System Identification for Robust Process Control

Nominal Models and Error Bounds

Richard G. Hakvoort

System Identification for Robust Process Control
Nominal Models and Error Bounds

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College van Dekanen aangewezen,
op dinsdag 29 november 1994 te 13.30 uur
door

Richard Gerrit HAKVOORT

werktuigbouwkundig ingenieur

geboren te Delft

Dit proefschrift is goedgekeurd door de promotor:
Prof. ir. O.H. Bosgra

Toegevoegd promotor:
Dr. ir. P.M.J. Van den Hof

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Hakvoort, Richard G.

System Identification for Robust Process Control:
Nominal Models and Error Bounds / Richard G. Hakvoort. -
Delft: Delft University of Technology,
Faculty of Mechanical Engineering and Marine Technology. - Ill.
Thesis Technische Universiteit Delft.
ISBN 90-370-0114-9
Subject headings: system identification / robust control / process control.

I devoted myself to study and to explore by wisdom all that is done under heaven.
Ecclesiastes 1:13a

So I turned my mind to understand, to investigate and to search out wisdom and the scheme of things.
Ecclesiastes 7:25a

To Eveline, my mother, Nelly

Preface

This thesis is the result of four years of Ph.D. research in the Mechanical Engineering Systems and Control Group at Delft University of Technology. Also due to the efforts of many other people, these four years have been a great time for me.

First of all, I would like to thank prof. Okko Bosgra for the confidence he put in me by letting me go my own way with my Ph.D. research.

Next, I thank Paul Van den Hof and Ton van der Weiden for their input into this work. I also give thanks to Ruud Schrama, Peter Bongers and Douwe de Vries for many fruitful discussions on the various subjects touched in this thesis.

With my roommates Hans Heintze and Raymond de Callafon I have had a very pleasant time, for which I wish to thank them. I also acknowledge the valuable support of Peter Valk and Cor Cremers.

I will always remember the 8 years of chess with Ewoud van Luik. The numerous games that we played, and our more or less serious conversations about other subjects, were an excellent way to spend our lunch breaks. I hope Ewoud finds someone else to play chess with, someone whom he can beat.

Special thanks are given to Dick de Roover. We had many pleasant discussions about our work and many other subjects. His input into this work is significant. In fact the identification results in Section 9.3 have been produced by Dick, for which I am grateful.

Next, I would like to thank my project-mates Peter van Overschee, Ghassan Murad, Heinz Falkus, Erik van Bracht and Jobert Ludlage. We had many fruitful discussions.

I thank the members of my family for their support during my studies. I am especially grateful to three women who have played an important role in my life, the three women I have dedicated this thesis to: my sister Eveline, my mother and my girlfriend Nelly.

Finally, I thank God for giving me strength each day to do my job.

Delft, October 1994 *Richard Hakvoort*

Summary

The performance of industrial processes can often be improved by implementation of an automatic feedback controller. In the design of a controller fruitfully use can be made of a model of the system to be controlled, provided proper account is given of the fact that a model is not perfect. If a *nominal model* and an upper bound on the inevitable *model error* are specified, a *robust* control design procedure can be used to design a controller with guaranteed performance for the system.

In this thesis the problem is considered of the *identification* of a nominal model and an explicit model error bound on the basis of measurement data. In particular, the objective is to identify a parametric nominal model of some specified order, and an upper bound on the H_∞- or ℓ_1-norm of the weighted additive model error, suited for use in H_∞- respectively ℓ_1-optimal robust control design procedures. Obviously, an overly *conservative* model error bound would unnecessarily restrict the achievable control performance. Therefore the identification should be carried out such that the upper bound on the model error is as small as possible, especially in the *control-relevant* frequency ranges, i.e. around the cross-over frequency. On the one hand, this requires the identification of a control-relevant nominal model, such that the model error is small in the control-relevant frequency ranges. On the other hand, the model error bound should tightly over-bound the model error in these frequency ranges.

To solve this problem of robust-control-oriented identification, a new identification procedure is developed in this thesis. First, on the basis of measurement data and prior information about system and noise a non-parametric *system uncertainty set* is identified, which consists of frequency response or pulse response system uncertainty regions. A procedure to construct such an uncertainty set is called an *uncertainty bounding identification* procedure. Next, on the basis of the system uncertainty set a parametric nominal model of a user-defined order is constructed, minimizing the H_∞- or ℓ_1-norm of a certain worst-case model error criterion. This problem is referred to as the problem of *system approximation* in H_∞- or ℓ_1-norm. Finally, on the basis of the system uncertainty set and the nominal model an upper bound on the H_∞- or ℓ_1-norm of the weighted model error is determined. The identification experiment design and weighting functions in the identification procedure can be tuned in order to emphasize the control-relevant frequency ranges.

Both a procedure for *deterministic* uncertainty bounding identification and a procedure for *probabilistic* uncertainty bounding identification are developed in this thesis. Both procedures utilize deterministic prior information about the system. In particu-

lar, the system is assumed to be linear, time-invariant, possibly infinite dimensional, and a bound is assumed on the pulse response coefficients of the system. In addition, the procedure for deterministic identification utilizes deterministic noise priors, particularly the unknown-but-bounded-amplitude noise assumption, which leads to deterministic system uncertainty regions. Moreover, a newly introduced *cross-covariance noise bound* can be used, which appears to have certain averaging properties which standard time-domain noise bounds lack. The deterministic system uncertainty regions are constructed by solving a set of linear programming problems.

The procedure for probabilistic uncertainty bounding identification utilizes stochastic noise priors, in particular the assumption that the noise is a realization of a stationary stochastic process, and independent of the input signal. This leads to probabilistic system uncertainty regions, or confidence regions, which are correct with a certain probability. The confidence regions are constructed by evaluating explicit bias and variance expressions of an instrumental variable estimate. A comparison between both procedures shows that the probabilistic approach generally leads to much smaller system uncertainty regions than the deterministic approach. In other words the stochastic noise assumption is much less conservative than the deterministic noise assumption.

The problem of system approximation in H_∞ is solved by developing a new *frequency response curve fit* procedure which minimizes a maximum amplitude criterion and guarantees stability of the resulting model. In the solution of the curve fit problem use is made of linear and nonlinear programming techniques. Besides providing a solution for the problem of *identification in* H_∞, the curve fit procedure also provides a solution for the H_∞*-optimal model reduction* problem. The problem of system approximation in ℓ_1 is solved by developing an algorithm for ℓ_1*-optimal model reduction* for linearly parametrized reduced order models. In this algorithm use is made of linear programming.

Practical applicability of the new identification procedure is shown by means of succesful application to a multivariable industrial glass tube manufacturing process.

Contents

Prologue

Chapter 1

Introduction

1.1 Robust Process Control

There are many processes, such as industrial processes or mechanical systems, which have an unsatisfactory natural behaviour. The process or system may exhibit inadequate performance or even unstable behaviour. In such a case a controller can be implemented in order to provide the desired stability and performance, which includes setpoint tracking, disturbance attenuation, improved efficiency and accuracy and decreased energy consumption of the process. Examples of systems to be controlled are industrial robots, chemical processes, power plants, aircrafts and vehicles. Another example is a glass-tube manufacturing process, an industrial process which is in particular addressed in this thesis.

A controller can be designed by many methods. In the early days of control theory control design was carried out by tuning parameters of a PID controller (Ziegler and Nichols, 1942; Bode, 1945). This control design method is still very popular in process industry, as it is a straightforward method to design a controller, and it can be carried out without much knowledge of system and control theory. However, the drawback of this method is that it can not be carried out easily in case the process is complex or multivariable, or if high performance requirements are imposed on the controlled process. In these cases a PID control law certainly does not give the maximum achievable performance.

For that reason in the sixties model-based control design methods were introduced, see for example Luenberger (1966), and also Franklin *et al.*(1986) and Chen (1987) for overviews. These control design methods make use of a state-space model of the system to be controlled, and the controller is calculated by minimizing some specified control criterion. The advantage of this approach is that it greatly facilitates the design of high-performing controllers for complex and multivariable systems, due to its less heuristic nature. A disadvantage of the approach, which was only recognized later, is that the control design method makes use of the so-called *certainty-equivalence* principle. This means that a controller is designed on the basis of a model of the system, as if the model is exactly equal to the system. The control design does not explicitly take into

account the fact that it is practically impossible to obtain a model which is identical to the process it describes.

Some time after the introduction of these control design methods, which are based on the certainty-equivalence principle, this disadvantage was recognized. It appeared that a controller, designed to have a certain performance for a model, need not show the same performance for the system. In the worst case the system might even be destabilized by the controller, whereas the controller works perfectly fine for the model. This discrepancy is due to the inevitable modelling error. In general a system is complex, with high-order, even nonlinear and time-varying, dynamics. A model is generally a more or less crude approximation of the system, with simplified dynamics, often linear and time-invariant, in order to be tractable for the control design procedure. Consequently, there is always a *model error*, which is defined to be the difference between the system and the model, i.e. the additive model error. Due to this model error a controller designed for the model need not perform identically for the system.

Once this insight had also been obtained by control theoreticians, they started to think about an alternative formulation of the control problem, which takes into account the presence of a model error. Based on the work of Youla *et al.* (1976a, 1976b), Zames (1981) and Doyle (1982), so-called *robust* control design and analysis procedures were developed. These control design methods are model-based and, therefore, applicable to complex and multivariable processes. But in addition these methods take into account the fact that the model is not a perfect description of the process to be controlled. In fact these control design methods yield a controller with guaranteed stability and performance not just for the model, but also for a set of models which slightly differ from this model.

Typically the system representation used in the robust control design methods are a model and a norm-bound on the model error. The model is often referred to as the *nominal model* and a deviation from this nominal model within the norm-bound is called a *model perturbation.* A robust control design method constructs a controller which shows satisfactory performance for both the nominal model and the set of perturbed models. A robust controller analysis procedure analyses the performance of a specified controller for all models within the specified norm-bounded set.

Various robust control design and analysis methods have been developed, which differ in the type of performance specifications that can be incorporated and the type of norm used to bound the model error. The so-called H_∞-optimal control design and analysis methods (Vidyasagar, 1985; Francis, 1987; Doyle *et al.*, 1989; Lunze, 1989; Maciejowski, 1989; Morari and Zafiriou, 1989; Kwakernaak, 1993) use an H_∞-norm to bound the model error, and these can typically handle frequency domain performance specifications. The μ-control design and analysis method (Doyle, 1982; Maciejowski, 1989; Packard and Doyle, 1993) is a method which can handle various types of model errors simultaneously, particularly bounds on model parameters and H_∞-bounds on the model error, or a frequency dependent model error bound. The so-called ℓ_1-optimal control design and analysis procedures (Dahleh and Pearson, 1987, 1988; Boyd and Barratt, 1991; Dahleh and Khammash, 1993) can handle an ℓ_1-norm bound on the model error. These procedures are well-suited to incorporate time-domain control specifications.

The introduction of the robust control design procedures offers the possibility to design high-performing controllers for complex and multivariable processes. Perhaps superfluously it is mentioned that using a robust control design procedure provides no carte blanche for the model-builder to deliver a bad nominal model. It is well-known that performance and robustness are conflicting. A large model error can certainly be taken into account in the robust control design procedure, but one cannot expect to achieve a high-performing controller then. On the other hand the best nominal model is not necessarily the most accurate one. Often a simple model is required in order to have a tractable control design problem, and in order to arrive at a low-order controller, which is simple enough to be practically implementable.

1.2 System Identification

1.2.1 Introduction

Obviously, a controller can be designed for a system if a model of this system is available. Robust control can be applied if in addition a model error bound in a suitable format is available. Basically there are two different ways to model a system:

- *White-box modelling* or *physical modelling.* A model is derived by formulating the physical laws which the system obeys, such as mass- and energy-balances or the laws of Newton.
- *Black-box modelling* or *identification.* A model is constructed by fitting a parametric model on measured data, without bothering about a possible physical interpretation of the model parameters.

A third way is the combination of these two:

- *Grey-box modelling.* Unknown parameters of a physical model, such as unknown masses and flows, are estimated from measurement data.

In this thesis attention is restricted to black-box modelling or identification. In many engineering problems identification has succesfully been applied. Sometimes identification is the only method that can be applied to derive a model of the system, for example due to high system complexity, or insufficient knowledge of the physical laws governing the system behaviour.

Also a model derived with identification does not perfectly describe the system. Basically there are two reasons for this.

- The system can show very complex behaviour, which is impossibly captured by a model of some desired simplicity:
  - The (nonlinear, time-varying, high-order) system is generally too complex to be perfectly describable by a model.
  - Often deliberately undermodelling is introduced as the application (control design) requires a linear, low-order model.

- The system dynamics can never be perfectly recovered from measurements due to partialness and corruptedness of the data:
  - The data are partial, i.e. the input signal in the past may be unknown (unknown initial conditions), and the number of measurement samples is finite. This implies that in particular the low-frequent system dynamics can not be accurately identified.
  - The data are corrupted by noise. In general, if the same input signal is applied twice, the measured output signals will not be identical. This is due to the presence of disturbances, such as unmeasurable input signals.

Surveys of applications of identification in process industry can be found in Gustavsson (1975), Freedman (1977), Prett *et al.* (1987) and Andersen *et al.* (1991). The standard approach is to identify a nominal model and judge the validity of this model by testing its ability to reproduce the given data and to predict other data. Next, a controller is designed for the nominal model without explicitly taking into account the model error. By means of a conservative design in general a controller is obtained which is sufficiently robust to perform satisfactorily for the system. However, the control performance certainly need not be the best achievable.

It is not necessary to do an overly conservative control design if there is reliable knowledge of the size of the model error. Robust control design can optimally trade-off performance and robustness if both a nominal model and a bound on the model error are available. Obviously, besides the specification of a nominal model, an important issue in identification should be the specification of an upper bound on the model error. In the next two subsections the current states-of-art in nominal model identification and model error bound identification are briefly described. It is investigated what existing identification methods have to offer, and if they indeed yield a nominal model and a model error bound which are suited to serve as a basis for high-performance robust process control.

1.2.2 Identification of a Nominal Model

As mentioned before an identified model is not a perfect description of the system. The inevitable model error restricts the achievable controller performance. One can try to reduce this effect as much as possible by identifying the nominal model such that the corresponding model error is the least unfavourable for the control design application. This means that the identification should take into account the fact that the model will be used to design a controller. Ways to do this are by means of the design of a suitable identification experiment, or the use of suitable weightings in the identification criterion. An identification procedure that takes into account the control application is called *control-oriented* or *feedback-relevant*. A control-oriented identification method yields a model which is good at the control-relevant frequencies. The control-relevant frequency range typically corresponds to the frequencies in the neighbourhood of the cross-over frequency of the controlled system. The importance of a good model fit for these frequencies has been recognized by e.g. Safonov *et al.* (1988), Jacobson *et al.* (1991), Tischler (1991), Schrama (1992a) and Bayard *et al.* (1992).

The standard method to identify a nominal model is the prediction error method (Ljung, 1987; Söderström and Stoica, 1989). Straightforward application of this method generally yields a model which accurately describes the system dynamics which are dominant in open loop, such as the low-frequent system dynamics. Such a model is in general well-suited for open loop model applications, such as prediction and fault detection, but possibly less suited for control design. In closed loop the system dynamics around the cross-over frequency are much more important than the low-frequent system dynamics.

In Wahlberg and Ljung (1986) it is shown that the design variables in the identification experiment, such as the input signal spectrum, determine the frequency distribution of the model error. In Hakvoort (1990, 1992a), Hakvoort *et al.* (1992,1994) and Zang *et al.* (1991, 1992) this is used to formulate an identification experiment which yields a prediction error model that is optimally suited for control design. In particular the identification is asymptotically optimally tuned towards H_2-optimal control. This is a nice result, but the obvious drawbacks are twofold. First, the results are only asymptotic in the number of measurement data. For a finite number of data no optimality results are available. Second, in general robust control utilizes the H_∞- or ℓ_1-norm. The identification procedure, optimized in view of H_2-optimal control, need not be optimal in view of H_∞- or ℓ_1-optimal control.

More recently the so-called subspace identification methods have been introduced (Moonen *et al.*, 1989; Moonen and Vandewalle, 1990; Van Overschee and De Moor, 1994; Verhaegen, 1994; Viberg, 1994). Similar to the prediction error method these subspace methods also tend to emphasize the low-frequent system dynamics. But contrary to the situation for the prediction error method, it is currently unclear how the design variables in the identification experiment exactly influence the frequency distribution of the model error of the identified subspace model. This makes it very hard to tune the identification method such that a nominal model is identified which is optimally suited for the design of a high-performing controller.

Within the identification community the need of delivering nominal models which are tuned towards the control application has been widely recognized. Much effort has been put into the development of feedback-relevant identification schemes (Balas and Doyle, 1990; Bitmead *et al.*, 1990; Liu and Skelton, 1990; Rivera *et al.*, 1990, 1992; Schrama, 1992a, 1992b; Bayard *et al.*, 1992; Lee *et al.*, 1992, 1993; Zang *et al.*, 1991, 1992; Schrama and Van den Hof, 1993; Hakvoort *et al.*, 1992, 1994). The reader is referred to Bitmead (1993), Gevers (1993), Van den Hof *et al.* (1993a) and Van den Hof and Schrama (1994) for overviews on this topic. Some of the proposed methods are more or less ad hoc, whereas some of them are shown to be optimal in some sense. However, typically this concerns H_2-optimality, the identification minimizing an H_2-error criterion. Moreover, most optimality results are asymptotic in the number of data. The problem remains to find an identification method which yields H_∞- or ℓ_1-optimal results in the non-asymptotic case.

Parallel to the developments in feedback-relevant identification as found in the references given above, a number of people have been working on a different approach to the identification problem. This concerns the so-called problem of identification in H_∞ (Helmicki *et al.*, 1991, 1992, 1993; Partington, 1991; Mäkilä, 1991a, 1991b; Mäkilä

and Partington, 1992; Gu and Khargonekar, 1992a, 1992b; Gu *et al.*, 1992; Akçay *et al.*, 1993; Bai and Raman, 1993) and the problem of identification in ℓ_1 (Jacobson *et al.*, 1992; Partington and Mäkilä, 1994; Theodosopoulos, 1994). The objective is to construct a nominal model in combination with an explicit H_∞- or ℓ_1-model error bound. Moreover, the nominal model is constructed such that it is optimal, in the sense that under certain asymptotic conditions (model order and number of data tend to infinity) the H_∞- or ℓ_1-norm of the model error vanishes. Obviously, this is an important step in the right direction, as the quality of the model is evaluated in the right norm. Nevertheless this approach suffers from serious drawbacks as well. The most serious drawback is that the quality of the nominal model can be very poor if the asymptotic conditions are not satisfied, which is in general the case. In particular no attempt is made to achieve a good model fit for the feedback-relevant frequencies.

1.2.3 Identification of a Model Error Bound

The identification community has very well been aware of the fact that an identified model is not perfect and that there is always a model error. Much effort has been put in characterizing the model error. The prediction error identification method provides a model error characterization by means of parameter variance expressions and implicit bias expressions (Ljung, 1987). Unfortunately this type of model error characterization does not coincide with the needs of robust control. In robust control a bound on the H_∞- or ℓ_1-norm of the model error is needed, or a frequency dependent upper bound on the model error. Classical identification methods do not provide this type of information.

As an answer to this shortcoming of classical identification methods, in the last few years much effort has been put into the development of new identification methods which yield explicit and calculable model error bounds. Frequency response or H_∞-model error bounds can be identified with the methods developed by Goodwin *et al.* (1990), Helmicki *et al.* (1991), Lamaire *et al.* (1991), Van den Boom *et al.* (1991), Zhu (1991), Ninness and Goodwin (1992), Wahlberg and Ljung (1992), Bayard (1992), Larimore (1993) and De Vries and Van den Hof (1992, 1993, 1994a, 1994b). And an ℓ_1-bound on the model error can be identified with the method described in Jacobson *et al.* (1992).

Two approaches in error bounding identification can be distinguished. On the one hand there is the *stochastic, probabilistic* or "*soft*" approach, in which use is made of stochastic noise assumptions. This leads to statistical (soft) model error bounds. Examples are the methods described in Ninnness and Goodwin (1992), Bayard (1992) and De Vries and Van den Hof (1993). On the other hand there is the *deterministic, non-probabilistic* or "*hard*" error bounding approach, in which use is made of deterministic noise assumptions. This leads to deterministic (hard) model error bounds. Examples are the methods of Helmicki *et al.* (1991) and Wahlberg and Ljung (1992).

