urban development: towards a theory of urban tourism, and its application to the case of Venice, Italy".

no. 3 Albert Jolink, "Liberté, Egalité, Rareté. The Evolutionary Economics of Léon Walras".

no. 5 Rudi M. Verburg, "The Two Faces of Interest. The problem of order and the origins of political economy and sociology as distinctive fields of inquiry in the Scottish Enlightenment".

no. 6 Harry P. van Dalen, "Economic Policy in a Demographically Divided World".

no. 8 Marjan Hofkes, "Modelling and Computation of General Equilibrium".

no. 12 Kwame Nimako, "Economic Change and Political Conflict in Ghana 1600-1990".

no. 13 Ans Vollering, "Care Services for the Elderly in the Netherlands. The PACKAGE model".

no. 15 Cees Gorter, "The dynamics of unemployment and vacancies on regional labour markets".

no. 16 Paul Kofman, "Managing primary commodity trade (on the use of futures markets)".

no. 18 Philip Hans Franses, "Model selection and seasonality in time series".

no. 19 Peter van Wijck, "Inkomensverdelingsbeleid in Nederland. Over individuele voorkeuren en distributieve effecten".

no. 20 Angela van Heerwaarden, "Ordering of risks. Theory and actuarial applications".

no. 21 Jeroen C.J.M. van den Bergh, "Dynamic Models for Sustainable Development".

no. 22 Huang Xin, "Statistics of Bivariate Extreme Values".

no. 23 Cees van Beers, "Exports of Developing Countries. Differences between South-South and South-North trade and their implications for economic development".

no. 24 Lourens Broersma, "The Relation Between Unemployment and Interest Rate. Empirical evidence and theoretical justification".

no. 26 Michel de Lange, "Essays on the Theory of Financial Intermediation".

no. 27 Siem-Jan Koopman, "Diagnostic checking and intra-daily effects in time series models".

no. 28 Richard Boucherie, "Product-form in queueing networks".

no. 29 Frank A.G. Windmeijer, "Goodness of Fit in Linear and Qualitative-Choice Models".

no. 30 M. Lindeboom, "Empirical duration models for the labour market".

no. 31 S.T.H. Storm, "Macro-economic Considerations in the Choice of an Agricultural Policy".

no. 33 R.W. van Zijp, "Austrian and New Classical Business Cycle Theories".

no. 34 J.A. Vijlbrief, "Unemployment Insurance and the Dutch Labour Market".

no. 35 G.E. Hebbink, "Human capital, labour demand and wages. Aspects of labour market heterogeneity".

no. 36 J.J.M. Potters, "Lobbying and pressure: theory and experiments".

no. 37 P. Boswijk, "Cointegration, identification and exogeneity. Inference in structural error correction models".

no. 38 M. Boumans, "A case of limited physics transfer. Jan Tinbergen's resources for re-shaping economics".

no. 42 Jan Kees Martijn, "Exchange-rate variability and trade: Essays on the impact of exchange-rate variability on international trade flows".

no. 43 Jos Verbeek, "Studies on Economic Growth Theory. The role of imperfections and externalities".

no. 45 F. Kalshoven, "Over Marxistische Economie in Nederland, 1883- 1939".

no. 46 W. Swaan, "Behaviour and Institutions under Economic Reform. Price Regulation and Market Behaviour in Hungary".

no. 47 J.J. Capel, "Exchange Rates and Strategic Decisions of Firms".

no. 48 M.F.M. Canoy, "Bertrand meets the Fox and the Owl - Essays in the Theory of Price Competition".

no. 50 Elles A.M. Bulder, "The social economics of old age: strategies to maintain income in later life in the Netherlands 1880-1940".

no. 51 Jaap Barendregt, "The Dutch Money Purge. The monetary consequences of German occupation and their redress after liberation, 1940-1952.

no. 55 Emiel F.M. Wubben, "Markets, Uncertainty and Decision-Making. A History of the Introduction of Uncertainty into Economics".

no. 57 R.M. de Jong, "Asymptotic Theory of Expanding Parameter Space Methods and Data Dependence in Econometrics".

no. 59 A. Lagendijk, "The internationalisation of the Spanish automobile industry and its regional impact. The emergence of a growth- periphery".

no. 60 B.M. Kling, "Life insurance, a non-life approach".

no. 64 Ralf Peeters, "System Identification Based on Riemannian Geometry: Theory and Algorithms".

no. 65 Olga Memedovic, "On the theory and measurement of comparative advantage. An empirical analysis of Yugoslav trade in manufactures with the OECD countries, 1970-1986".

no. 68 René A. Belderbos, "Strategic Trade Policy and Multinational Enterprises: Essays on Trade and Investment by Japanese Electronics Firms".

no. 69 Inge T. van den Doel, "Dynamics in Cross-Section and Panel Data Models".