Each of the methods found in literature has its own features, its advantages, and its disadvantages. Some of these methods are more or less ad hoc, whereas other ones are profound. Some methods, such as the ones of Wahlberg and Ljung (1992) and Helmicki *et al.* (1991), suffer from the serious drawback that they yield conservative

results. This simply means that the model error bound is larger than necessary, or in other words that with the same information (data and assumptions about system and noise) a smaller model error bound should be obtainable. Obviously, the use of a conservative model error bound unnecessarily restricts the achievable performance of the robust controller. This is especially the case if the model error bound is conservative in the feedback-relevant frequency ranges.

In the next chapter the various error bounding identication methods are discussed more thoroughly. At this moment it is mentioned that especially the method of Bayard (1992) and the one of De Vries and Van den Hof (1993, 1994b) have significant practical relevance. These have especially been developed for complex mechanical systems and have succesfully been applied to these systems, yielding tight model error bounds. Also applicability of the first method to industrial processes has been claimed in Rivera *et al.* (1993). However, both methods heavily rely on a proper design of the identification experiment, in particular the use of multi-sinusoidal input excitation. For mechanical systems it is generally not problematic to apply such an input signal. But in industrial applications there is often not much freedom with respect to the choice of input signal, as experimentation time is expensive in an industrial environment. In that case the identification procedure should be able just to get the maximum out of the data that are available.

1.3 Problem Formulation

On the basis of the discussion in the previous sections, requirements can be formulated which an identification method should satisfy in order to be applicable to high-performance robust process control.

Any identification procedure will yield a model which is only an approximation of the system. As indicated in Section 1.1, this is not the end of the story, as modern robust control design procedures can easily cope with model errors. These procedures can be applied if a norm-bound on the model error is available. Consequently, a controller with guaranteed robust stability and performance can be designed under the following condition:

> A nominal model and an upper bound on the model error should be available in a format which can be used in a robust controller design or analysis procedure.

This means that the identification should come up with a nominal model and a model error bound in a proper format. If these are available, a controller can be designed which optimally trades off performance with the required robustness. An appropriate format of the model error bound is a bound on the H_∞- or ℓ_1-norm of the model error, or a frequency dependent upper bound on the model error.

However, it is not sufficient that the identification just yields a nominal model and a model error bound in the proper format. Once a bound on the model error in the proper format is available, it is straightforward to design a controller with guaranteed performance for the system. But this may very well be a low performance. A high

control performance can only be achieved if the model is an accurate description of the system, and the model error bound is correspondingly small.

Unfortunately it is in general impossible to identify a model which accurately describes all the dynamics of the system. Choices have to be made concerning which dynamics should be modelled and which can be discarded. In practice this means that the system can be identified accurately in certain frequency ranges, at the price of a worse fit for other frequencies. Consequently, the design of a high-performing robust controller imposes the following requirement on the nominal model and the model error bound:

> The upper bound on the model error, and hence the model error itself, should be small in the frequency range which is important for the control application, and may be larger in a frequency range which is less important.

Note that it is not sufficient that the nominal model is tuned towards the control application, such that the model error is small at the control-relevant frequencies. As indicated above it is also necessary that the upper bound on the model error is small for these frequencies, as it is this upper bound that is used in robust control design and analysis. An important tool for the tuning of the upper bound on the model error is the choice of uncertainty structure, and in particular the choice of weighting matrices in the uncertainty structure.

An optimal tuning of the model error and the model error bound towards the control application requires the specification of a *feedback-relevant identification criterion*. The identification should minimize a criterion which is related to the criterion used in the control-design procedure. On the one hand, this requires the formulation of the right weightings in the identification. On the other hand, this requires the use of the right norm. In particular, if H_∞-optimal (ℓ_1-optimal) control design procedures are applied, the identification procedure should minimize the H_∞-norm (ℓ_1-norm) of the model error. In addition it is important that optimality of the identification procedure does not heavily rely on the asymptotic case of the number of data samples tending to infinity. The procedure should also work adequately if a finite number of data are available, as is always the case in practice. These issues are more thoroughly discussed in Chapter 2.

Finally, it is important to note that the specific nature of industrial processes puts additional requirements on the identification procedure. In particular these are the following:

- The identification procedure should be able to handle arbitrary input signals, as often these can not be chosen freely due to production or safety requirements.

- The identification procedure should be MIMO (multi-input multi-output) applicable, as many industrial processes are multivariable.

From the discussion in Section 1.2 it becomes clear that the identification procedures presently available in literature do not provide a complete solution for all aspects of the identification problem as outlined here. This motivates the development of new

identification algorithms, both for the construction of a nominal model and for the construction of a model error bound.

This leads to the formulation of the problem considered in this thesis. The main objective of this thesis is to develop an identification procedure which solves the following identification problem:

> GIVEN A PROCESS WITH UNKNOWN DYNAMICS. IDENTIFY A NOMINAL MODEL AND AN EXPLICIT MODEL ERROR BOUND SUITED FOR THE DESIGN OF A HIGH-PERFORMING ROBUST CONTROLLER.

As discussed previously, this implies that

- The nominal model should be identified in a control-oriented way. On the one hand, this requires the use of the right weightings, such that the model error is small in the feedback-relevant frequency ranges. On the other hand, this requires the use of the right norm, i.e. the same norm as used in the robust control design procedure, in particular the H_∞- and ℓ_1-norm. Optimality of the nominal model should be proven for the non-asymptotic case, i.e. for a finite number of data.

- The model error bound should be identified in a control-oriented way as well. On the one hand, this requires the use of the right norm, the H_∞- or ℓ_1-norm. On the other hand, this requires the use of the right weightings in the uncertainty structure, such that the upper bound on the model error is accurate for the feedback-relevant frequencies.

- The identification procedure should be applicable to practical industrial processes. Among other things this implies that the input signal often can not be chosen freely, and is possibly generated in closed loop. Further, a low model complexity is required and the identification procedure should be MIMO applicable. Moreover, it should be taken into account that the number of measurement data is finite, and sometimes even small or very small.

The problem, as formulated above, is tackled in this thesis by providing new tools for identifying nominal models and error bounds, which meet the requirements listed. Algorithms are developed to identify a nominal model, such that either an H_∞-error criterion or an ℓ_1-error criterion is minimized. Also, algorithms are developed to identify an upper bound on the model error of this nominal model. Typically the model error bound is a frequency dependent upper bound on the model error, or a bound on the H_∞- or ℓ_1-norm of the model error.

1.4 Synopsis of the Thesis

This thesis consists of a Prologue, three main parts and an Epilogue.

- The Prologue consists of this chapter and the next chapter. In the next chapter the problem formulation is worked out in much more detail, giving exact mathematical formulations of the problems considered in this thesis. Moreover, an

identification procedure is proposed which represents the basic control-oriented identification approach adopted in this thesis.

- Part I consists of Chapter 3, 4 and 5. There the problem is considered of identifying explicit and non-conservative system uncertainty bounds and model error bounds. Both a procedure for deterministic identification and a procedure for probabilistic identification are developed and analysed.
- Part II consists of Chapter 6 and 7. There the problem is considered of identifying a nominal model. Procedures to identify a nominal model which is optimal in H_∞- or ℓ_1-sense are developed and analysed.
- Part III consists of Chapter 8 and 9 and describes the practical application. First, in Chapter 8 a practical identification procedure is presented, which consists of a combination of tools developed in the Chapters 2 till 7. Next, in Chapter 9 this practical identification procedure is tested on an industrial glass tube manufacturing process.
- The Epilogue consists of conclusions, presented in Chapter 10.

Chapter 2

Identification for Robust Control

2.1 Introduction

In this chapter the issues considered in this thesis are described in detail. The problem formulation in Chapter 1 is worked out and subproblems are formulated which are solved in the sequel of the thesis.

As in general identification is carried out on the basis of a sequence of sampled data, this quite naturally leads to the estimation of discrete-time models. Moreover, in control applications often use is made of digital computers, which requires the design and implementation of discrete-time controllers. For these reasons in this thesis the main focus is on discrete-time systems.

The outline of the chapter is as follows. First, in the next section some robust control theory is recapitulated, which motivates the identification problems considered in this thesis. Then in Section 2.3 the identification problems are investigated in some detail. This leads to the introduction of an identification procedure, which contains the basic ingredients necessary to solve the problem of identification for high-performance robust process control as posed in Section 1.3. In Section 2.4 the various steps of this identification procedure are discussed. Finally, in Section 2.5 an outline is given of the contents of the remaining chapters of this thesis.

2.2 Robust Control

2.2.1 Introduction

As explained in Section 1.1 modern robust control design procedures can cope with a model error, provided an explicit upper bound on the model error is specified. Typically this is done by means of a norm-bound. The two norms used for this are the H_∞- and the ℓ_1-system norm. Besides that, these norms are well-suited to formulate performance specifications for the controlled system. This is due to the close relation of the H_∞-norm and ℓ_1-norm with certain signal norms. In this section these issues are explained in some detail.

2.2.2 Norms

In this subsection some definitions of signal and system norms are given. For more details one is referred to Desoer and Vidyasagar (1975) or Boyd and Barratt (1991, Ch. 4,5).

Let $\{x(t)\}_{t=0,1,\ldots,\infty}$ be a signal, with each $x(t) \in \mathbb{R}^{q'}$. The ℓ_2-norm of $\{x(t)\}$ is defined as

$$\|x\|_2 = \sqrt{\sum_{t=0}^{\infty} |x(t)|^2},$$

where $|x(t)|$ denotes the standard Euclidean vector-norm of $x(t)$,

$$|x(t)| = \sqrt{\sum_{k=0}^{q'} x_k(t)^2}.$$

The ℓ_∞-norm of $\{x(t)\}$ is defined as

$$\|x\|_\infty = \sup_t |x(t)|_\infty$$

where $|x(t)|_\infty$ is defined as

$$|x(t)|_\infty = \max_k |x_k(t)|.$$

Let $G(z)$ be a discrete-time, LTI (linear, time-invariant), causal, possibly infinite dimensional, multivariable system, where z denotes the z-transform variable. The system operates on signals as $z(t) = G(q)x(t)$, with q the forward shift operator, i.e. $q^p x(t) = x(t+p)$, and with $x(t) \in \mathbb{R}^{q'}$ and $z(t) \in \mathbb{R}^{p'}$. Note that in the terminology as used here, the system and its transfer function are considered equivalent.

The ℓ_2-induced norm of the MIMO (multi-input multi-output) system G is defined as

$$\|G\|_{\ell_2,\text{ind}} := \sup_{\|x\|_2 \leq 1} \|Gx\|_2 .$$

The system G is called ℓ_2-stable if its ℓ_2-induced norm is finite. In that case the ℓ_2-induced norm is equal to the H_∞-norm of the system,

$$\|G\|_{\ell_2,\text{ind}} = \|G(z)\|_\infty ,$$

with

$$\|G(z)\|_\infty := \sup_{\omega \in [0,\pi]} \sigma_{\max}\left(G(e^{i\omega})\right),$$

where $\sigma_{\max}(\cdot)$ denotes the maximum singular value. In this thesis mainly the H_∞-norm of SISO (single-input single-output) systems are of interest, which is given by

$$\|G(z)\|_\infty = \sup_{\omega \in [0,\pi]} \left|G(e^{i\omega})\right|,$$

for scalar $G(z)$. Often it is assumed that the system's pulse response decays exponentially, i.e.

$$G(z) = \sum_{k=0}^{\infty} g(k) z^{-k},$$

with $g(k)$ the Markov-parameters of dimension $p' \times q'$, which are bounded by some bound $\bar{g}(k)$ with converges exponentially to zero for $k \to \infty$. This implies a smoothness assumption on the system's frequency response. In that case the H_∞-norm of the scalar system $G(z)$ is equal to

$$\|G(z)\|_\infty = \max_{\omega \in [0,\pi]} \left| G(e^{i\omega}) \right|.$$

The ℓ_∞-induced norm of the multivariable system G is defined as

$$\|G\|_{\ell_\infty,\text{ind}} := \sup_{\|x\|_\infty \leq 1} \|Gx\|_\infty .$$

The system G is called ℓ_∞-stable if its ℓ_∞-induced norm is finite. In that case the system has a pulse response representation,

$$G(z) = \sum_{k=0}^{\infty} g(k) z^{-k},$$

and the ℓ_∞-induced norm is equal to the ℓ_1-norm of the system,

$$\|G\|_{\ell_\infty,\text{ind}} = \|G(z)\|_{\ell_1} ,$$

with

$$\|G(z)\|_{\ell_1} := \max_i \sum_{j=1}^{q'} \sum_{k=0}^{\infty} |g_{ij}(k)|.$$

The H_∞- and ℓ_1-norms are related to each other as follows (Boyd and Doyle, 1987; Dahleh and Khammash, 1993),

$$\|G(z)\|_\infty \leq c_1 \|G(z)\|_{\ell_1} \leq c_2(n) \|G(z)\|_\infty ,$$

where c_1 is a constant depending only on the dimension of the matrix G, and c_2 is a linear function of the McMillan degree n of $G(z)$. For scalar systems $c_1 = 1$, which means that the ℓ_1-norm overbounds the H_∞-norm. Note that for finite dimensional systems ℓ_2-stability is equivalent to ℓ_∞-stability. For infinite dimensional systems ℓ_∞-stability implies ℓ_2-stability, but ℓ_2-stability does not imply ℓ_∞-stability. Generally in this thesis systems are assumed to belong to $\mathcal{A}$, the algebra of ℓ_∞-stable, linear, time-invariant, causal operators (which implies ℓ_2-stability).

The H_∞- and ℓ_1-norms satisfy the triangle inequality,

$$\|G(z) + H(z)\| \leq \|G(z\| + \|H(z)\| ,$$

for $G(z)$ and $H(z)$ of equal dimensions. Due to the fact that they are induced norms, they also satisfy the multiplicative property,

$$\|G(z)H(z)\| \leq \|G(z)\| \, \|H(z)\| ,$$

for $G(z)$ and $H(z)$ of proper dimensions.

2.2.3 Robustness Analysis

The H_∞-norm is the system norm used in H_∞-optimal feedback design (Zames, 1981; Francis, 1987; Doyle *et al.*, 1989; Morari and Zafiriou, 1989; Kwakernaak, 1993). The ℓ_1-norm is used in ℓ_1-optimal controller design (Vidyasagar, 1986; Dahleh and Pearson, 1987, 1988; McDonald and Pearson, 1991; Dahleh and Khammash, 1993). Both performance and robustness specifications can be realized in either of these two control design settings. Due to the character of the norms involved, the H_∞-control design methods are especially suited to realize frequency domain specifications, and ℓ_1-control design methods to realize time-domain specifications.

In both the H_∞- and the ℓ_1-setting robust stability and performance can be analysed by means of a small-gain theorem. First, it is shown that a robust performance problem can be recast into a robust stability problem. Consider closed loop configuration I depicted in Figure 2.1, which corresponds to a general performance robustness analysis problem. Here M is a stable, LTI, causal transfer function, containing dy-

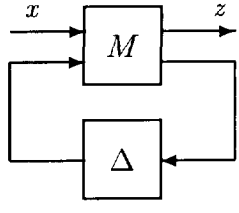

Fig. 2.1: Closed loop configuration I: performance robustness.

namics of the closed loop of model and controller. Δ is an uncertainty block, which is assumed to belong to the class $\mathcal{D}(p)$ given by,

$$\mathcal{D}(p) = \left\{ \Delta = \text{diag}\,(\Delta_1, \ldots, \Delta_p) \;\middle|\; \Delta_i \text{ is LTIC and } \|\Delta_i\|_{\ell_2,\text{ind}} \leq 1 \right\} \tag{2.1}$$

in the H_∞-case, and

$$\mathcal{D}(p) = \left\{ \Delta = \text{diag}\,(\Delta_1, \ldots, \Delta_p) \;\middle|\; \Delta_i \text{ is NLTVC and } \|\Delta_i\|_{\ell_\infty,\text{ind}} \leq 1 \right\} \tag{2.2}$$

in the ℓ_1-case. Here LTIC stands for linear, time-invariant and causal, and NLTVC stands for nonlinear, time-varying and causal. The ℓ_∞-induced norm of a nonlinear operator is defined as

$$\|\Delta_i\|_{\ell_\infty,\text{ind}} = \sup_{x \neq 0} \frac{\|\Delta_i x\|_\infty}{\|x\|_\infty}.$$

For simplicity it is assumed that each Δ_i is scalar.

The map from x to z in Figure 2.1 is denoted as T_{zx}. The closed loop system in Figure 2.1 is said to achieve *robust stability* if the closed loop is stable for all $\Delta \in \mathcal{D}(p)$. The closed loop system is said to achieve *robust performance* if

$$\sup_{\Delta \in \mathcal{D}(p)} \|T_{zx}\| < 1,$$

where the norm is the ℓ_2-induced norm in the H_∞-case, and the ℓ_∞-induced norm in the ℓ_1-case. Obviously, robust performance implies nominal performance and robust stability, where "nominal" refers to the situation with $\Delta = 0$.

Now consider closed loop configuration II depicted in Figure 2.2, which corresponds to a stability robustness problem, and which is formed from closed loop I by feeding z back to x through a fictitious perturbation Δ_P, satisfying $\|\Delta_P\| \leq 1$.

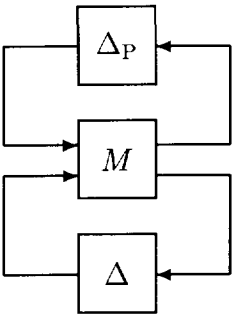

Fig. 2.2: Closed loop configuration II: stability robustness.

In the H_∞-case Δ_P is LTIC and the ℓ_2-induced norm is used. In the ℓ_1-case Δ_P is NLTVC and the ℓ_∞-induced norm is used. If the perturbations Δ_P and Δ in Figure 2.2 are taken together, the perturbation class is equal to $\mathcal{D}(p+1)$. It appears that the performance robustness problem of Figure 2.1 is equivalent to the stability robustness problem of Figure 2.2 in the sense that the closed loop system in Figure 2.1 achieves robust performance if and only if the closed loop system of Figure 2.2 achieves robust stability. This has been shown in Doyle (1982) for the H_∞-case and in Khammash and Pearson (1991, 1993) in the ℓ_1-case.

As a robust performance analysis problem can be recast into an equivalent robust stability analysis problem, it suffices to consider the latter. Some results concerning robust stability analysis are discussed next. Consider closed loop configuration III depicted in Figure 2.3. Again, M is assumed to be stable, LTI and causal, with

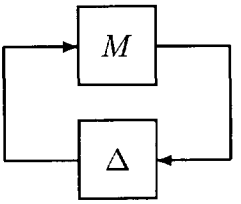

Fig. 2.3: Closed loop configuration III: stability robustness.

dimension $p \times p$, and Δ is assumed to be an element of the perturbation class $\mathcal{D}(p)$.

The following theorem gives a necessary and sufficient condition for robust stability of the closed loop in the H_∞-case.

Theorem 2.2.1 (Doyle, 1982) *The closed loop of Figure 2.3 possesses robust stability for all $\Delta \in \mathcal{D}(p)$, as specified by (2.1), if and only if*

$$\sup_{\omega \in [0,\pi]} \mu\left(M(e^{i\omega})\right) < 1,$$

where $\mu(\cdot)$ denotes the structured singular value, defined in Doyle (1982).

In Packard and Pandey (1993) it has been shown that the structured singular value is a continuous function of ω in case the perturbations are complex-valued, as is the case here and elsewhere in this thesis. This means that in general the robust stability condition becomes

$$\max_{\omega \in [0,\pi]} \mu\left(M(e^{i\omega})\right) < 1.$$

In the case the uncertainty block Δ is scalar ($p = 1$), the condition in the theorem simplifies to $\|M(z)\|_\infty < 1$.

The following theorem gives necessary and sufficient conditions for robust stability of the closed loop in the ℓ_1-case.

Theorem 2.2.2 (Khammash and Pearson, 1991, 1993) *The closed loop of Figure 2.3 possesses robust stability for all $\Delta \in \mathcal{D}(p)$, as specified by (2.2), if and only if*

$$\rho(\widehat{M}) < 1,$$

where $\rho(\cdot)$ denotes the spectral radius, and

$$\widehat{M} = \begin{bmatrix} \|M_{11}(z)\|_{\ell_1} & \cdots & \|M_{1p}(z)\|_{\ell_1} \\ \vdots & & \vdots \\ \|M_{p1}(z)\|_{\ell_1} & \cdots & \|M_{pp}(z)\|_{\ell_1} \end{bmatrix},$$

M_{ij} denoting element ij of matrix M.

In the scalar case ($p = 1$) the condition reduces to $\|M(z)\|_{\ell_1} < 1$. Note that the condition in the theorem is necessary and sufficient only if the perturbation class includes nonlinear or time-varying perturbations. If the perturbations are LTI, the condition is sufficient but not necessary. In that case the μ-condition is necessary and sufficient. Always the following inequality holds,

$$\rho(\widehat{M}) \geq \sup_{\omega \in [0,\pi]} \mu\left(M(e^{i\omega})\right).$$

The robust performance synthesis problem goes one step further than the robust performance analysis problem discussed above. Then the objective is to construct a controller such that the system is stabilized and robust performance is achieved. This problem is much more difficult. It is, for example, considered in Packard and Doyle

(1993) for the H_∞- or μ-case, and in Khammash and Pearson (1993) and Dahleh and Khammash (1993) for the ℓ_1-case.

Obviously, modern robust control design procedures can handle an upper bound on the H_∞- or ℓ_1-norm of the model error for robust stability and performance analysis or robust controller synthesis. This motivates the development of identification methods in this thesis which provide these upper bounds.

2.3 System Identification for High-Performance Robust Control

In this section the system identification problem as formulated in Section 1.3 is investigated in some detail. A solution strategy for the identification problem is proposed, which is quite non-standard. During the discussion a number of subproblems are formulated, which are further explored in the sequel of this thesis. In the discussion the main focus is on SISO systems. MIMO extensions are given in the course of the thesis.

Denote the (unknown) discrete-time *system* to be identified and controlled by $G_0(z)$. It is assumed that this system is an element of the set $\mathcal{A}$, i.e. that it is LTI, ℓ_∞-stable and causal. Note that the system is not necessarily finite-dimensional.

Remark 2.3.1 *It is emphasized that the linearity assumption does not automatically mean that the identification methods developed in this thesis are not applicable to nonlinear processes. In fact many tools described in this thesis are succesfully applicable to nonlinear systems, provided these systems have only small nonlinearities. In the application to the glass tube manufacturing process (Chapter 9) this has been taken care of by input excitation around a fixed operating point. The system can then be considered approximately linear in this operating point.*

The identification objective is to construct a finite-dimensional, LTI, stable and causal *nominal model* $\widehat{G}(z)$ and an explicit bound on the *model error* $G_0(z) - \widehat{G}(z)$ on the basis of measurement data. The nominal model and model error bound should be identified such that they are suited for use in high-performance robust control. The complexity (order) of the model is determined by the user. Often this order is restricted in view of the control application. For many robust control design procedures the order of the resulting controller is proportional to the order of the nominal model. As in industrial process applications often only low-order controllers can be implemented, this necessarily restricts the order of the nominal model.

The measurement data are given by N input data samples $u(t)$ and N measured output data samples $y(t)$. The disturbances or noise are assumed to be additive to the output,

$$y(t) = G_0(q)u(t) + e_0(t), \ t = 1, \ldots, N,$$

with $\{e_0(t)\}$ the noise process. The input signal may have been generated in closed loop. If that is the case it is assumed that also some external reference signal $r(t)$, $t = 1, \ldots, N$, is available, which is uncorrelated to the noise process, but correlated to the input signal. In general the collection of measurement data is preceded by an

identification *experiment design*, in which the model application is taken into account. This issue is discussed in Subsection 2.4.2.

In this thesis attention is restricted to the following weighted additive *uncertainty set*,

$$\mathcal{G} := \left\{ G(z) \;\middle|\; G(z) = \widehat{G}(z) + \Delta(z)W^{-1}(z),\ \|\Delta(z)\| \leq \delta \right\},$$

with the norm being either the H_∞- or the ℓ_1-norm. Here $W(z)$ is some LTI, finite dimensional, stable and stably invertible *weighting*, and δ is a norm-bound on the weighted model error. The structure of $\mathcal{G}$ is such that the set is perfectly suited to serve as basis for robust control design, as apparent from the discussion in Section 2.2. In view of the control application there is a user-defined restriction on the order of $W(z)$. This order determines the complexity of the uncertainty structure, and as such influences the complexity of the control design and the order of the resulting controller.

The objective is to identify $\mathcal{G}$, i.e. to construct $\widehat{G}(z)$, $W(z)$ and δ on the basis of measurement data, in such a way that the following statement can be made,

$$G_0(z) \in \mathcal{G}.$$

Consequently, a robust controller $C(z)$ designed to work adequately for all systems $G(z)$ in the set $\mathcal{G}$, also works adequately for the system $G_0(z)$. Unfortunately, it appears impossible to identify a bound on the model error from partial and corrupted data. This is impossible unless *prior assumptions* are made about the system and the noise corrupting the data, see also Ljung (1992), Wahlberg and Ljung (1992) and Hjalmarsson (1993a). If nothing is known about the noise, the measured output could just be only noise. If nothing is known about the system, it could be time-varying with radically changing dynamics just outside the measurement interval. In Subsection 2.4.3 it is explained what system and noise assumptions are precisely used in the identification procedures developed in this thesis. Formally speaking, the measured data in combination with the assumed prior information yields a set of *feasible* systems $\mathcal{F}$,

$$\mathcal{F} := \{G(z) \mid y(t) = G(z)u(t) + e(t),\ \forall t,\ \text{and } G(z),\ \{e(t)\} \text{ satisfy prior ass.}\},$$

which is the set of all systems consistent with data and prior information. If the assumptions about system and noise are correct, then

$$G_0(z) \in \mathcal{F}.$$

By definition the set $\mathcal{F}$ reflects the fundamental uncertainty about the system $G_0(z)$. But obviously, it is not suited for use in robust control design due to the fact that it is implicitly defined and highly unstructured. On the basis of this set $\mathcal{F}$, a model and a bound on the model error will be constructed.

In this thesis two different types of noise assumptions are utilized. The first type is a *deterministic* or unknown-but-bounded noise assumption. Typically this concerns the assumption that the amplitude of the noise is bounded by some known bound. This yields a deterministic feasible system set $\mathcal{F}$, which contains the system $G_0(z)$ with

probability 1. The second type of noise assumption is a *stochastic* noise assumption or the assumption that the noise is a stationary stochastic process and independent of the input signal. This yields a probabilistic feasible system set $\mathcal{F}$, which contains the system $G_0(z)$ with a certain probability.

Obviously, a noise assumption should be used which reflects reality, the assumption should be correct. In some cases a stochastic noise assumption would not be correct, due to the presence of non-stationary disturbances. But in many cases both a deterministic and a stochastic assumption about the noise is realistic. It is then important to choose that assumption that yields the smallest feasible system set, in order to reduce the system uncertainty as much as possible. An important issue in this thesis is the comparison of both types of noise assumptions. In particular in Part I of this thesis it is evaluated which type of noise assumption ultimately leads to the smallest model error bounds, i.e. is the least *conservative* type of noise assumption.

The identification objective can now be formulated as to construct the uncertainty set $\mathcal{G}$ on the basis of the feasible system set $\mathcal{F}$, in such a way that $\mathcal{G}$ outerbounds $\mathcal{F}$. More formally, the identification procedure is a mapping

$$\text{Identification: } \mathcal{F} \longrightarrow \mathcal{G},$$

such that

$$\mathcal{F} \subseteq \mathcal{G}.$$

Basically one would like to construct $\mathcal{G}$ such that it is exactly equal to the set $\mathcal{F}$. But in practice this is impossible, as the set $\mathcal{F}$ is highly unstructured, whereas $\mathcal{G}$ is highly structured. The set $\mathcal{G}$ is necessarily an approximation of the set $\mathcal{F}$. The fact that $\mathcal{G}$ outerbounds $\mathcal{F}$ then guarantees that $G_0(z) \in \mathcal{G}$. The outerbounding introduces some *conservatism*, which is unavoidable but negatively influences the control design. The designed controller has to be robust with respect to the set $\mathcal{G}$ instead of the smaller set $\mathcal{F}$, which further restricts the achievable controller performance. It is important that the conservatism in the outerbounding is kept as small as possible, in the sense that it causes as little controller performance degradation as possible. This implies that little conservatism is allowed in the feedback-relevant frequeny ranges, and possibly more in the other frequency ranges.

The identification problem formulated above is generally too complex to solve in one step. It is very difficult, if not impossible, to construct an identification algorithm which does the job in one step, i.e. identifies $\widehat{G}(z)$, $W(z)$ and δ on the basis of the set $\mathcal{F}$ in a feedback-relevant way. This is due to the fact that $\mathcal{F}$ is a highly unstructured set, often containing all kinds of implicit relations. The gap between this set and the highly structured and explicitly defined set $\mathcal{G}$ is just too big to take in one step. Therefore in this thesis the identification problem is tackled by adopting a two-step procedure. The starting point is the set $\mathcal{F}$, the set of systems consistent with data and prior information. In the first identification step this set $\mathcal{F}$ is outerbounded by an intermediate set $\mathcal{S}$, which is in turn outerbounded by the set $\mathcal{G}$ in the second step. This intermediate set $\mathcal{S}$ is a non-parametric *system uncertainty set* consisting of explicit relations, but still with only little structure. In this thesis two different types of system uncertainty sets $\mathcal{S}$ are considered. The first type specifies explicit bounds

on the system's frequency response,

$$\mathcal{S} = \left\{ G(z) \,\middle|\, G(e^{i\omega}) \in \mathcal{P}(\omega),\ \omega \in [0, \pi] \right\},$$

where, for each ω, $\mathcal{P}(\omega)$ is a specified region in the complex plane, referred to as an uncertainty region.

Remark 2.3.2 *Note the subtle difference in notation with respect to the frequency-dependency in $\mathcal{P}(\omega)$ and $G(e^{i\omega})$. The former just specifies the fact that $\mathcal{P}$ is frequency dependent, whereas the latter refers to the frequency response of a transfer function $G(z)$ evaluated on the unit circle. Consequently, the notation $G(\omega)$ would refer to some frequency response without assuming the existence of a generating transfer function $G(z)$.*

The second type of system uncertainty set specifies explicit bounds on the system's pulse response,

$$\mathcal{S} = \left\{ G(z) \,\middle|\, G(z) = \sum_{k=0}^{\infty} p(k) z^{-k},\ p(k) \in [p_{\mathrm{l}}(k), p_{\mathrm{u}}(k)],\ k = 0, \ldots, \infty \right\},$$

with $p_{\mathrm{l}}(k)$, $p_{\mathrm{u}}(k) \in \mathbb{R}$ specified lower and upper bounds. These definitions of the set $\mathcal{S}$ are a little bit crude. Precise definitions are given in Subsection 2.4.4.

The two-step identification problem can schematically be depicted as

$$\text{Identification: } \mathcal{F} \longrightarrow \mathcal{S} \longrightarrow \mathcal{G},$$

such that

$$\mathcal{F} \subseteq \mathcal{S} \subseteq \mathcal{G}.$$

In particular, the mapping from $\mathcal{F}$ to $\mathcal{S}$ is referred to as an *uncertainty bounding identification procedure*. Dependent on the type of prior information, this concerns deterministic or probabilistic identification. In fact, in Chapter 4 the problem is considered of constructing $\mathcal{S}$ on the basis of data and deterministic (non-probabilistic) noise assumptions. And in Chapter 5 the problem is considered of constructing $\mathcal{S}$ on the basis of data and stochastic (probabilistic) noise assumptions. Obviously, conservatism can hardly be avoided in this identification step. But in Chapter 4 and 5 it is attempted to introduce as little conservatism as possible.

The mapping from $\mathcal{S}$ to $\mathcal{G}$ can in fact not be called identification, as no data are involved. It concerns the construction of a nominal model $\widehat{G}(z)$, uncertainty weighting $W(z)$ and model error bound δ on the basis of the system uncertainty set $\mathcal{S}$. As, obviously, conservatism is involved in this step, the construction should be carried out in a feedback-relevant way. It is shown in Subsection 2.4.6 that it is possible to formulate a certain *global* identification criterion. In this criterion feedback-relevance can be incorporated by means of a suitable choice of weighting function which puts additonal emphasis on the frequencies around the cross-over frequency. Minimization of this criterion would all in once yield an optimal nominal model $\widehat{G}(z)$, an optimal weighting $W(z)$ and an optimal value δ. However, formulating the optimization problem is one

thing, solving it is another thing. For computational reasons it appears very hard to solve this global identification problem. For that reason the problem of constructing $\mathcal{G}$ on the basis of $\mathcal{S}$ is split up into two steps as well.

In the first step a nominal model $\widehat{G}(z)$ is constructed on the basis of the set $\mathcal{S}$, minimizing a certain feedback-relevant *nominal identification criterion*. In the second step the uncertainty weighting $W(z)$ and model error bound δ are constructed on the basis of the set $\mathcal{S}$ and the specified nominal model $\widehat{G}(z)$. The first step, the construction of a nominal model, is in this thesis referred to as a *system approximation* problem. Schematically this two-step procedure can be depicted as follows,

$$\text{System approximation: } \mathcal{S} \longrightarrow \widehat{G}(z),$$

and

$$\mathcal{S},\ \widehat{G}(z) \longrightarrow W(z),\ \delta.$$

In particular the following nominal identification criterion is considered in the system approximation step,

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'(z) \right) \right\|,$$

with $\mathcal{M}$ some specified *model set*. The norm used in this criterion is the same norm as used in the control design, i.e. either the H_∞-norm or the ℓ_1-norm. $\widetilde{W}(z)$ is some specified stable and stably invertible, LTI, finite dimensional weighting function, which can be used to emphasize the need of a good model fit in the feedback-relevant frequency ranges. The nominal model $\widehat{G}(z)$ is identified in such a way that the H_∞- or ℓ_1-norm of the worst-case model error is minimized, *also in the non-asymptotic case*, i.e. if a finite number of data is available. This is possible due to the fact that the set $\mathcal{S}$ is defined also for a finite number of measurement data. The fact that the procedure is optimal for the asymptotic as well as the non-asymptotic case, is an important property of the identification procedure. This as opposed to standard identification methods for which only asymptotic optimality results are available. In Subsection 2.4.5 and in Chapter 7 more is said about this system approximation problem.

The second step concerns the construction of an uncertainty weighting $W(z)$ and model error bound δ. For a given weighting $W(z)$ the model error bound is defined as

$$\delta = \max_{G(z) \in \mathcal{S}} \left\| \left(G(z) - \widehat{G}(z) \right) W(z) \right\|,$$

i.e. the norm of the worst-case weighted model error with respect to the system uncertainty set $\mathcal{S}$. Again, the norm is either the H_∞- or the ℓ_1-norm. For the H_∞-case this problem has also been considered in e.g. De Vries and Van den Hof (1992, 1994a) and Bayard and Yam (1993). The analogous formulation for the ℓ_1-case is new. On the one hand, the issue is the construction of a suitable weighting $W(z)$ which captures the shape of the model error. On the other hand, the actual calculation of δ according to its definition needs to be investigated. These issues are further explored in Subsection 2.4.6.

The final outcome of the identification procedure sketched above is the uncertainty set $\mathcal{G}$. Similar to current practice in standard identification procedures this identification result may be subjected to *model validation*. More is said about this in Subsection 2.4.7.

2.4 Identification Procedure

2.4.1 Introduction

In the previous section a procedure has been sketched, which aims at the identification of a nominal model and a model error bound suited for use in high-performance robust control. The main items in this identification procedure are schematically depicted in Figure 2.4. All steps in this scheme are passed through sequentially in the direction of the arrows. The central item in this identification scheme is the system uncertainty set $\mathcal{S}$, which is identified on the basis of data and prior information about system and noise. Note that the set of feasible systems, $\mathcal{F}$, is not explicitly shown in Figure 2.4. The reason for this is that this set is rather artificial. In case deterministic noise assumptions are used, as in Chapter 4, it is straightforward to define the set $\mathcal{F}$, although by implicit relations. However, in case stochastic noise assumptions are used, as in Chapter 5, it is very hard to give a consistent definition of the set $\mathcal{F}$. Therefore, in Figure 2.4, and in Chapter 5, less emphasis is put on this set $\mathcal{F}$.

The items in the identification scheme in Figure 2.4 have already been discussed briefly. In the following subsections the steps in the identification scheme are discussed in more detail. In the course of this discussion precise formulations are given of the subproblems which are investigated and solved in the sequel of this thesis.

2.4.2 Experiment Design

The first step in the entire identification procedure is experiment design, i.e. choice of input signal. In general terms the input signal should be sufficiently rich such that the important system dynamics can be identified from the data (Ljung, 1987). To be more specific, the identification experiment should be such that the system can be accurately identified in the frequency ranges important for the control application, without allowing too much model error in the other frequency ranges.

Besides the intended use of the model, possible experimental restrictions have to be taken into account as well, such as measurement time, input amplitude constraints and smoothness conditions on the input signal. Such restrictions are not uncommon for industrial processes. Typically sine-sweep experiments can not be carried out due to measurement time restrictions, and impulse response experiments can not be carried out due to input amplitude and smoothness restrictions. Moreover, sometimes it is unavoidable that closed loop experiments are carried out, due to instability of the uncontrolled process, or the demand that the nominal process operation is not much affected by the experiments. Concluding, the experiment design should be driven by the control objective and the process restrictions.

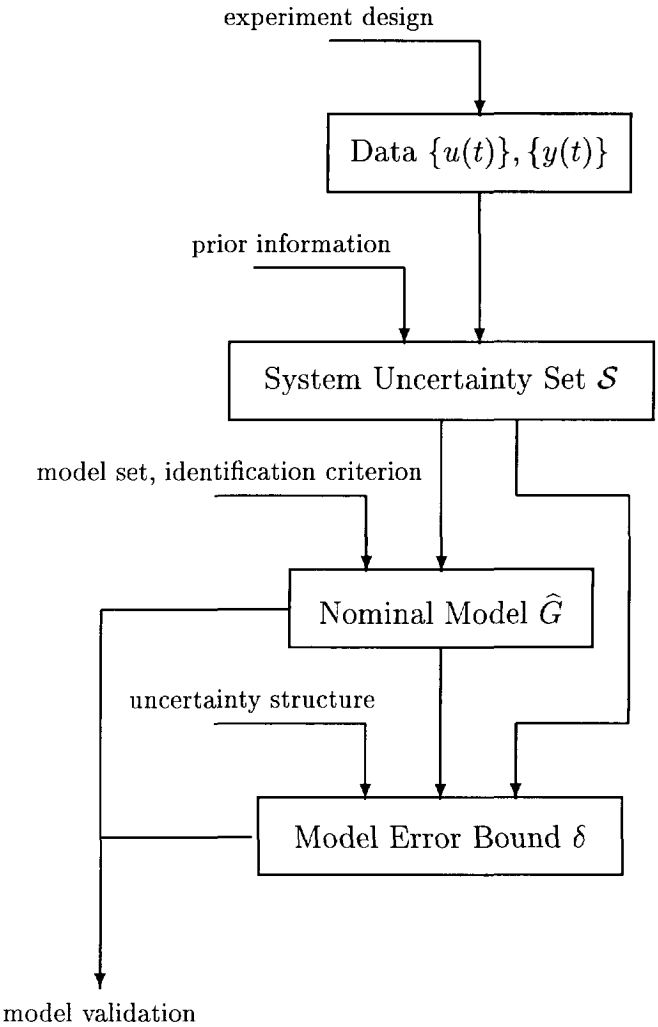

Fig. 2.4: Identification procedure.

Many researchers have considered the problem of (optimally) designing an identification experiment, see e.g. Goodwin and Payne (1977), Gevers and Ljung (1986), Wahlberg and Ljung (1986) and Ljung (1987) in the prediction error framework, and Pronzato and Walter (1989), Mäkilä (1991b) and Tse *et al.* (1991) in a worst-case identification setting.

In the context of the identification approach developed in this thesis the identification experiment in particular influences the size and shape of the set of feasible systems $\mathcal{F}$, and as such the size and shape of the system uncertainty set $\mathcal{S}$. More input energy

in a certain frequency region reduces the system uncertainty for these frequencies. It is important that the identification experiment is performed such that the system can be accurately identified in feedback-relevant frequency ranges, such as near the cross-over frequency. In fact this means that, if possible, the identification experiment should be tuned towards the control objective.