no. 70 Michael Dell, "Maximum Price Regulations and Resulting Parallel and Black Markets".

no. 74 John Baggen, "Duurzame mobiliteit. Duurzame ontwikkeling en de voorzieningenstructuur van het personenvervoer in de Randstad".

no. 75 Ron Boschma, "Looking through a window of locational opportunity. A long-term spatial analysis of techno-industrial upheavals in Great-Britain and Belgium.

no. 78 Rob van de Wijngaert, "Trade Unions and Collective Bargaining in The Netherlands".

no. 81 Rob Zuidwijk, "Complementary Triangular Forms for Pairs of Operators and Matrices".

no. 83 E. Eggink, "A Symmetric Approach to the Labor Market: An application of the SEM method".

No. 84 A.F. Correljé, "The Spanish Oil Industry: Structural change and modernization".

No. 85 André Leliveld, "Social security in developing countries; operation and dynamics of social security mechanisms in rural Swaziland".

No. 87 Erno Kuiper, "Farmers, Prices and Rational Expectations".

Subseries B. Business Economics

no. 4 Rob Buitendijk, "Towards an Effective Use of Relational Database Management Systems".

no. 7 P.J. Verbeek, "Two Case Studies on Manpower Planning in an Airline".

no. 9 T.C.R. van Someren, "Innovatie, emulatie en tijd. De rol van de organisatorische vernieuwingen in het economische proces".

no. 10 M. van Vliet, "Optimization of manufacturing system design".

no. 11 R.M.C. van Waes, "Architectures for Information Management. A pragmatic approach on architectural concepts and their application in dynamic environments".

no. 14 Shuzhong Zhang, "Stochastic Queue Location Problems".

no. 17 Paul Th. van de Laar, "Financieringsgedrag in de Rotterdamse maritieme sector, 1945-1960".

no. 25 E. Smeitink, "Stochastic Models for Repairable Systems".

no. 32 H.E. Romeijn, "Global Optimization by Random Walk Sampling Methods".

no. 39 J.B.J.M. de Kort, "Edge-disjointness in combinatorial optimization: problems and algorithms".

no. 40 J.F.J. de Munnik, "The Valuation of Interest Rates Derivative Securities".

no. 41 Jeroen C.A. Potjes, "Empirical Studies in Japanese Retailing".

no. 44 R.H. van het Kaar, "Medezeggenschap bij fusie en ontvlechting".

no. 49 H.M. Kat, "The Efficiency of Dynamic Trading Strategies in Imperfect Markets".

no. 52 Nanda Piersma, "Combinatorial Optimization and Empirical Processes".

no. 53 M.J.C. Sijbrands, "Strategische en Logistieke Besluitvorming. Een empirisch onderzoek naar informatiegebruik en instrumentele ondersteuning".

no. 54 Erik van der Sluis, "Heuristics for Complex Inventory Systems. Deterministic and stochastic problems".

no. 56 Volkert J. Batelaan, "Organizational Culture and Strategy. A study of cultural influences on the formulation of strategies, goals, and objectives in two companies".

no. 58 Dirk P.M. de Wit, "Portfolio Management of Common Stock and Real Estate Assets. An Empirical Investigation into the Stochastic Properties of Common Stock and Equity Real-Estate".

no. 61 Jan Grootendorst, "De markthuur op kantorenmarkten in Nederland".

no. 62 Marian Dingena, "The Creation of Meaning in Advertising. Interaction of Figurative Advertising and Individual Differences in Processing Styles".

no. 63 R.T. Lie, "Economische dynamiek en toplocaties. Locatiekarakteristieken en prijsontwikkeling van kantoren in een aantal grote Europese steden".

no. 66 Sven Fischer, "The paradox of information technology management".

no. 67 Philip A. Stork, "Modelling International Financial Markets: An Empirical Study".

no. 71 Bo Chen, "Worst Case Performance of Scheduling Heuristics".

no. 72 Ellen Christiaanse, "Strategic Advantage and the Exploitability of Information Technology. An Empirical Study of the Effects of IT on Supplier-Distributor Relationships in the US Airline Industry".

no. 73 Lei Lei, "User Participation and the Success of Information System Development. An Integrated Model of User-Specialist Relationships".

no. 76 Anco Sneep, "Innovation management in the Agro-food industry".

no. 79 M. Boogaard, "Defusing the Software Crisis: Information Systems Flexibility Through Data Independence".

no. 80 Erik M. Vermeulen, "Corporate Risk Management. A multi-factor approach".

no. 82 Marc H. Goedhart, "Financial Planning in Divisionalised Firms: Models and Methods".

no. 86 Yvonne M. Prince, "Price Cost Margins in Dutch Manufacturing: With an Emphasis on Cyclical and Firm Size Effects".