In this thesis the problem of *optimally* designing the identification experiment is not considered. In the examples and application described, an ad hoc approach is adopted that the input signal should contain relatively more energy in the feedback-relevant frequency range. This is automatically the case if experiments are performed in closed loop. Indeed often the optimal feedback-relevant identification experiment design appears to be in closed loop, see e.g. Hansen *et al.* (1989), Schrama (1992a), Zang *et al.* (1992) and Hakvoort *et al.* (1994). The optimal experiment design then requires knowledge of the controller to be designed. This can only be solved in an iterative procedure of identification and subsequent control design, see Schrama (1992b).

2.4.3 Prior Information

As observed before, system uncertainty bounds can only be identified if prior assumptions are made about the system $G_0(z)$ and the noise process $\{e_0(t)\}$. It is important that these prior assumptions satisfy the following requirements:

- *Correctness* or *validity* of the prior assumptions, and the possibility to check this. For example, the assumption of zero noise is, though convenient, invalid for most, if not all, practical situations. It is also important that the validity of the prior assumptions can be checked on data.

- *Non-conservatism* of the prior assumptions. Specifically, does the prior information about the noise really capture the essential features of the noise process? If not, unnecessarily large error bounds will result. This is called conservatism, as with another set of prior assumptions smaller error bounds can be obtained, using the same data set.

- *Employability* of the prior assumptions in the procedure to estimate a system uncertainty set. Typically this requirement is conflicting with the previous ones, but obviously it is important. If one is able to specify a nice set of prior assumptions about system and noise, but there is no procedure which can handle this prior information in the derivation of system uncertainty regions, it is the end of the exercise. In this thesis new procedures are developed which can handle various kinds of noise assumptions, and which are very flexible. However, always a compromise has to be made between employability and practical validity.

Remark 2.4.1 *With respect to the required correctness of the prior assumptions the following remark can be made. Often, because of the required employability of the prior assumptions, simplifying assumptions are made, such as linearity of the system $G_0(z)$ or stationarity of the noise $\{e_0(t)\}$. Such a simplifying assumption can be justified if first, it does not deviate too much from reality, and second, the identification procedure*

is robust to small errors in the prior assumptions. The identification procedure is called robust if small errors in the prior assumptions give rise to a small error in the identified model error bound.

In the procedures developed in this thesis typically the following assumptions are made:

- The system $G_0(z)$ is assumed to belong to $\mathcal{A}$. Let a system representation be as follows,

$$G_0(z) = \sum_{k=0}^{\infty} g_0(k) B_k(z),$$

where $\{B_k(z)\}_{k=0,\ldots,\infty}$, is some user-defined set of basis functions given by

$$B_k(z) = \sum_{k'=0}^{\infty} b_k(k') z^{-k'}, \ k = 0, \ldots, \infty,$$

for given and known scalar pulse response parameters $b_k(k')$. These basis functions can, for example, be chosen as the pulse functions $B_k(z) = z^{-k}$, or the Laguerre functions $B_k(z) = \sqrt{1-a^2}z(1-az)^k/(z-a)^{k+1}$ (Wahlberg, 1991), or Kautz functions (Wahlberg, 1994), or generalized orthonormal basis functions, see Heuberger *et al.* (1994) and Ninness (1993, Ch. 2). More is said about these basis functions in Appendix 4.B in Chapter 4.

The (unknown) coefficients $g_0(k)$ can be considered as generalized pulse response parameters of the system $G_0(z)$. Although they are unknown, they are assumed to be bounded by

$$|g_0(k)| \leq \bar{g}(k), \ k = 0, \ldots, \infty,$$

for given $\bar{g}(k)$.

- In order to cope with unknown initial conditions the input signal in the past is assumed to be bounded by

$$|u(t)| \leq \bar{u}, \ \forall t \leq 0,$$

for some given $\bar{u}$. This bound may result from actuator constraints and need not be very tight, as in general its influence on the identification result is restricted.

- The noise $\{e_0(t)\}$ is additive to the output,

$$y(t) = G_0(q)u(t) + e_0(t), \ t = 1, \ldots, N,$$

and satisfies a deterministic or a stochastic noise assumption.

In this thesis both an uncertainty bounding identification procedure is developed which employs deterministic noise assumptions, and one which employs stochastic noise assumptions. A typical deterministic noise assumption is the unknown-but-bounded-amplitude noise assumption

$$|e_0(t)| \leq \bar{e}(t), \ t = 1, \ldots, N,$$

for some specified bounds $\bar{e}(t)$. This noise assumption is also utilized in the so-called bounded error, parameter bounding or set membership identification literature, see Walter and Piet-Lahanier (1990) and Milanese and Vicino (1991) for overviews. This noise assumption will lead to deterministic uncertainty regions and error bounds, which are correct with probability 1. Of course this is only under the condition that the prior information, such as the specified noise bound, is correct. If the fact is incorporated that it is in general very hard, if not impossible, to prove that the prior information is correct, and that often the prior information has to be estimated from data, then it is clear that no absolute guarantees can be given.

It is well known that the unknown-but-bounded-amplitude noise assumption can be quite conservative if the noise is a random process. This is due to the following two reasons:

- No averaging properties of the noise are incorporated in this noise description. In particular the noise is assumed to be possibly heavily correlated to the input signal.
- The noise bounds have to be chosen conservative, in order to guarantee their correctness.

In close relation to this it has been shown by Ninness (1993) that a procedure employing this noise description is very sensitive to noise outliers. For that reason in this thesis, in Chapter 3, also alternatives are presented for the standard time-domain noise constraint as given above. In particular this concerns the introduction of cross-covariance and frequency-domain noise constraints. As opposed to the time-domain noise constraints, the cross-covariance noise constraints appear to possess certain averaging properties and are more robust to noise outliers.

A typical stochastic noise assumption is

> The noise process $\{e_0(t)\}$ is stationary, satisfying $e_0(t) = H_0(q)w_0(t)$ for some stable $H_0(q)$, with $\{w_0(t)\}$ a sequence of independent random variables with zero mean values, variances λ_0 and bounded fourth moments. Moreover, the noise $\{e_0(t)\}$ is uncorrelated to the input signal $\{u(t)\}$ (in open loop) or the external reference signal $\{\bar{r}(t)\}$ (in closed loop).

Typically this type of noise assumption leads to statistical uncertainty regions and model error bounds, which are correct with any prespecified probability, for example probability 0.999. This in turn means that a robust controller can be designed which has guaranteed performance with high probability. This is an acceptable situation for most engineers, not in the least because it is consistent with engineering practice where no absolute guarantees exist and small failure risks are unavoidable.

As mentioned before the prior information should be correct. This means that, if the noise is a deterministic signal, bounded by some known bounds, then the deterministic noise assumption is, obviously, the one that should be used. If, on the other hand, the noise is a realization of a stochastic process, possibly unbounded or with unknown bounds, then the stochastic noise assumption should be used. However, if the noise is both, it is bounded and behaves stochastic, then it could be unclear which noise

assumption should be used. It is too cheap to favour the deterministic approach for the reason that this approach yields error bounds which are correct with probability 1, where the stochastic approach yields error bounds which are correct with, for example, probability 0.999. If the price paid for this additional certainty is a substantial increase in the size of the error bounds, the statistical error bounds are definitely preferable. The fact that in this thesis both a procedure for deterministic identification and a procedure for probabilistic uncertainty bounding identification are developed, which utilize the same system priors, offers the nice possibility to compare the outcomes of these methods, if applied to the same data set.

2.4.4 System Uncertainty Set

In the identification procedure adopted in this thesis the data and prior information are used to construct a system uncertainty set, denoted by $\mathcal{S}$. New uncertainty bounding identification algorithms are developed for this purpose, which can handle the data and prior information as specified in Subsection 2.4.3. Besides the requirements on the prior information utilized in the procedure to derive a system uncertainty set, it is important that the identification procedure itself satisfies the following requirements:

- *Correctness* of the resulting uncertainty set. Not only the prior information should be correct, but it should also be processed in a correct way. For example, one may think of a procedure with nonlinear optimizations. In that case easily a local optimum can be obtained and, consequently, the results can be incorrect. Besides that, it is also important to consider if the procedure yields non-asymptotic results, or if it relies on asymptotic results, e.g. in the number of data samples. In the latter case one should evaluate the rate of convergence, how accurate are these asymptotic results if they are applied to a non-asymptotic situation? This is important as in many industrial applications only short data records are available.

- *Non-conservatism* of the procedure. Because of computational aspects it is often not possible to find a one-to-one mapping from data and a priori information (the set $\mathcal{F}$) to a posteriori information, the system uncertainty set $\mathcal{S}$. In general discrepancies are included. It is important that these discrepancies do not affect the correctness of the results, and therefore they are necessarily conservative, which means that they increase the size of the resulting uncertainty set. This is called conservatism, and it adds to the conservatism in the prior assumptions. It is important that this conservatism is kept as small as possible.

- *Applicability* of the identification procedure. The procedure should be able to handle arbitrary input signals, possibly generated in closed loop, and for multivariable processes with simultaneous excitation of the various input signals.

- *Insensitivity* to the prior assumptions and *robustness* to errors in the priors. The procedure should be rather insensitive to prior information that one is not sure about. And the procedure should be robust to small nonlinearities and noise outliers in order to be practically applicable.

- *Proper format* of the resulting uncertainty set. Obviously, the procedure should yield a system uncertainty set in a format suited for further use in model error bounding procedures, or even directly in robust control synthesis or analysis. Uncertainty intervals for the system's frequency or pulse response are considered suitable a posteriori information.

- *Restricted computational complexity* of the identification algorithm. Though current computers are extremely fast, computation time can still be a fundamental limitation.

In this thesis two types of system uncertainty sets $\mathcal{S}$ are considered. The first type involves convex frequency response uncertainty regions $\mathcal{P}(\omega_j)$ in the complex plane for a finite number (l) of (user-defined) frequencies as specified by the set Ω,

$$\Omega = \{\omega_1, \ldots, \omega_l\},\ 0 \leq \omega_1 < \cdots < \omega_l \leq \pi.$$

In combination with a uniform bound β_1 on the first derivative of the system's frequency response, which can be established by using the prior bound on the system's (generalized) pulse response parameters, this gives a uniform bound on the system's frequency response. So, this first type of a posteriori information is given by the frequency response system uncertainty set

$$\mathcal{S} = \left\{ G(z) \in \mathcal{A} \;\middle|\; G(e^{i\omega_j}) \in \mathcal{P}(\omega_j),\ \forall \omega_j \in \Omega,\ \left| \frac{\mathrm{d}\, G(e^{i\omega})}{\mathrm{d}\,\omega} \right| \leq \beta_1,\ \forall \omega \in [0, \pi] \right\}. \quad (2.3)$$

This type of system uncertainty set appears useful in view of H_∞- or μ-optimal controller analysis and synthesis. Once a nominal model $\widehat{G}(z)$ and a weighting $W(z)$ have been specified, it appears straightforward to evaluate the H_∞-norm of the worst-case weighted model error,

$$\max_{G(z)\in\mathcal{S}} \left\| \left(G(z) - \widehat{G}(z) \right) W(z) \right\|_\infty = \max_{G(z)\in\mathcal{S}} \max_{\omega\in[0,\pi]} \left| \left(G(e^{i\omega}) - \widehat{G}(e^{i\omega}) \right) W(e^{i\omega}) \right|,$$

where the definition of the H_∞-norm has been substituted. Moreover, this type of uncertainty set appears especially useful in formulating and solving H_∞-optimal nominal identification problems.

The second type of a posteriori information involves pulse response uncertainty intervals $[p_\mathrm{l}(k), p_\mathrm{u}(k)]$, possibly with a finite dimensional, stable and stably invertible weighting $W(z)$ incorporated. In particular this second type is given by the pulse response system uncertainty set

$$\mathcal{S} = \left\{ G(z) \in \mathcal{A} \;\middle|\; G(z) = W^{-1}(z) \sum_{k=0}^{\infty} p(k) z^{-k},\ p(k) \in [p_\mathrm{l}(k), p_\mathrm{u}(k)],\ k = 0, \ldots, \infty \right\}. \quad (2.4)$$

This type of system uncertainty set appears useful in view of ℓ_1-robust control design and analysis. Once a nominal model $\widehat{G}(z)$ has been specified, it appears straightforward

to evaluate the ℓ_1-norm of the worst-case weighted model error,

$$\max_{G(z)\in\mathcal{S}} \left\| \left(G(z) - \widehat{G}(z)\right) W(z) \right\|_{\ell_1} = \max_{p(k)\in[p_l(k),p_u(k)]} \sum_{k=0}^{\infty} |p(k) - \widehat{p}(k)|,$$

where the definition of the ℓ_1-norm has been substituted, and $\widehat{p}(k)$ denote the pulse response parameters of $\widehat{G}(z)W(z)$. Moreover, this system uncertainty set appears very useful in formulating and solving ℓ_1-optimal nominal identification problems.

In Helmicki *et al.* (1991) an identification problem is formulated which makes use of system information in a format similar to (2.3). The identification procedure sketched in that paper resembles the procedure adopted here. In both cases frequency response uncertainty regions are identified, which are next used to identify a nominal model. However, in Helmicki *et al.* (1991) quite different requirements are imposed on the data, the nominal model and the model error bound, as discussed in Subsection 2.4.5. The formulation of the pulse response system uncertainty set $\mathcal{S}$ in (2.4) is new. Such an uncertainty set has not been used in a system identification setting before.

Remark 2.4.2 *Note that in this thesis just one symbol, $\mathcal{S}$, is used for both system uncertainty sets (2.3) and (2.4). From the context it will always be clear which is meant, as the first is only used in an H_∞-identification setting and the second only in an ℓ_1-identification setting. Furthermore, note that in the ℓ_1-case the system uncertainty set (2.4) is a function of the specified weighting $W(z)$, though this is not made explicit in the notation.*

For both types of a posteriori information typically the following statement can be made,

$$G_0(z) \in \mathcal{S}, \text{ w.p. } \geq \alpha,$$

provided the assumed prior information is correct. Here α denotes some confidence level, $\alpha = 1$ in case of deterministic prior assumptions, and $\alpha < 1$ in case of stochastic prior assumptions. The system uncertainty set $\mathcal{S}$ typically reflects the fundamental uncertainty about the system $G_0(z)$ due to partialness and corruptedness of the data.

Note that there is some conservatism intrinsically present in the formats of $\mathcal{S}$ as presented here, as the interaction between frequencies ω_j or pulse response parameters $p(k)$ is lost. It may be possible to derive uncertainty intervals $[p_l(k), p_u(k)]$ which are non-conservative for each pulse response parameter separately. This then means that, with the information available in data and prior information, the possibility can not be excluded that the system attains an extreme value for any k. However, it can, for example, be very unlikely that the system attains a maximum for k and a minimum for $k+1$, which possibility is not excluded in the given description of $\mathcal{S}$. The same holds for the uncertainty regions $\mathcal{P}(\omega_j)$ as a function of ω_j.

In Chapter 4 a procedure for deterministic uncertainty bounding identification is developed, which can handle the deterministic noise assumptions mentioned in Subsection 2.4.3, both time-domain noise constraints and cross-covariance and frequency-domain noise constraints. With this method both frequency and pulse response uncertainty intervals can be identified. These are constructed by solving a set of linear programming (LP) problems.

In Chapter 5 a procedure for probabilistic uncertainty bounding identification is developed, which can handle the stochastic noise assumption as described in the previous section. The procedure in fact involves the explicit calculation of the bias and variance error of an instrumental variable or FIR output error estimate. This statistical method fully employs the averaging properties and random nature of the noise. An a posteriori trade-off can be made between size of the confidence regions and probability of correctness. This stochastic uncertainty bounding method also yields both frequency response and pulse response confidence regions.

It is shortly discussed how these two new methods differ from other deterministic and stochastic uncertainty bounding identification methods as found in literature. The interested reader is also referred to Hjalmarsson (1993a) and De Vries (1994) for recent overviews.

The procedure for deterministic identification as described in Wahlberg and Ljung (1992), utilizes the unknown-but-bounded-amplitude noise constraint. As indicated before, this can be a conservative noise assumption. Moreover, the procedure adds much conservatism, due to the use of ellipsoidal outer bounding algorithms. For those reasons this procedure typically yields frequency domain system uncertainty regions which are too large to be practically useful.

The methods for deterministic identification described in Lamaire *et al.* (1991), Van den Boom *et al.* (1991) and De Vries and Van den Hof (1992, 1994a), utilize frequency domain noise constraints. These procedures only work satisfactorily with sinusoidal input excitation, otherwise overly large system uncertainty regions will be identified.

The procedure for probabilistic uncertainty bounding identification as described in De Vries and Van den Hof (1993, 1994b) and De Vries (1994), requires periodic input excitation, which restricts industrial applicability. However, the method has succesfully been applied to mechanical systems, where less severe input signal restrictions exist.

The stochastic method presented in Bayard (1992) assumes gaussian noise with known noise generating filter. The method requires sinusoidal excitation. Succesful application to a mechanical system has been reported.

In Bai (1991) and Younce and Rohrs (1992) the noise-free situation is considered.

The stochastic approach described in Zhu (1989) is an approach based on asymptotic variance expressions for prediction error identification. This method only yields correct results if a large number of measurement data is available.

In the procedure described in Goodwin *et al.* (1990), Ninness and Goodwin (1992) and Ninness (1993) stochastic assumptions are made about both the noise and the system. However, it seems rather difficult to prove validity of the assumption that the unmodelled dynamics are a realization of a stochastic process. Moreover, due to numerical complexity of the method, application is restricted to low-order systems. In Schoukens and Pintelon (1994) the following remark is made about this uncertainty bounding identification method:

> "The simulations show that the technique does not give an exact idea about the model errors, but a rough guess about the order of magnitude of possible model errors is obtained."

Obviously, this is rather unsatisfactory. An uncertainty bounding identification method should be able to do better than this.

The deterministic identification method described in Helmicki *et al.* (1991) has a very restricted applicability, as a sinusoidal input signal is needed to identify frequency response uncertainty regions. Similarly, the method described in Jacobson *et al.* (1992) is too restrictive, as pulse or step input signals are needed for the construction of ℓ_1-model error bounds.

In standard parameter bounding identification methods (Norton, 1987; Clement and Gentil, 1988; Belforte *et al.* 1990; Walter and Piet-Lahanier, 1990; Norton and Veres, 1991; Milanese and Vicino, 1991) the system uncertainty set is represented by parameter bounds in a parametrized transfer function. Although these identification methods have originally been developed for purposes other than robust control design, it is sometimes claimed (Norton, 1987) that the parameter bounds can be used in μ-control synthesis or analysis, by means of a real-valued structured uncertainty block. However, to the author's opinion, applicability in control problems is very restricted due to the following difficulties:

- Standard parameter bounding identification utilizes the unknown-but-bounded-amplitude noise assumption, which can give overly large parameter bounds, due to conservatism of this noise prior. But, as mentioned before, this problem is dealt with in this thesis by the introduction of alternative and less conservative noise priors.

- The parametrized transfer function should have the same complexity (order) as the system, in order to obtain correct results. However, besides that the system order is generally unknown, in practice one is often interested in a model of reduced complexity, both for identifiability reasons and because of applicability in the control design. An interesting approach is taken in Lau *et al.* (1990) and Kosut *et al.* (1992), where an attempt is made to solve this problem by using both structured parametric and unstructered nonparametric uncertainty. The latter then takes account of the undermodelling. However, the nonparametric uncertainty is bounded by an a priori specified H_∞-bound, where it seems more to the point to estimate this bound from data.

- Even for low-order models the μ-synthesis problem can easily become too complicated to be computationally feasible.

- There is no straightforward frequency domain interpretation of bounds on the parameters of a parametrized transfer function. For example, it is well known that a small perturbation of a denominator parameter may cause an arbitrarily large perturbation of the frequency response. On the other hand, it is easy to construct examples where large parameter perturbations correspond to arbitrarily small frequency response perturbations, especially if there is almost pole-zero cancellation. However, many feedback properties, such as stability, have a frequency domain interpretation. It is important to be able to tune the (upper bound on the) model error towards the control objective. This can only be done

if there is the possibility to include frequency domain weighting functions. A parameter bounding identification setting seems unfit for this.

2.4.5 Model Set and Identification Criterion

Once the system uncertainty set $\mathcal{S}$ has been identified, the next step in the identification scheme of Figure 2.4 is the selection of a model set and identification criterion for the identification of a nominal model. Typically the choice of model set and model order is a trade-off between

- *Accuracy.* The model set should be flexible enough, and the model order high enough, to be able to capture the relevant system dynamics, with respect to the intended model application.
- *Applicability.* Typically a linear model set is required such that the nominal model can be used in linear control design procedures. The model order should not be too high, in order to achieve a reasonable complexity of the control design procedure and resulting controller.

In this thesis the following linear (SISO, discrete-time) model set is considered,

$$\mathcal{M}: \ \widehat{G}'(z) = \frac{n(z)}{d(z)}, \ n(z) = \sum_{k=0}^{n} n_k z^k, \ d(z) = z^d + \sum_{k=0}^{d-1} d_k z^k,$$

with additionally the requirement that the model is stable. The numerator degree n and denominator degree d are user-defined. It is mentioned that the problem of optimal model order selection, though certainly of interest, is outside the scope of this thesis.

Once the model set has been selected, it is necessary to specify the identification criterion. Note that the particular choice of identification criterion is less important if the model error is small in all frequency ranges, see e.g. Caines and Baykal-Gürsoy (1989), where L_∞-consistency of L_2-estimators is discussed. However, if a large bias can be expected, for example due to deliberate undermodelling as the application requires a relatively simple model, it is important to tune the model error towards the application (control design). This can be done by choosing an appropriate identification criterion. If the identification criterion is tuned towards the control objective, the identification of the nominal model is called *control-oriented* or *feedback-relevant*.

In this thesis the following nominal identification criterion is considered,

$$\min_{\widehat{G}\in\mathcal{M}} \max_{G\in\mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}(z) \right) \right\|, \tag{2.5}$$

with $\widetilde{W}(z)$ some stable and stably invertible, LTI, finite dimensional weighting function, and the norm being either the H_∞- or the ℓ_1-norm. The norm used in the identification thus coincides with the norm used in the robust control design. The criterion is a worst-case error criterion in the sense that the maximum model error

with respect to the system uncertainty set $\mathcal{S}$ is minimized. The weighting can be chosen control-relevant, by which is meant here that additional emphasis is put on the closed-loop relevant frequencies. For example, the choice

$$\widetilde{W}(z) = \frac{C(z)}{1 + C(z)G(z)} \tag{2.6}$$

emphasizes frequencies near the critical point -1, and is as such feedback-relevant, see Schrama (1992a, Ch. 2). Another possibility is to use a desired sensitivity as a weighting function (Rivera *et al.*, 1992). In this thesis it is not further investigated which choice of weighting $\widetilde{W}(z)$ in (2.5) is optimal in view of the control application.

The problem of specifying an identification criterion which is *optimally* tuned towards the control objective has had considerable attention in literature, see e.g. Gevers (1993) for a discussion on the subject. Control-oriented identification methods have been developed in Bayard *et al.* (1992), Zang *et al.* (1992), Lee *et al.* (1993) and by many others. In this area there are many open questions and unsolved problems. In the sequel of this subsection attention is paid to the problem of specifying an optimal feedback-relevant identification criterion. The problem is only shortly addressed. It is certainly not claimed that all questions will have been answered at the end of this subsection.

An interesting and promising approach to the control-oriented identification problem is taken in Schrama (1992a, 1992b), where a connection is made between a certain quite general control objective and the nominal identification criterion. This control objective concerns the following minimization problem,

$$\min_{C} \left\| W_1 T(\widehat{G}, C) W_2 \right\|, \tag{2.7}$$

for some given nominal model $\widehat{G}$ and weighting matrices W_1 and W_2, with the matrix T defined as

$$T(G, C) := \begin{bmatrix} I \\ G \end{bmatrix} (I + CG)^{-1} \begin{bmatrix} I & C \end{bmatrix}.$$

A high controller performance corresponds to a small norm of the weighted T-matrix. In Chapter 8 and 9 a control design procedure is used which utilizes the H_∞-norm in the minimization. Also, certain H_2- or ℓ_1-control design objectives can be recast into this format. In Schrama (1992a) it is argued that one is interested in the performance for the true system G_0, given by

$$\left\| W_1 T(G_0, C) W_2 \right\|.$$

Using the triangle inequality the following bound can be established for this performance,

$$\left\| W_1 T(G_0, C) W_2 \right\| \leq \left\| W_1 T(\widehat{G}, C) W_2 \right\| + \left\| W_1 \left(T(G_0, C) - T(\widehat{G}, C) \right) W_2 \right\|.$$

It is then argued that the identification should determine a nominal model $\widehat{G}$ as the solution to the following minimization problem,

$$\min_{\widehat{G}\in\mathcal{M}} \left\| W_1 \left(T(G_0, C) - T(\widehat{G}, C) \right) W_2 \right\|, \tag{2.8}$$

for some given controller C. In that case the controller performance for the system G_0 is made small, though not minimized, by the combination of control design and identification, as follows from the triangle inequality. As the control design step utilizes a given nominal model $\widehat{G}$, and the identification a given controller C, an iterative procedure of identification and control design is unavoidable. Such an iterative procedure is typical for most control-oriented identification schemes.

Basically this approach is perfectly sound, if the identification could really minimize the specified criterion. However, this criterion contains the unknown system G_0. The solution provided in Schrama (1992a) therefore heavily relies on the asymptotic situation $N \to \infty$. In that case perfect implicit or explicit system knowledge becomes available, in the sense that the variance error vanishes, see e.g. the consistency results for prediction error identification in Ljung (1987). Note that this drawback holds for many currently available control-oriented identification schemes. Often optimality of the scheme can only be shown in the asymptotic case. It is clear that for finite N the given identification criterion (2.8) can not be minimized. A modification is needed, which is presented here.

In the identification procedure adopted in this thesis, a system uncertainty set $\mathcal{S}$ has become available. This uncertainty set contains the system G_0 (with high probability). The knowledge (or lack of knowledge) about the system G_0 is represented by this set $\mathcal{S}$. This is true both in the asymptotic case and the non-asymptotic case, i.e. for a finite number of data. The following identification criterion is proposed as a natural modification of the criterion (2.8),

$$\min_{\widehat{G}\in\mathcal{M}} \max_{G\in\mathcal{S}} \left\| W_1 \left(T(G, C) - T(\widehat{G}, C) \right) W_2 \right\|, \tag{2.9}$$

for some given controller C. This is supported by the following triangle inequality,

$$\max_{G\in\mathcal{S}} \|W_1 T(G, C) W_2\| \leq \left\| W_1 T(\widehat{G}, C) W_2 \right\| + \max_{G\in\mathcal{S}} \left\| W_1 \left(T(G, C) - T(\widehat{G}, C) \right) W_2 \right\|.$$

The controller design minimizes the first contribution at the right, and the identification the second contribution at the right-hand side. Therefore the worst-case cost at the left and, consequently, the cost for the system G_0, is made small, though not necessarily minimized, by the combination of identification and control.

Note that, once the uncertainty set $\mathcal{S}$ and controller C are specified, the given identification criterion is well-posed. However, actual minimization may be problematic for numerical reasons. In Schrama (1992a) the problem has been simplified by considering the ℓ_2-norm in the identification criterion, whereas the H_∞-norm has been used in the control design criterion. This has been done as many standard identification tools exist which minimize an ℓ_2-norm, see e.g. Ljung (1987) for time-domain methods, and

Pintelon *et al.* (1993) and Bayard (1994) for frequency-domain methods. However, it is essential for the triangle inequality to be applicable that the norms used in identification and control are identical. And robust control design methods generally utilize an H_∞- or ℓ_1-norm.

In this thesis identification algorithms are developed which do utilize the H_∞- or ℓ_1-norm. However, attention is restricted to the criterion function (2.5). No solution is provided for the more general problem (2.9) for the reason that this is a rather complicated problem. As mentioned before the problem (2.5) is referred to as a *system approximation* problem. It is not an identification problem in the classical sense as no measurement data are involved in its formulation. Obviously, the system uncertainty set $\mathcal{S}$ is identified on the basis of data and, consequently, the nominal model $\widehat{G}(z)$ is based on data in this indirect fashion. If in the criterion (2.5) the model error is measured by its H_∞-norm (ℓ_1-norm), this concerns the problem of *system approximation in H_∞ (ℓ_1)*. If a model $\widehat{G}(z)$ minimizes the H_∞-norm (ℓ_1-norm) of the identification criterion (2.5), the nominal model is called *H_∞-optimal (ℓ_1-optimal)*.

It appears that $\mathcal{S}$ in the format given by (2.3) is useful if the H_∞-norm is considered in the criterion (2.5). In Chapter 7 an (approximate) solution is provided for this H_∞-optimal system approximation problem. In the solution fruitfully use is made of a frequency response curve fit procedure which is developed in Chapter 6. The set $\mathcal{S}$ as given by (2.4) appears useful if the ℓ_1-norm is used in (2.5). In Chapter 7 also a solution is provided for this ℓ_1-optimal system approximation problem. In the solution fruitfully use is made of an ℓ_1-optimal model reduction procedure which is developed in Chapter 7 as well.

This subsection is ended with a discussion of the problem of *identification in H_∞* as formulated by Helmicki *et al.* (1991) and studied extensively by many authors (Partington, 1991; Mäkilä, 1991a, 1991b; Mäkilä and Partington, 1992; Helmicki *et al.*, 1992, 1993; Gu and Khargonekar, 1992a, 1992b; Gu *et al.*, 1992; Akçay *et al.*, 1993; Bai and Raman, 1993). In this identification approach a system condidence set $\mathcal{S}$ as given in (2.3) is utilized as starting point for the construction of a nominal model with model error bound. Many identification algorithms have been developed for this identification setting. Due to the explicit H_∞-error bound, these identification methods are claimed to be "*inherently control-oriented*". Nevertheless, although the papers cited certainly contain valuable theoretical results, to the author's opinion the algorithms developed are generally not suited to identify models which can fruitfully be used in a control design procedure. These methods easily lead to bad nominal models in combination with an overly large model error bound, as also shown in Ninness (1993).

The inferiority of the nominal models is explained by the fact that the identification procedures are evaluated only for their asymptotic behaviour. The notion of *(robust) convergence* is considered the most important property for identification in H_∞. This means that the identification procedure is evaluated for the situation that

- the noise level tends to zero,
- the number of measurements tends to infinity,
- the model order tends to infinity.

Unfortunately, no practical situation exists where any of these is satisfied. In practice there is nonzero noise, a finite number of measurements, and a finite, or even small, model order. However, hardly any analysis (let alone feedback-relevant analysis) is made of the quality of the nominal model for the non-asymptotic situation, otherwise than by means of a conservative upper bound on the model error.

The conservativeness of the upper bound on the model error is partly due to the fact that the noise is assumed to behave worst-case deterministic, which is generally a pessimistic noise characterization as discussed before. But the conservativeness of the model error bound is also due to the fact that it is computed *a priori*, solely on the basis of the prior information assumed available about system and noise. This yields an upper bound on the H_∞-norm of the model error for the worst-case system and noise realization. This a priori worst-case H_∞-bound is needed to prove the convergence properties of the nominal identification algorithm for any system and any noise realization, but it is doubtful if it is useful as model error bound in robust control design. In general neither the system nor the noise realization is worst-case. It should, obviously, be possible to compute much smaller error bounds *a posteriori*, i.e. using data, which contain information about the specific system and noise realization.

2.4.6 Uncertainty Structure and Model Error Bound

Solving the system approximation problem (2.5) yields a nominal model $\widehat{G}(z)$. The next step in the identification procedure depicted in Figure 2.4 is the selection of an uncertainty structure and subsequently the calculation of an upper bound on the model error. Note that this can be carried out for any nominal model. It is not compulsory to use an identification procedure which minimizes (2.5). Basically any identification algorithm may be used, which is considered relevant for the problem at hand, such as the control-oriented identification procedures described in Rivera *et al.* (1992), Zang *et al.* (1992), Schrama (1992a), Lee *et al.* (1993), etc. As shown later, on the basis of the identified system uncertainty set $\mathcal{S}$ a model error bound can be established for any nominal model $\widehat{G}(z)$, irrespective of the applied nominal identification procedure.

Typically in this step of the identification procedure conservatism is added as an upper bound on the model error is needed in a format simple enough for use in the control design stage. In particular this restricts the order of the weighting function in the uncertainty representation. If much conservatism is added, it is necessary to do this such that it has the smallest impact on the control objective. In other words, if possible only little conservatism should be added in the closed-loop-relevant frequency range, such as near the cross-over frequency. A large upper bound on the model error for those frequencies would unnecessarily restrict the achievable robust performance, and as such the achievable controller performance for the system. More conservatism can be added for other frequency ranges, such as the low- and high-frequency ranges, but always such that the achievable robust control performance is not affected too much either. Altogether this means that the uncertainty structure should be chosen feedback-relevant, see also Schrama (1992a, Ch. 6).

In this thesis attention is restricted to the following weighted additive uncertainty structure, however without claiming that this is the only uncertainty structure of

interest,

$$G_0(z) = \widehat{G}(z) + \Delta(z)W^{-1}(z), \tag{2.10}$$

with $W(z)$ an arbitrary finite-dimensional, LTI, stable and stably invertible, weighting.

For the given uncertainty structure the problem of specifying an uncertainty description which is (optimally) tuned towards the control objective, boils down to the problem of (optimally) specifying a weight $W(z)$. Unfortunately, in the ℓ_1-case the weighting $W(z)$ must be specified a priori, *before* the set $\mathcal{S}$ given by (2.4) can actually be identified. In the H_∞-case the weighting $W(z)$ is not incorporated in the definition of the set $\mathcal{S}$, see (2.3). Therefore in the H_∞-case this weighting can be specified a posteriori, *after* the system uncertainty set $\mathcal{S}$ and the nominal model $\widehat{G}(z)$ have been identified. For that reason in the sequel of this subsection the discussion about optimal uncertainty weight selection is restricted to the H_∞-case.

Consider the frequency response uncertainty set $\mathcal{S}$ specified in (2.3). For any given nominal model $\widehat{G}(z)$ the following bounds $\delta(\omega_j)$ can easily be established,

$$\delta(\omega_j) = \max_{G(\omega_j)\in\mathcal{P}(\omega_j)} \left| G(\omega_j) - \widehat{G}(e^{i\omega_j}) \right|,$$

which is a frequency dependent upper bound on the additive model error. Now this non-parametric upper bound on the model error, $\delta(\omega_j)$, should be translated into a parametric upper bound, $|W^{-1}(e^{i\omega_j})|\delta$, with $\|\Delta(z)\|_\infty \leq \delta$. This concerns the specification of a bound δ, but in particular it concerns the choice of a suitable weighting function $W(z)$. For each frequency ω_j the parametric upper bound must be larger than the non-parametric upper bound $\delta(\omega_j)$. In general only for a few frequencies it can be achieved that the parametric upper bound equals the non-parametric upper bound, such that there is no conservatism for these frequencies.

In Scheid *et al.* (1991) a procedure is presented to construct a stable, minimum-phase weighting $W(z)$ of some specified order such that $\left|W^{-1}(e^{i\omega_j})\right|$ tightly overbounds the frequency dependent model error bound. The procedure minimizes the following criterion,

$$\min_{W(z)} \max_{\omega_j\in\Omega} \left| \left(\left|W^{-1}(e^{i\omega_j})\right|^2 - \delta^2(\omega_j) \right) \bar{W}^2(\omega_j) \right|,$$

or equivalently

$$\min_{W(z)} \max_{\omega_j\in\Omega} \left| W^{-1}(e^{i\omega_j})\bar{W}(\omega_j) \right|,$$

subject to the constraint that $\left|W^{-1}(e^{i\omega_j})\right|$ overbounds $\delta(\omega_j)$,

$$\left|W^{-1}(e^{i\omega_j})\right| \geq \delta(\omega_j),\ j = 1, \ldots, l.$$

Here $\bar{W}(\omega_j)$ is a weighting function which can be used to emphasize the need of a good fit at some frequencies, such as near the cross-over frequency. Obviously, the use of the resulting optimal weighting $W(z)$ in combination with the norm-bound $\|\Delta(z)\|_\infty \leq 1$ is a correct uncertainty description, at least for the frequencies in the set Ω. If for the same choice of weight $W(z)$ the intersample frequency response behaviour is taken into account as well, the norm-bound typically increases.

It appears that the choice of uncertainty structure is closely related to the choice of nominal identification criterion. In fact a *global* identification objective can be formulated for simultaneous identification of the nominal model $\widehat{G}(z)$ and the uncertainty weighting $W(z)$ such that the upper bound on the model error, $\left|W^{-1}(e^{i\omega})\right|$, is minimized in some weighted sense,

$$\min_{\widehat{G},W} \max_{\omega\in[0,\pi]} \left|W^{-1}(e^{i\omega})W^*(e^{i\omega})\right| \text{ s.t.}$$
$$\left|W^{-1}(e^{i\omega})\right| \geq \max_{G\in\mathcal{S}} \left|G(e^{i\omega}) - \widehat{G}(e^{i\omega})\right|, \ \forall\omega\in[0,\pi].$$

The constraint in this optimization problem represents the fact that the upper bound on the model error should be larger than the model error itself. The weighting $W^*(z)$ is assumed to represent the global identification objective. By this is meant that, if a small model error bound is desired for a certain frequency range, $\left|W^*(e^{i\omega})\right|$ should be chosen large for these frequencies. The optimal nominal model $\widehat{G}(z)$ will be such that the worst-case model error is relatively small in this frequency range. And the optimal uncertainty weighting $W(z)$ will be such that the model error bound tightly overbounds the worst-case model error in this frequency range.

This global identification problem combines the problem of identifying a good nominal model and the problem of specifying a suitable weighting in the uncertainty structure. One can try to solve this global problem as a combined optimization problem over $\widehat{G}(z)$ and $W(z)$. However, this is a difficult optimization problem. In the approach in this thesis two subproblems are considered instead, which provide an approximate solution for the global identification problem. The first subproblem is with respect to the nominal model, and the second one with respect to the weighting in the uncertainty structure. First, the nominal model $\widehat{G}(z)$ is determined such that the worst-case model error is minized,

$$\min_{\widehat{G}\in\mathcal{M}} \max_{G\in\mathcal{S}} \max_{\omega\in[0,\pi]} \left|\left(G(e^{i\omega}) - \widehat{G}(e^{i\omega})\right) W^*(e^{i\omega})\right|.$$

This first problem corresponds to the nominal system approximation problem (2.5), for the choice $\widetilde{W}(z) = W^*(z)$. Next, the uncertainty structure weighting $W(z)$ is determined such that the upper bound on the worst-case model error, $\left|W^{-1}(e^{i\omega})\right|$, is minimized,

$$\min_{W(z)} \max_{\omega\in[0,\pi]} \left|W^{-1}(e^{i\omega})W^*(e^{i\omega})\right| \text{ s.t.}$$
$$\left|W^{-1}(e^{i\omega})\right| \geq \max_{G\in\mathcal{S}} \left|G(e^{i\omega}) - \widehat{G}(e^{i\omega})\right|, \ \forall\omega\in[0,\pi],$$

with the nominal model $\widehat{G}(z)$ as specified by the first step. This second problem corresponds to the problem considered in Scheid *et al.* (1991) as discussed above, for the choice $\bar{W}(\omega_j) = W^*(e^{i\omega_j})$. The close connection between the nominal identification problem (2.5) and the problem of uncertainty structure selection, becomes clear from the fact that the weightings $\widetilde{W}$ and $\bar{W}$ in fact should be chosen identical.

Remark 2.4.3 *As mentioned before it is rather difficult, if not impossible, to formulate an analogous optimization problem for constructing a suitable weight $W(z)$ if the ℓ_1-norm is considered. This is due to the fact that the pulse response uncertainty set $\mathcal{S}$ given by (2.4) is constructed for an a priori specified weighting $W(z)$. This weight can not be determined a posteriori like in the H_∞-case discussed above. This is in turn due to the fact that in the frequency domain multiplications of transfer functions can be performed pointwise, for each frequency separately. However, in the time domain a multiplication of two transfer functions does not correspond to multiplication of the corresponding pulse response sequences, but to a convolution of these.*

The problem of optimally specifying the weight $W(z)$ is not further discussed in this thesis.

Once the weighting $W(z)$ is given, a bound δ on the H_∞- or ℓ_1-norm of $\Delta(z)$ can be calculated,

$$\delta = \max_{G(z)\in\mathcal{S}} \left\| \left(G(z) - \widehat{G}(z)\right) W(z) \right\| .$$

In case of the H_∞-norm the system uncertainty set $\mathcal{S}$ as given in (2.3) can be used to calculate δ, properly taking into account the worst-case intersample frequency response behaviour. In case of the ℓ_1-norm the set $\mathcal{S}$ as given in (2.4) can be used. Details are given in Chapter 4, in the Subsections 4.4.3 and 4.5.3.

Note that, if the weight $\widetilde{W}(z)$ in the system approximation criterion (2.5) is chosen identical to the weight $W(z)$ in the uncertainty description (2.10), the nominal model $\widehat{G}(z)$ is determined such that the upper bound on the model error δ is minimized.

The model error description (2.10) in combination with the norm-bound $\|\Delta(z)\| \leq \delta$, leads to the uncertainty set

$$\mathcal{G} := \left\{ G(z) \;\middle|\; G(z) = \widehat{G}(z) + \Delta(z)W^{-1}(z),\ \|\Delta(z)\| \leq \delta \right\},$$

which by construction has the outer-bounding property

$$\mathcal{S} \subseteq \mathcal{G}.$$

This uncertainty set $\mathcal{G}$ can be used in robust control synthesis and analysis procedures, such as μ- or ℓ_1-control design procedures.

Note that it is not always necessary to use a model error description in some specified uncertainty structure, like the one in (2.10). For example, the control design procedure described in Horowitz (1979, 1982), referred to as Quantitative Feedback Theory, directly uses frequency responses to design a controller, without even needing a nominal model. Moreover, in robustness analysis it appears very well possible to directly apply the system's frequency response uncertainty regions $\mathcal{S}$ as given by (2.3) in order to prove robust stability (or robust performance). In the SISO case use can be made of the Nyquist stability criterion. An example of this is given in Chapter 8. In the MIMO case μ can be evaluated for each frequency separately. This is carried out in the MIMO application in Chapter 9.

2.4.7 Model Validation

The final step in the identification procedure, and in many other identification procedures, is model validation, a verification if the identified model describes the system satisfactorily. This is always done on the basis of a fresh data set, i.e. measurement data which have not been used in the identification before. Standard model validation is validation of the nominal model (Ljung, 1987). One can, for example, compute the simulation or prediction error of the nominal model for a fresh data set, and verify whether it is sufficiently uncorrelated to the input signal. Notice that it is important that the validation is performed in view of the intended model application. It is, for example, well-known that a nominal model suited for high-performance control design, does not necessarily provide a good description of the uncontrolled system. And conversely a good fit on an open loop system response need not imply that the nominal model is suited for high-performance control design, see Dailey and Lukich (1988), Jacobson *et al.* (1991), Schrama (1992a, Ch. 2) and Hakvoort *et al.* (1994).

In fact the identification of the model error bound is a kind of validation as well. The model error bound correctly quantifies the accuracy of the nominal model, provided the prior information utilized is correct. This implies that validation of a nominal model with its model error bound is equivalent to validation of the prior assumptions. Fresh data can be used then to check if the prior assumptions about system and noise are likely to be correct. In Poolla *et al.* (1994) and Smith and Doyle (1992) this is performed by investigating if there exists a system $G(z)$ in the uncertainty set $\mathcal{G}$, in combination with a noise realization $e(t)$ satisfying some noise bounds, such that the validation data $\{u(t)\}$ and $\{y(t)\}$ are explained, in the sense that $y(t) = G(q)u(t) + e(t)$, $\forall t$.

In the framework of the identification procedure adopted in this thesis yet another approach to the validation of the nominal model with its model error bound makes sense. For the new data set one may compute the corresponding system uncertainty set $\mathcal{S}$, and check if this uncertainty set intersects with the one derived for the identification data set. If the uncertainty sets (with respect to a high probability α) do not intersect, the prior assumptions are invalidated. Either the noise assumptions, such as noise bounds, or the system assumptions, such as linearity, are invalid.

2.5 Outline of the Thesis

In Part I of this thesis attention is paid to the problem of identifying system uncertainty bounds and model error bounds. Algorithms are developed for both deterministic and stochastic noise assumptions. In Part II attention is paid to problem (2.5), the problem of identifying an H_∞- or ℓ_1-optimal model. Algorithms are developed which (partially) solve this identification problem. In Part III the elements of the first two parts are combined into a practical identification procedure, which is applied to an industrial process. Now a more detailed description is given of the contents of each chapter.

Chapter 3: Consistent Parameter Bounding Identification

Part I starts with this chapter, in which alternatives are introduced for the standard time-domain noise constraints in bounded error identification. The alternatives include cross-covariance and frequency domain noise constraints. In contrast with the time-domain noise constraints these possess certain averaging properties, that particularly become apparent in the formulation of consistency results. The new types of noise bounds are of use in the procedure for deterministic uncertainty bounding identification described in Chapter 4.

The results of this chapter are also of independent interest, as they can be used in parameter bounding identification, providing it consistency properties which standard parameter bounding identification with time-domain noise constraints lacks. For that reason in this chapter a parameter bounding identification setting is adopted. The main results of this chapter are also reported in Hakvoort *et al.* (1993).

Chapter 4: A Procedure for Deterministic Uncertainty Bounding Identification

In this chapter a procedure for deterministic uncertainty bounding identification is developed. The noise constraints described in Chapter 3 can be used in this procedure. The use of orthonormal basis functions allows for the incorporation of approximate knowledge of the pole locations of the system. Both frequency and pulse response uncertainty regions can be identified. It is also shown how these uncertainty regions can be used to derive an upper bound on the H_∞- or ℓ_1-norm of the model error. Attention is paid to the problem of estimating the required prior information from data. The MIMO situation is considered as well. Some results of this chapter are also reported in Hakvoort and Van den Hof (1993) and Hakvoort (1992b, 1994).

Chapter 5: A Procedure for Probabilistic Uncertainty Bounding Identification

In this chapter a procedure for probabilistic uncertainty bounding identification is developed. The procedure is closely related to the uncertainty bounding procedure of Chapter 4, as the same system priors are used. In a similar fashion orthonormal basis functions can be used to incorporate approximate system knowledge. Both frequency and pulse response confidence regions can be identified. The MIMO situation is considered as well. Some results of this chapter are also reported in Hakvoort and Van den Hof (1994b).

Chapter 6: A Frequency Response Curve Fit Procedure

Part II starts with this chapter in which a frequency response curve fit procedure is developed. In this curve fit procedure a parametrized transfer function is fitted to a set of frequency response data, minimizing a maximum amplitude criterion. In the curve fit procedure a parametrization of all stable models of some specified order is used. Consequently, the resulting curve fit model is always stable. MIMO extensions are given as well. The curve fit algorithm can fruitfully be applied to solve the problem of identification in H_∞. However, the results of this chapter are considered to be of

independent interest as well. Possible applications of the curve fit procedure are model reduction in H_∞-norm, or H_∞-optimal transformations from discrete to continuous time and vice versa. The results of this chapter can also be found in Hakvoort (1993a) and Hakvoort and Van den Hof (1994a).

Chapter 7: System Approximation in H_∞ and ℓ_1

In this chapter the problems are considered of identifying an H_∞ or ℓ_1-optimal nominal model, as defined in (2.5). Algorithms are developed which (approximately) solve these problems. The curve fit procedure developed in Chapter 6 appears indispensable for a solution in the H_∞-case. In the ℓ_1-case an essential step in the identification procedure is ℓ_1-optimal model reduction. As no tools are available so far which solve that, an algorithm is developed which (partially) solves this model reduction problem. The identification procedures are analysed and extensions to the MIMO case are presented. Some results as reported in this chapter can be found in Hakvoort (1992b, 1993b).

Chapter 8: A Robust-Control-Oriented Identification Procedure

This is the first chapter of Part III, which describes the practical application of the theory presented in Parts I and II. In this chapter a practically applicable, robust-control-oriented identification procedure is proposed. The procedure builds on the identification procedure depicted in Figure 2.4. In fact the results of Part I are used to specify the items "prior information" and "system uncertainty set" in this identification procedure. And the results of Part II are used to specify "model set and identification criterion" and "nominal model" in this procedure. The procedure is evaluated in the light of the requirements which an identification procedure should meet to be applicable in a process industry environment.

Chapter 9: Application to a Glass Tube Manufacturing Process

The identification procedure proposed in Chapter 8 is applied to a multivariable industrial process. The identified nominal model is used to design a model-based controller. Before implementation on the process robust stability of the controller is evaluated using the identified model error bounds.

Chapter 10: Conclusions and Perspectives

In this final chapter the results are discussed, conclusions are drawn and recommendations are given for future research.

Part I

Identification of System Uncertainty Bounds

Chapter 3

Consistent Parameter Bounding Identification

3.1 Introduction

This chapter describes alternatives for the standard time-domain noise constraints in bounded error identification. These alternatives are represented by linear constraints, to be used in parameter bounding identification by means of linear programming. The results of this chapter are applied in the procedure for deterministic uncertainty bounding identification described in the next chapter. However, the results are considered of independent interest as well. Therefore the chapter is written such that it can be read independently of the other chapters in this thesis. A parameter bounding identification setting is adopted with a linearly parametrized model set. A FIR parametrization with generalized basis functions, as used in the next chapter, is a special case of this parametrization.

The literature on set membership, bounded error or parameter bounding identification is quite extensive by now. See Walter and Piet-Lahanier (1990), Milanese and Vicino (1991) and Norton and Veres (1991) for overviews on this topic. The idea is to calculate a parameter set of minimal size using measurement data and certain deterministic bounds on the noise. To clarify the discussion consider the discrete-time linear regression model

$$y(t) = \phi^T(t)\theta + e(t),\ t = 1, \ldots, N, \tag{3.1}$$

where $y(t)$ is the measured output at time t, $\phi(t)$ the $n \times 1$ regression vector, θ the $n \times 1$ parameter vector, $e(t)$ the equation error or residual and N the number of samples. This linear regression model can describe a large class of systems, including multi-input single-output and nonlinear systems. This means that in this chapter attention is not restricted to the LTI identification setting adopted in Chapter 2. The observations are assumed to be generated by

$$y(t) = \phi^T(t)\theta_0 + e_0(t),\ t = 1, \ldots, N, \tag{3.2}$$

with $\{e_0(t)\}$ an unknown stochastic noise process. It is emphasized that $\{e_0(t)\}$ is not assumed to be white noise or uncorrelated to the regression vector $\{\phi(t)\}$. Thus, the data may be generated in closed loop and the regression vector may contain samples of the output signal $\{y(t)\}$. Basically $\{e_0(t)\}$ may also account for undermodelling, but in this chapter undermodelling is not considered.

In parameter bounding identification the parameter vector θ is bounded on the basis of certain bounds on $\{e(t)\}$. The most common procedure is to bound the amplitude of the residuals in the time domain,

$$e_l(t) \leq e(t) \leq e_u(t), \ t = 1, \ldots, N, \tag{3.3}$$

see e.g. Walter and Piet-Lahanier (1990). The feasible parameter set is then defined as

$$\Theta_N = \left\{ \theta \mid e_l(t) \leq y(t) - \phi^T(t)\theta \leq e_u(t), \ t = 1, \ldots, N \right\}.$$

Next, an orthotopic outer bounding parameter set can be constructed by calculating

$$\theta_k^{(l)} = \min_{\theta \in \Theta_N} \theta_k, \ \theta_k^{(u)} = \max_{\theta \in \Theta_N} \theta_k, \tag{3.4}$$

for $k = 1, \ldots, n$, which requires solving $2n$ linear programming problems with n unknowns subject to $2N$ linear inequality constraints, see Milanese and Belforte (1982) for details. If $\{e_0(t)\}$ also satisfies the residual bounds (3.3), then the parameter vector θ_0 is guaranteed to be in the identified set.

Note that the feasible parameter set Θ_N may be unbounded. This can happen if the data are insufficiently informative with respect to the chosen model parametrization. Basically this would imply that one or more parameter bounds as identified in (3.4) are infinite. If this is undesired, prior parameter bounds can be included in the definition of Θ_N.

Fogel and Huang (1982) and Veres and Norton (1991) have shown that under certain conditions the outer bounding parameter set converges to θ_0 if $N \to \infty$, provided the noise $\{e_0(t)\}$ is at sufficiently many time instants arbitrarily close to the specified noise bounds, without exceeding them. However, it seems impossible to meet this requirement in many practical situations. In practice the noise bounds have to be chosen conservative in order to guarantee their correctness. Therefore in general there is no consistency in parameter bounding identification. A similar situation is encountered in the field of identification in H_∞, see e.g. Helmicki *et al.* (1991) and Gu and Khargonekar (1992a, 1992b). There deterministic bounds on the noise are assumed, and basically only consistency (convergence) is established under the condition that the noise level tends to zero. This highlights the demerit of only using the bounds (3.3). The explanation for this lack of consistency in the presence of noise is that the noise is assumed to be able to take a worst-case realization within the noise bounds. In particular the residuals $\{e(t)\}$ can be heavily correlated with the input signal.

For prediction error type of identification procedures it is known that if the number of data samples tends to infinity the parameter estimate converges to the true parameter vector θ_0 under fairly general conditions, also for nonzero noise, see e.g. Ljung (1987). This is due to the fact that stochastic or averaging properties are present in

this identification setting. The objective of this chapter is to adjust the parameter bounding identification such that a similar consistency property is obtained in the presence of noise. This is achieved by introducing alternative noise bounds, which have a stochastic interpretation.

Sometimes it is claimed that for small data sets, N small, stochastic assumptions on the noise may not be justifiable. The noise bounds (3.3) would then be a much more realistic description of the noise. However, Monte Carlo simulations in Chapter 5 show that this certainly need not be true. A stochastic noise description can still be very realistic in case of a small data set. Moreover, if a large data set is available, the noise characterization (3.3) is definitely overly pessimistic compared to a stochastic description.

Stochastics have already been introduced into parameter bounding identification by Fogel and Huang (1982), where basically a stochastic interpretation is given for the noise bounds (3.3). In Veres and Norton (1989) the situation is considered that the noise has a bounded auto-covariance, or cross-covariance with a specified signal, but only for the purpose of model structure selection. Fogel (1979) considers bounds on the ℓ_2-norm or energy of the noise, but this neither leads to consistency. In Gustafsson and Mäkilä (1993) a bound on the ℓ_1-norm of the noise is proposed. But the use of this noise bound leads to huge linear programming problems and a consistency property lacks as well.

In this chapter a cross-covariance bound on the noise is introduced into parameter bounding identification as an alternative for the standard time-domain bound on the noise (3.3). The sample covariance between the residuals and some given signal is bounded. Typically this signal is chosen such that it is correlated to the input signal, but uncorrelated to the noise process. In an open loop experimental situation the (filtered) input signal meets the specifications. In closed loop operation some external reference signal can be taken. The cross-covariance noise bounds can be represented by a small number of linear inequalities, which can be used to calculate an outer bounding parameter set by means of linear programming, similar to (3.4). Sufficient conditions are derived which make the feasible parameter region converge to the parameter vector θ_0. These conditions are very general, basically only persistence of excitation is required, and it is not required that the specified bounds are tight. A stochastic interpretation is given of the cross-covariance bounds on the noise. Also, a procedure is presented to estimate correct bounds from measurement data.

Similar consistency results appear obtainable with other types of noise bounds, which also give rise to linear constraints usable in parameter bounding identification with linear programming techniques. A frequency-domain bound on the noise is introduced into parameter bounding identification. More specifically the amplitude of the discrete Fourier transform (DFT) of the residuals, the square-root of its periodogram, is bounded for a set of specified frequencies. This bound is also utilized in frequency-domain identification procedures, see Lamaire *et al.* (1991) and De Vries and Van den Hof (1992, 1994a). As parameter bounding identification adopts a time-domain setting, the bound requires a translation into linear constraints in the time domain, which is presented in this chapter. A stochastic analysis of the noise bound is presented, as well as sufficient conditions for consistency.

Finally, similar results are shown to be obtainable for a time-domain bound on the noise in combination with measurement averaging and periodic excitation. The use of repeated experiments has also been suggested by Mäkilä (1991b) in an ℓ_1-identification setting. In De Vries and Van den Hof (1992, 1994a, 1994b) repeated experiments have succesfully been applied in an H_∞-identification setting.

The outline of the chapter is as follows. In Section 3.2 the cross-covariance constraint on the noise is elaborated and consistency results are derived. In Section 3.3 it is discussed how to estimate cross-covariance noise bounds from data in a statistically reliable way. In Section 3.4 the frequency-domain constraint on the noise is considered. Next, in Section 3.5 the time-domain noise constraints are combined with repeating experiments. In Section 3.6 a simulation example is shown. Finally, in Section 3.7 the results are discussed.

3.2 Cross-Covariance Constraints on the Noise

First, the cross-covariance constraints on the noise are introduced into parameter bounding identification. Consider the following linear constraints,

$$c_l(p) \leq \frac{1}{\sqrt{N}} \sum_{t=1}^{N} r_p(t)e(t) \leq c_u(p), \ p = 1, \ldots, s, \tag{3.5}$$

yielding the feasible parameter set,

$$\Theta_N = \left\{ \theta \;\middle|\; c_l(p) \leq \frac{1}{\sqrt{N}} \sum_{t=1}^{N} r_p(t)(y(t) - \phi^T(t)\theta) \leq c_u(p), \ p = 1, \ldots, s \right\}, \tag{3.6}$$

where $c_l(p)$ and $c_u(p)$ are specified bounds and $\{r_p(t)\}$ is some specified signal, typically equal to a delayed and/or filtered signal that is correlated to the regression vector $\{\phi(t)\}$ but uncorrelated to the noise process $\{e_0(t)\}$, as explained later. The constraints (3.5) restrict the set of accepted residuals, and therefore the feasible parameter set. As the constraints in (3.6) are linear in the parameter vector θ they can easily be included in the linear programming problems (3.4).

Of course it is desirable to specify the bounds $c_l(p)$ and $c_u(p)$ on the sample covariance of the residuals with the signals $\{r_p(t)\}$ such that they are satisfied by the true noise process $\{e_0(t)\}$. Analogously to the noise bounds (3.3) it is possible to give a deterministic interpretation for the bounds (3.5), which then does not require the noise being looked upon as a stochastic process. However, it appears that a nice probabilistic interpretation of the bounds (3.5) exists if the noise process $\{e_0(t)\}$ has some stochastic properties, which can often be justified in practical situations. In fact (3.5) then boils down to the assumption that the noise process $\{e_0(t)\}$ is uncorrelated to the signals $\{r_p(t)\}$. The underlying noise assumption is that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} r_p(t)e_0(t) = 0,$$

which is a common assumption in identification, see e.g. Ljung (1987).

For the analysis some technical assumptions on the signals $\{e_0(t)\}$, $\{r_p(t)\}$ and $\{\phi(t)\}$ are needed. The assumptions about $\{e_0(t)\}$ are:

Assumption 3.2.1 *The noise process $\{e_0(t)\}$ is stationary, with $e_0(t) = H_0(q)w_0(t)$ for some stable $H_0(q)$, and where $\{w_0(t)\}$ is a sequence of independent random variables with zero mean values, variances λ_0 and bounded fourth moments.*

The assumptions about $\{r_p(t)\}$ are:

Assumption 3.2.2 *Each signal $\{r_p(t)\}$ is quasi-stationary (Ljung, 1987, Ch. 2), i.e. its auto-covariance function*

$$R_{r_p}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Er_p(t+\tau)r_p(t)$$

exists $\forall\, \tau$. Moreover, $\{r_p(t)\}$ satisfies

$$r_p(t) = \bar{r}_p(t) + R_p(t,q)\tilde{r}_p(t),$$

where for each p $\{\bar{r}_p(t)\}$ is a bounded deterministic signal, $\{R_p(t,q),\ t = 1,2,\ldots\}$ is a uniformly stable family of filters and $\{\tilde{r}_p(t)\}$ is a sequence of independent random variables with zero mean values, variances $\lambda_{p,t}$ and bounded fourth moments.

The assumptions about $\{\phi(t)\}$ are:

Assumption 3.2.3 *Each signal $\{\phi_k(t)\}$, this is element k of vector $\phi(t)$, is quasi-stationary and it satisfies*

$$\phi_k(t) = \bar{\phi}_k(t) + S_k(t,q)\tilde{\phi}_k(t),$$

where $\{\bar{\phi}_k(t)\}$ is a bounded deterministic signal, $\{S_k(t,q),\ t = 1,2,\ldots\}$ is a uniformly stable family of filters and $\{\tilde{\phi}_k(t)\}$ is a sequence of independent random variables with zero mean values, variances $\mu_{k,t}$ and bounded fourth moments.

And the assumptions about joint properties of $\{r_p(t)\}$ and $\{\phi_k(t)\}$ are:

Assumption 3.2.4 *For each p and k the signals $\{r_p(t)\}$ and $\{\phi_k(t)\}$ are jointly quasi-stationary, i.e. they are both quasi-stationary, and the cross-covariance function*

$$R_{r_p\phi_k}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Er_p(t+\tau)\phi_k(t)$$

exists. Moreover, the signal $\begin{bmatrix} \tilde{r}_p(t) \\ \tilde{\phi}_k(t) \end{bmatrix}$ *has covariances $\nu_{p,k,t}$ and bounded fourth moments.*

Using these assumptions a stochastic interpretation of the constraints (3.5) can be established. In this interpretation the notion of uncorrelation is strengthened to independence.

Theorem 3.2.5 *Suppose that* $\{e_0(t)\}$ *and* $\{r_p(t)\}$ *are independent and that they satisfy the assumptions 3.2.1 and 3.2.2 respectively. Denote*

$$\Lambda_p^N := E\left(\frac{1}{\sqrt{N}}\sum_{t=1}^{N} r_p(t)e_0(t)\right)^2,$$

$$\Lambda_p := \lim_{N\to\infty} \Lambda_p^N,$$

and

$$R_{r_p}^N(\tau) := \frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|} Er_p(t)r_p(t+|\tau|),\ \tau = -N+1,\ldots,N-1.$$

Then

$$\textbf{(i)}\quad \Lambda_p^N = \sum_{\tau=-N+1}^{N-1} \frac{N-|\tau|}{N} R_{r_p}^N(\tau)R_{e_0}(\tau),$$

$$\textbf{(ii)}\quad \Lambda_p = \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau)R_{e_0}(\tau),$$

$$\textbf{(iii)}\quad \frac{1}{\sqrt{N}}\sum_{t=1}^{N} r_p(t)e_0(t) \overset{N\to\infty}{\longrightarrow} \mathcal{N}(0,\Lambda_p),$$

where $\mathcal{N}(0,\Lambda_p)$ *denotes the normal distribution with mean 0 and variance* Λ_p.

Proof: See Appendix 3.A. □

On the one hand, the bound (3.5) is a hard or deterministic bound, on the other hand, a probabilistic interpretation has been given in the above theorem. Note that the parts (ii) and (iii) of Theorem 3.2.5 have been stated in Ljung (1987, Pr. 16T.1) without proof.

Remark 3.2.6 *In part (iii) of Theorem 3.2.5 a convergence factor* $f_N = 1/\sqrt{N}$ *has been used. The consequence of this choice is that an asymptotic distribution could be derived, which is independent of* N. *For example, asymptotically in* N *a 0.9995 probability region is obtained by choosing* $c_u(p) = -c_l(p) = 3.5\sqrt{\Lambda_p}$. *However, other convergence factors can be of interest as well. For example, the choice* $f_N = 1/\sqrt{N\log\log N}$ *gives by the law of the iterated logarithm (Heunis, 1988) that, w.p. 1,*

$$\lim_{N\to\infty}\left|\frac{1}{\sqrt{N\log\log N}}\sum_{t=1}^{N} r_p(t)e_0(t)\right| < \sqrt{\Lambda_p}.$$

See Ljung and Wahlberg (1992) for an extensive treatment of convergence factors. Note that from a practical viewpoint there is no difference, as for finite N a certain choice of $c_l(p)$ and $c_u(p)$ in (3.5) always corresponds to a certain probability of correctness, calculable with Theorem 3.2.5. Moreover, $c_l(p)$ and $c_u(p)$ may be chosen functions of N, although this has not been made explicit in the notation.

Theorem 3.2.5 also states that Λ_p can be evaluated by considering the second order statistics of $\{r_p(t)\}$ and $\{e_0(t)\}$ separately. The second order statistics of the signals $\{r_p(t)\}$ will generally be known exactly, the statistics of the noise process $\{e_0(t)\}$ have to be estimated from data. How to estimate these statistics is the subject of Section 3.

The covariance bound (3.5) is especially useful if $\{r_p(t)\}$ is a signal that is correlated to the regression vector sequence $\{\phi(t)\}$, but uncorrelated to the noise $\{e_0(t)\}$. This follows from the following consistency result.

Theorem 3.2.7 *Suppose that the signal $\{e_0(t)\}$ satisfies the constraints (3.5) for given signals $\{r_p(t)\}$, $p = 1, \ldots, s$, and given and finite $c_l(p)$ and $c_u(p)$. Suppose that $\{r_p(t)\}$ and $\{\phi_k(t)\}$ satisfy the assumptions 3.2.2, 3.2.3 and 3.2.4. If the matrix*

$$R = \begin{bmatrix} R_{r_1\phi_1}(0) & \cdots & R_{r_1\phi_n}(0) \\ \vdots & & \vdots \\ R_{r_s\phi_1}(0) & \cdots & R_{r_s\phi_n}(0) \end{bmatrix}$$

has full column rank, then the feasible parameter region Θ_N, defined in (3.6), converges to the true parameter vector θ_0,

$$\lim_{N\to\infty} \max_{\theta\in\Theta_N} |\theta - \theta_0| = 0,$$

with probability 1.

Proof: See Appendix 3.A. □

This theorem thus provides a consistency result for bounded error identification without requiring tight error bounds. If the values $c_u(p)$ and $c_l(p)$ in (3.5) are chosen too large, convergence will still take place. The cross-covariance noise bounds possess a certain averaging property which is not present in standard parameter bounding identification with time-domain noise bounds. Also, note that this consistency result holds *without* assuming stationary noise: assumption 3.2.1 has not been used in Theorem 3.2.7.

Though the bounds may be chosen conservative, they are required to be correct for the convergence result to hold. Theorem 3.2.5 shows that the bounds in (3.5) can be chosen such that they will be correct with any prespecified probability. Moreover, in Remark 3.2.6 a probability 1 result is given. These results are applicable if the signals $\{r_p(t)\}$ are independent of the noise $\{e_0(t)\}$, and the noise is a realization of a stationary stochastic process. No assumptions on the distribution or colour of the noise $\{e_0(t)\}$ are needed.

The other conditions of Theorem 3.2.7 are not very restrictive. The matrix R will have full column rank if the identification experiment is sufficiently informative and the signals $\{r_p(t)\}$ have been chosen suitably. These signals should be chosen such that they are correlated with the regression vector $\{\phi(t)\}$. If identification takes place in open loop the (filtered and/or delayed) input signal is a suitable choice. If identification takes place in closed loop a (filtered and/or delayed) external reference signal is an appropriate choice.

Note that the assumption of quasi-stationarity of the signals $\{r_p(t)\}$ and $\{\phi_k(t)\}$ can be relaxed. Basically it is not necessary that the matrix R in Theorem 3.2.7 exists. It is merely essential that the matrix

$$\frac{1}{N}\sum_{t=1}^{N}\begin{bmatrix} r_1(t)\phi_1(t) & \cdots & r_1(t)\phi_n(t) \\ \vdots & & \vdots \\ r_s(t)\phi_1(t) & \cdots & r_s(t)\phi_n(t) \end{bmatrix}$$

is bounded and has full column rank for N large enough. But convergence of this matrix for $N \to \infty$ is not needed. Moreover, the assumption that $c_l(p)$ and $c_u(p)$ are finite can be relaxed. These bounds are allowed to tend to infinity, provided the divergence is slower than $\sqrt{N}$.

Remark 3.2.8 *Note that the use of cross-covariance bounds is closely related to the instrumental variable identification method, see Söderström and Stoica (1989, Ch. 8). In fact the signals $\{r_p(t)\}$ can be regarded as instrumental variables. Also, note the close connection of the consistency result of Theorem 3.2.7 to consistency for instrumental variable identification techniques, see Ljung (1987, Ch. 8). In the present parameter bounding identification setting as well as the instrumental variable identification setting, consistency has been shown under fairly general conditions. For example, in both cases there is still consistency if the input signal is correlated to the noise process (closed loop identification).*

3.3 Estimating the Cross-Covariance Bounds from Data

In bounded error identification an a priori specification of noise bounds is required. In case time-domain noise constraints (3.3) are used, the bounds $e_l(t)$ and $e_u(t)$ need to be specified. In case cross-covariance noise constraints (3.5) are used the bounds $c_l(p)$ and $c_u(p)$ need to be specified. It may be possible that these bounds are known a priori, e.g. from physical laws. However, it is not at all imaginary that this is not the case, and that measurement data have to be used to establish the noise bounds. Unfortunately, in the parameter bounding literature very little attention is paid to the problem of estimating the noise bounds from data.

In this section the problem is considered of estimating appropriate cross-covariance bounds from data. Theorem 3.2.5 shows that this boils down to estimating Λ_p^N, which is related to the second order noise statistics. If knowledge of the noise statistics is not available from physical contemplations about the process, measurements have to

be used to estimate these. This is a valid procedure if the noise is stationary, i.e. the statistical properties do not change in time. In that case any measurement sequence may be used to estimate the noise statistics. The estimated statistics will then remain to hold for the measurement sequence used in the parameter bounding identification procedure.

Notice that if an exact value for Λ_p^N is not obainable, an upper bound is still of use. If Λ_p^N is overestimated, the resulting noise bounds $c_l(p)$ and $c_u(p)$ are conservative but correct (with a certain specified probability) and the resulting feasible parameter region will be correct, i.e. will contain the true parameter vector. Even if the noise bounds are chosen conservative, the consistency result given in Theorem 3.2.7 remains valid.

Suppose that there is available a measurement sequence generated by (3.2). As mentioned above this need not be the same measurement sequence as the one used in the parameter bounding identification procedure, it may be an independent data set. In this section it is assumed that the regression process $\{\phi(t)\}$ is uncorrelated to the noise process $\{e_0(t)\}$. This implies that measurements have to take place in open loop and that the regression vector may only contain filtered and/or delayed samples of the input signal, as the output signal is disturbed by noise. This does not necessarily imply a restriction to FIR identification. Also, identification with Laguerre polynomials (Wahlberg, 1991) and identification with generalized orthonormal polynomials (Heuberger and Bosgra, 1990; Heuberger *et al.*, 1994) fit in this setting. Let there be available a nominal model $\widehat{\theta}$, which has been obtained independently of the given data set, but, for example, by physical modelling or identification based on another dataset. The prediction or output error $\widehat{e}(t)$ for this nominal model $\widehat{\theta}$ is given by

$$\widehat{e}(t) := y(t) - \phi^T(t)\widehat{\theta} = \psi(t) + e_0(t), \tag{3.7}$$

with

$$\psi(t) := \phi^T(t)\left(\theta_0 - \widehat{\theta}\right). \tag{3.8}$$

The idea is to use this prediction error in order to estimate the second order statistics of the noise process. First, some technical assumptions are made with respect to the signal $\psi(t)$.

Assumption 3.3.1 *The signal $\{\psi(t)\}$ is quasi-stationary and it satisfies*

$$\psi(t) = \bar{\psi}(t) + P(t,q)\tilde{\psi}(t),$$

where $\{\bar{\psi}(t)\}$ is a bounded deterministic signal, $\{P(t,q),\ t = 1,2,\ldots\}$ is a uniformly stable family of filters and $\{\tilde{\psi}(t)\}$ is a sequence of independent random variables with zero mean values, variances μ_t and bounded fourth moments. Moreover, the auto-covariance function of the signal $\{\psi(t)\}$ is exponentially decaying, i.e. $R_\psi(\tau) \leq M\rho^\tau$, $\forall\,\tau$, for certain finite M and $\rho < 1$.

A sufficient condition for assumption 3.3.1 to hold, is that each element of the regression

vector $\phi(t)$ satisfies assumption 3.3.1 and $|\theta_0 - \widehat{\theta}|$ is finite. Next, denote

$$\widehat{R}_{\widehat{e}}^N(\tau) := \frac{1}{N - |\tau|} \sum_{t=1}^{N-|\tau|} \widehat{e}(t)\widehat{e}(t + |\tau|),$$

and consider the following estimate for Λ_p^N,

$$\widehat{\Lambda}_p^N = \sum_{\tau=-w(N)}^{w(N)} c_w(\tau) \frac{N - |\tau|}{N} R_{r_p}^N(\tau) \widehat{R}_{\widehat{e}}^N(\tau), \tag{3.9}$$

where $c_w(\tau)$ is a window-function, similar to the ones used in spectral analysis, see Ljung (1987, Ch. 6). Notice that the quantities $R_{r_p}^N(\tau)$ appearing in the estimate $\widehat{\Lambda}_p^N$ are assumed to be known precisely. This is a realistic situation, as the signals $\{r_p(t)\}$ are generally user-determined. These signals may be deterministic and completely known, or realizations of a stochastic process, with known second order statistics, or mixed deterministic-stochastic, with given auto-covariance function.

In order to be able to establish a useful convergence results, some technical assumptions are made with respect to the window used in the estimate (3.9). These assumptions, which are taken from Hjalmarsson (1993a, Th. 3.1), in fact state that the window should converge to 1 if $N \to \infty$, but slow enough in comparison to the number of data N.

Assumption 3.3.2 *The sequence of integers $\{w(N)\}$ is a positive, monotonously increasing sequence such that for some $k > \frac{5}{2}$,*

$$\lim_{N\to\infty} \frac{w(N)}{\sqrt{N}/\log^k N} = C,$$

for some finite C. And the real-valued window-function $c_w(\tau)$ is such that

$$|c_w(\tau)| < C, \ \forall w, \tau,$$

and

$$\lim_{w\to\infty} c_w(\tau) = 1, \ \forall \tau.$$

The following theorem states that, under the given assumptions, the estimate $\widehat{\Lambda}_p^N$ asymptotically overbounds Λ_p^N.

Theorem 3.3.3 *Consider Λ_p^N as given in part (i) of Theorem 3.2.5, and the estimate $\widehat{\Lambda}_p^N$ defined in (3.9) with a window which satisfies assumption 3.3.2. Suppose that $\{e_0(t)\}$ satisfies assumption 3.2.1 and $\{r_p(t)\}$ assumption 3.2.2. Moreover, suppose that the estimate $\widehat{\theta}$, used in (3.7), has been established independently of the noise process $\{e_0(t)\}$, and that the signal $\{\phi(t)\}$ is uncorrelated to $\{e_0(t)\}$, and additionally*

satisfies certain weak conditions (in particular conditions (2.11), (3.13) and (3.14) in Hjalmarsson, 1993a). Finally, suppose that $\{\psi(t)\}$ satisfies assumption 3.3.1. Then,

$$\textbf{(i)} \quad \lim_{N\to\infty} \widehat{\Lambda}_p^N = \Lambda_p + \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau) R_\psi(\tau), \ \text{w.p. } 1,$$

$$\textbf{(ii)} \quad \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau) R_\psi(\tau) \geq 0,$$

Proof: See Appendix 3.A. □

Consequently, asymptotically correct cross-covariance bounds $c_\mathrm{l}(p)$ and $c_\mathrm{u}(p)$ can be established as the estimated variance $\widehat{\Lambda}_p^N$ is overbiased. The conservatism decreases if the nominal model $\widehat{\theta}$ becomes a more accurate description of the true system θ_0. In the special case that the system has not been excited, implying $\phi(t)$ to be equal to zero, the signal $\psi(t)$ vanishes and $\widehat{e}(t)$ equals $e_0(t)$, yielding an asymptotically unbiased estimate for the noise statistics.

Remark 3.3.4 *It is desirable to introduce as little conservatism as possible when establishing the noise bounds $c_\mathrm{l}(p)$ and $c_\mathrm{u}(p)$, in order to avoid that unnecessarily large parameter sets are identified in the parameter bounding identification procedure. Part (ii) of Theorem 3.3.3 shows that asymptotically the estimate $\widehat{\Lambda}_p^N$ overbounds Λ_p^N. As mentioned only in the special case that $\widehat{e}(t)$ equals $e_0(t)$ the estimate is unbiased, yielding minimal conservatism. However, there is another special situation where it is possible to derive a non-conservative estimate of the noise statistics, with a procedure different from the one described above. This is the case if a repeated experiment has been performed, i.e. if*

$$\phi(t+T) = \phi(t), \ t = 1, \ldots, T,$$

for some period time T. If the regression vector only contains (filtered) samples of the input, this is realized by applying a periodic input signal. Now consider the signal

$$\widehat{e}(t) := (y(t+T) - y(t))/\sqrt{2}, \ t = 1, \ldots, T,$$

which with (3.2) can be written as

$$\widehat{e}(t) = (\phi(t+T) - \phi(t))\theta_0/\sqrt{2} + (e_0(t+T) - e_0(t))/\sqrt{2} = (e_0(t+T) - e_0(t))/\sqrt{2}.$$

The signal $\widehat{e}(t)$ actually appears to have second order statistics identical to those of the noise process $e_0(t)$ (in the asymptotic case $T \to \infty$). Also, if the regression process is not perfectly periodic, e.g. due to different initial conditions, the given signal $\widehat{e}(t)$ can still be used to estimate the second order noise statistics of $e_0(t)$. Using an argument similar to the one used in Theorem 3.3.3 it can be shown that the estimate $\widehat{\Lambda}_p^N$ will then be overbiased, yielding correct bounds $c_\mathrm{l}(p)$ and $c_\mathrm{u}(p)$.

In Subsection 5.6.1 in Chapter 5 more is said about the properties of the estimate $\widehat{\Lambda}_p^N$.

3.4 Frequency-Domain Constraints on the Noise

It appears that consistency results in parameter bounding identification, similar to those of Theorem 3.2.7, can be obtained with another type of noise constraint. This concerns a frequency-domain bound on the noise. Consider the function $E(\omega_j)$ defined by

$$E(\omega_j) := \frac{1}{\sqrt{N}} \sum_{t=1}^{N} e(t) e^{-i\omega_j t}, \tag{3.10}$$

which for $\omega_j = 2\pi j/N$, $j = 1, \ldots, N$, is the discrete Fourier transform of the sequence $\{e(1), \ldots, e(N)\}$. With some abuse of terminology $E(\omega_j)$ is called the discrete Fourier transform of the signal $\{e(t)\}$, no matter what ω_j is, and keeping in mind that it is dependent on N. By bounding its amplitude, the square root of the so-called periodogram, the residuals are bounded in the frequency domain. Consider the constraints,

$$|E(\omega_j)| \leq f(\omega_j), \ \omega_j \in [0, \pi], \ j = 1, \ldots, l,$$

for some specified bounds $f(\omega_j)$. This type of noise constraint is also used in Lamaire *et al.* (1991) and De Vries and Van den Hof (1992, 1994a) in a frequency-domain identification setting. Substituting $e(t) = y(t) - \phi^T(t)\theta$ gives the feasible parameter set,

$$\Theta_N = \left\{ \theta \mid \left| \frac{1}{\sqrt{N}} \sum_{t=1}^{N} (y(t) - \phi^T(t)\theta) e^{-i\omega_j t} \right| \leq f(\omega_j), \ \omega_j \in [0, \pi], \ j = 1, \ldots, l, \right\}.$$

If parameter outer bounding by linear programming is carried out, linear constraints are required. However, the constraints given above are nonlinear due to the fact that the DFT is a complex-valued quantity. Fortunately it appears possible to approximate each nonlinear constraint by a number of linear constraints. Consider the following linear constraints,

$$f_{\mathrm{l}}(\omega_j, \alpha_{k'}) \leq \frac{1}{\sqrt{N}} \sum_{t=1}^{N} e(t) \cos(\omega_j t - \alpha_{k'}) \leq f_{\mathrm{u}}(\omega_j, \alpha_{k'}), \tag{3.11}$$

$$\omega_j \in [0, \pi], \ j = 1, \ldots, l, \ \alpha_{k'} \in [0, \pi], \ k' = 1, \ldots, m,$$

yielding the feasible parameter set,

$$\Theta_N = \left\{ \theta \ \middle| \ \begin{array}{c} f_{\mathrm{l}}(\omega_j, \alpha_{k'}) \leq \dfrac{1}{\sqrt{N}} \displaystyle\sum_{t=1}^{N} \cos(\omega_j t - \alpha_{k'})(y(t) - \phi^T(t)\theta) \leq f_{\mathrm{u}}(\omega_j, \alpha_{k'}) \\ \omega_j \in [0, \pi], \ j = 1, \ldots, l, \ \alpha_{k'} \in [0, \pi], \ k' = 1, \ldots, m \end{array} \right\}, \tag{3.12}$$

where $f_{\mathrm{l}}(\omega_j, \alpha_{k'})$ and $f_{\mathrm{u}}(\omega_j, \alpha_{k'})$ are specified bounds, ω_j specified frequencies and $\alpha_{k'}$ specified phase shifts. As the constraints in (3.12) are linear in θ, they can easily be included in linear programming problems like the ones in (3.4).

The relation of the constraints (3.11) to the amplitude of the DFT of $\{e(t)\}$ is investigated. This amplitude can be written as

$$|E(\omega_j)| = \left|\frac{1}{\sqrt{N}}\sum_{t=1}^{N} e(t)e^{-i\omega_j t}\right| = \max_{\alpha\in[0,2\pi]} \mathrm{Re}\left(e^{i\alpha}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} e(t)e^{-i\omega_j t}\right) = \tag{3.13}$$

$$= \max_{\alpha\in[0,2\pi]} \frac{1}{\sqrt{N}}\sum_{t=1}^{N} e(t)\mathrm{Re}\left(e^{-i\omega_j t+\alpha}\right) = \max_{\alpha\in[0,2\pi]} \frac{1}{\sqrt{N}}\sum_{t=1}^{N} e(t)\cos(\omega_j t-\alpha),$$

where $\mathrm{Re}\,(\cdot)$ denotes the real part of $\cdot$. This shows that the bound (3.11) approximates the bound on $|E(\omega_j)|$. The approximation improves if more phase shifts $\alpha_{k'}$ are used.

This can be stated more formally. For that purpose, but also for future reference, first an auxiliary lemma is presented. This lemma is an important lemma, which is also used in other parts of this thesis, in particular in the Chapters 4, 6 and 7.

Lemma 3.4.1 *Consider the function* $f_{m'}(x) : \mathbb{C} \to \mathbb{R}$, *defined by*

$$f_{m'}(x) = \max_{k'=1,\ldots,m'} \mathrm{Re}\left(e^{2\pi\frac{k'}{m'}i}x\right),$$

with $m' \geq 3$. *Then*

$$\textbf{(i)}\quad f_{m'}(x) \leq |x| \leq \frac{f_{m'}(x)}{\cos\left(\frac{\pi}{m'}\right)},$$

$$\textbf{(ii)}\quad \lim_{m'\to\infty} f_{m'}(x) = |x|.$$

Proof: See Appendix 3.A. □

Moreover, it is easy to show that the bounds in (i) are tight in the sense that there exists an x such that the lower bound becomes equality, and there is an x such that the upper bound becomes equality. The lemma in fact says that the amplitude of a complex number can be calculated approximately by checking a number of different orientations in the complex plane. Another interpretation of the upper bound in part (i) of the lemma is that a circle is approximated by an equilateral polygon. In Figure 3.1 this is illustrated for $m' = 4$ (square) and $m' = 8$ (octagon).

Returning to (3.13), the following proposition is easily established.

Proposition 3.4.2 *Let* $E(\omega_j)$ *be defined by (3.10) and* $\alpha_{k'}$ *be given by* $\alpha_{k'} = \pi k'/m$, $k' = 1,\ldots,m$, *then*

$$\max_{k'}\left|\frac{1}{\sqrt{N}}\sum_{t=1}^{N} e(t)\cos(\omega_j t-\alpha_{k'})\right| \leq |E(\omega_j)| \leq \frac{\max_{k'}\left|\frac{1}{\sqrt{N}}\sum_{t=1}^{N} e(t)\cos(\omega_j t-\alpha_{k'})\right|}{\cos(\frac{\pi}{2m})}.$$

Proof: Follows immediately from Lemma 3.4.1 by substituting $m' = 2m$. □

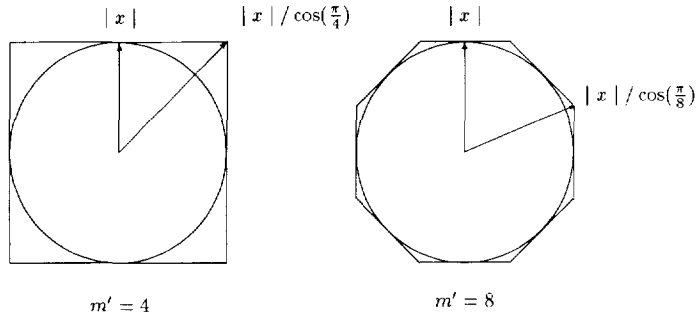

Fig. 3.1: Illustration of part (i) of Lemma 3.4.1 for $m' = 4$ and $m' = 8$: a circle is outerbounded by an equilateral polygon.

This proposition states that (3.11) bounds the amplitude of the residuals $\{e(t)\}$ in the frequency domain. This apparently boils down to excluding the residuals to contain sinusoids with the specified frequencies ω_j, $j = 1, \ldots, l$. For $\omega_j = 0$ (3.11) is a bound on the mean value of the residuals. From Proposition 3.4.2 it can also be concluded that the most natural choice for the bounds $f_l(\omega_j, \alpha_{k'})$ and $f_u(\omega_j, \alpha_{k'})$ is

$$f_u(\omega_j, \alpha_{k'}) = -f_l(\omega_j, \alpha_{k'}) = f(\omega_j), \; k' = 1, \ldots, m.$$

The frequency-domain bounds on the noise are especially useful if the system is excited by sinusoids. This follows from the following consistency result.

Theorem 3.4.3 *Suppose that the signal $\{e_0(t)\}$ satisfies the constraints (3.11) for given and finite $f_l(\omega_j, \alpha_{k'})$ and $f_u(\omega_j, \alpha_{k'})$, $j = 1, \ldots, l$, $\alpha_{k'} = \pi k'/m$, $k' = 1, \ldots, m \geq 2$. Suppose that $\{\phi_k(t)\}$ satisfies assumption 3.2.3. Denote for each j and k,*

$$\Psi_k(\omega_j) := \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E\phi_k(t) e^{-i\omega_j t}.$$

If the matrix

$$\Psi = \begin{bmatrix} \mathrm{Re}\,(\Psi_1(\omega_1)) & \cdots & \mathrm{Re}\,(\Psi_n(\omega_1)) \\ \vdots & & \vdots \\ \mathrm{Re}\,(\Psi_1(\omega_l)) & \cdots & \mathrm{Re}\,(\Psi_n(\omega_l)) \\ \mathrm{Im}\,(\Psi_1(\omega_1)) & \cdots & \mathrm{Im}\,(\Psi_n(\omega_1)) \\ \vdots & & \vdots \\ \mathrm{Im}\,(\Psi_1(\omega_l)) & \cdots & \mathrm{Im}\,(\Psi_n(\omega_l)) \end{bmatrix}$$

has full column rank, then the feasible parameter region Θ_N, defined in (3.12), converges to the true parameter vector θ_0,

$$\lim_{N \to \infty} \max_{\theta \in \Theta_N} |\theta - \theta_0| = 0,$$

with probability 1.

Proof: See Appendix 3.A. □

Again, this is a consistency result for bounded error identification without requiring tight error bounds. If the values $f_l(\omega_j, \alpha_{k'})$ and $f_u(\omega_j, \alpha_{k'})$ are chosen too large, convergence will still take place under the conditions given.

Remark 3.4.4 *In general $\Psi_k(\omega_j)$ is unequal to zero if $\{\phi_k(t)\}$ contains a sinusoid with frequency ω_j. A sinusoid $s(t) = a\sin(\omega_j t + \beta)$ namely has the well-known property that its periodogram is unbounded, as the following relation holds,*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} s(t)e^{-i\omega_j t} = \begin{cases} \frac{a}{2}(\sin\beta - i\cos\beta), & \omega_j \in (0,\pi), \\ a\sin\beta, & \omega_j = 0,\pi. \end{cases}$$

Consequently, the consistency result of Theorem 3.4.3 will only hold if the system has been excited by a sum of at least $n/2$ sinusoids. This type of excitation has also been exploited by De Vries and Van den Hof (1992, 1994a, 1994b) and Bayard (1992a).

It is desirable to specify the bounds $f_l(\omega_j, \alpha_{k'})$ and $f_u(\omega_j, \alpha_{k'})$ such that they are satisfied by the true noise process $\{e_0(t)\}$. The bounds (3.11) may be regarded as entirely deterministic, which does not require any stochastic assumptions about the noise. However, analogously to the bounds (3.5), a nice probabilistic interpretation exists in case the noise has some stochastic properties. In the following proposition probabilistic properties of the periodogram of the noise are evaluated.

Proposition 3.4.5 *Suppose that $\{e_0(t)\}$ satisfies assumption 3.2.1. Let $E_0(\omega_j)$ be the DFT of $\{e_0(t)\}$, defined analogously to (3.10), then*

$$|E_0(\omega_j)|^2 \xrightarrow{N\to\infty} \begin{cases} \frac{1}{2}\Phi_{e_0}(\omega_j)\chi^2(2), & \omega_j \in (0,\pi), \\ \Phi_{e_0}(\omega_j)\chi^2(1), & \omega_j = 0,\pi, \end{cases}$$

where $\Phi_{e_0}(\omega_j)$ denotes the auto-spectrum of the process $\{e_0(t)\}$, defined by

$$\Phi_{e_0}(\omega_j) = \sum_{\tau=-\infty}^{\infty} R_{e_0}(\tau)e^{-i\tau\omega_j},$$

and $\chi^2(n)$ denotes the chi-squared distribution with n degrees of freedom.

Proof: See Brillinger (1981, Theorem 5.2.6). The result can also be established along the lines of the proof of Theorem 3.2.5. □

Again, note that the bound (3.11) is a hard or deterministic bound, which, however, enables a probabilistic interpretation. For example, asymptotically in N a 0.999 probability region is obtained by choosing $f_u(\omega_j, \alpha_{k'}) = -f_l(\omega_j, \alpha_{k'}) = c\sqrt{\Phi_{e_0}(\omega_j)}$, $k' = 1, \ldots, m$, with $c = 2.63$ if $\omega_j \neq 0, \pi$ and $c = 3.29$ if $\omega_j = 0, \pi$.

Remark 3.4.6 *The information required to establish the frequency-domain bounds on the noise consists of $\Phi_{e_0}(\omega_j)$, see Proposition 3.4.5. Analogously to the procedure of Section 3.3 it is possible to use the prediction error to (conservatively) estimate the noise spectrum. With the definitions and assumptions of Section 3.3 it follows that*

$$\Phi_{\hat{e}}(\omega_j) = \Phi_{\psi}(\omega_j) + \Phi_{e_0}(\omega_j) \geq \Phi_{e_0}(\omega_j),$$

where the latter inequality follows from the fact that an auto-spectrum is nonnegative. The spectrum of the prediction error can be estimated with standard techniques for spectral estimation, see Brillinger (1981). Further details are omitted.

3.5 Time-Domain Constraints on the Noise Revisited

It appears that a consistency result similar to the ones presented above can be achieved with the standard time-domain noise constraints in combination with periodic excitation and measurement averaging. Consider one measurement sequence, which is split up into $m+1$ parts, each of length T.

$$y(t+k'T) = \phi^T(t+k'T)\theta + e(t+k'T),\ t = 1,\ldots,T,\ k' = 0,\ldots,m,$$

$$\frac{1}{\sqrt{m+1}}\sum_{k'=0}^{m} y(t+k'T) = \frac{1}{\sqrt{m+1}}\sum_{k'=0}^{m} \phi^T(t+k'T)\theta + \frac{1}{\sqrt{m+1}}\sum_{k'=0}^{m} e(t+k'T).$$

Now instead of bounding $e(t)$ the averaged residuals are bounded by requiring,

$$m_{\mathrm{l}}(t) \leq \frac{1}{\sqrt{m+1}}\sum_{k'=0}^{m} e(t+k'T) \leq m_{\mathrm{u}}(t),\ t = 1,\ldots,T. \tag{3.14}$$

These constraints yield the feasible parameter set

$$\Theta_N = \left\{ \theta \;\middle|\; m_{\mathrm{l}}(t) \leq \frac{1}{\sqrt{m+1}}\sum_{k'=0}^{m} y(t+k'T) - \phi^T(t+k'T)\theta \leq m_{\mathrm{u}}(t),\ t = 1,\ldots,T \right\}, \tag{3.15}$$

which can be used to set up linear programming problems for parameter bounding like the ones in (3.4).

As mentioned, the idea is to combine the averaging in combination with repeated experiments. The following consistency result shows the benefits of this approach.

Theorem 3.5.1 *Suppose that the signal $\{e_0(t)\}$ satisfies the constraints (3.14) for given and finite $m_{\mathrm{l}}(t)$ and $m_{\mathrm{u}}(t)$, $t = 1,\ldots,T$. Also, suppose that $\{\phi_k(t)\}$ satisfies assumption 3.2.3. Denote for each k,*

$$M_{\phi_k}(t) := \lim_{m\to\infty} \frac{1}{m+1}\sum_{k'=0}^{m} E\phi_k(t+k'T).$$

If the matrix

$$M = \begin{bmatrix} M_{\phi_1}(1) & \cdots & M_{\phi_n}(1) \\ \vdots & & \vdots \\ M_{\phi_1}(T) & \cdots & M_{\phi_n}(T) \end{bmatrix}$$

has full column rank, then the feasible parameter region Θ_N, defined in (3.15), converges to the true parameter vector θ_0,

$$\lim_{m\to\infty} \max_{\theta\in\Theta_N} |\theta - \theta_0| = 0,$$

with probability 1.

Proof: See Appendix 3.A. □

A necessary condition for this consistency property to hold is that there is some deterministic, repetitive component in the regression vector $\{\phi(t)\}$. If, for example, $\phi(t + k'T) = \phi(t),\ \forall\, t, k'$, then $M_{\phi_k}(t) = \phi_k(t),\ \forall\, t$.

Remark 3.5.2 *In fact by applying a periodic input signal in combination with averaging, the signal-to-noise ratio improves, whereas the number of data samples to be used in the identification procedure remains small. This data reduction is an attractive property as the complexity of the bounded error identification problem increases with the number of constraints used in the linear programming problems (3.4). Also in case m is small, the effect of the procedure described will be that the signal-to-noise ratio improves. Note that this method of improving the signal-to-noise ratio, by applying repeated experiments and data reduction, is not restricted to the set membership identification setting but can be applied in combination with any identification method.*

Again, the bound (3.14) can be considered entirely deterministic without giving a stochastic interpretation. But stochastic theory can fruitfully be applied when evaluating the meaning of these bounds. The following result is then obtained.

Theorem 3.5.3 *Suppose that $\{e_0(t)\}$ satisfies assumption 3.2.1. Then,*

$$\frac{1}{\sqrt{m+1}} \sum_{k'=0}^{m} e_0(t + k'T) \overset{m,T\to\infty}{\longrightarrow} \mathcal{N}(0, \Lambda),\ \Lambda = R_{e_0}(0),\ \forall\, t.$$

Proof: See Appendix 3.A. □

This result states that if the number of repetitions increases, the distribution of the averaged noise converges to the normal distribution.

Remark 3.5.4 *The information required to establish the time-domain bounds on the noise in combination with repeated experiments consists of $R_{e_0}(0)$. Analogously to the results of Section 3.3 the estimate $R_{\hat{e}}(0)$, which is based on the prediction error, can be used as a conservative estimate of this quantity as*

$$R_{\hat{e}}(0) = R_{\psi}(0) + R_{e_0}(0) \geq R_{e_0}(0).$$

3.6 Example

A simulation example is presented in which the performance of time-domain noise bounds in parameter bounding identification is compared to the performance of the new types of noise bounds, as introduced in this chapter.

Consider the data generating system

$$y(t) = \theta_0^{(1)}u(t) + \theta_0^{(2)}u(t-1) + \theta_0^{(3)}u(t-2) + e_0(t),\ \theta_0^{(1)} = 2,\ \theta_0^{(2)} = 1,\ \theta_0^{(3)} = 0.6,$$

with the noise process given by

$$e_0(t) = w_0(t) + 0.8w_0(t-1) + 0.2w_0(t-2) + 0.1w_0(t-3),$$

where $\{w_0(t)\}$ is a white noise process uniformly distributed between -0.25 and 0.25. The input signal $\{u(t)\}$ is chosen to be

$$u(t) = \begin{cases} 0, & t \leq 0, \\ \sin(\frac{\pi t}{4}) + \sin(\frac{\pi t}{2}), & t = 1, \ldots, N,\ N = 800. \end{cases}$$

The following parameter bounding identification procedures have been carried out.

1. Parameter bounding with bounds on the amplitude of the noise as in (3.3). As $|e_0(t)| \leq 0.25(1 + 0.8 + 0.2 + 0.1) = 0.525$, the bounds have been chosen $e_u(t) = -e_l(t) = 0.525, \forall t$. Altogether 6 linear programming problems have had to be solved, each with 3 unknowns subject to 1600 inequality constraints.

2. Parameter bounding with bounds on the cross-covariance of the noise as in (3.5). The signals $\{r_p(t)\}$ have been chosen $r_p(t) = u(t + \tau_p)$, $\tau_p = 1 - p$, $p = 1, \ldots, 8$, which with Theorem 3.2.5 yields $\Lambda_p = 0.0388$, $p = 1, \ldots, 8$. The bounds $c_u(p) = -c_l(p) = 3\sqrt{\Lambda_p}$, $p = 1, \ldots, 8$ have been chosen. Hence, each bound is satisfied with probability 0.997. Altogether 6 linear programming problems have had to be solved, each with 3 unknowns and 16 constraints. According to the rule of Bonferroni (Manoukian, 1986) the resulting parameter region is correct with a probability larger than or equal to $1 - 8(1 - 0.997) = 0.98$.

3. Parameter bounding with bounds on the periodogram of the noise as in (3.11). The frequencies $\omega_1 = \pi/4$ and $\omega_2 = \pi/2$ were selected, and $\alpha_{k'} = \pi k'/m$, $k' = 1, \ldots, 4$. Straighforward calculations show that $\Phi_{e_0}(\omega_1) = 0.2473$ and $\Phi_{e_0}(\omega_2) = 0.1534$. The bounds $f_u(\omega_j, \alpha_{k'}) = -f_l(\omega_j, \alpha_{k'}) = 2.146\sqrt{\Phi_{e_0}(\omega_j)}$, $j = 1, 2$, $k' = 1, \ldots, 4$ have been chosen, which corresponds to a 0.99 confidence interval for each frequency seperately. Thus, 6 linear programming problems with each 3 unknowns subject to 16 constraints have had to be solved. The resulting parameter region is correct with probability larger than or equal to $1 - 2(1 - 0.99) = 0.98$.

4. Parameter bounding with bounds on the averaged noise as in (3.14). The input sequence is periodic with a period of 8 samples. The last 99 periods were averaged, the first period has been used to settle the initial conditions of the system.

The bounds have been set to $m_u(t) = -m_l(t) = 3\sqrt{R_{e_0}(0)} = 0.563$, $t = 1, \ldots, 8$. This means that each linear programming problem has 3 unknowns and 16 inequality constraints. The probability that θ_0 is an element of the resulting parameter region is at least $1 - 8(1 - 0.997) = 0.98$.

The resulting upper and lower bounds on the parameters are shown in Table 3.1. It

	$\theta_1^{(l)}$	$\theta_1^{(u)}$	$\theta_2^{(l)}$	$\theta_2^{(u)}$	$\theta_3^{(l)}$	$\theta_3^{(u)}$
Procedure 1	1.890	2.117	0.862	1.127	0.468	0.707
Procedure 2	1.964	2.044	0.963	1.046	0.557	0.637
Procedure 3	1.954	2.048	0.981	1.028	0.554	0.648
Procedure 4	1.959	2.062	0.955	1.056	0.542	0.644
Exact values	2.0		1.0		0.6	

Table 3.1: Results parameter bounding identification.

is concluded that for the first and third parameter the tightest bounds are obtained with the cross-covariance bounds on the noise, for the second parameter this is the case with the frequency-domain bounds on the noise. The parameter uncertainty intervals estimated with cross-covariance noise constraints, procedure 2, are a factor 3 smaller than the ones estimated with time-domain noise constraints, procedure 1. The uncertainty interval for the second parameter estimated with the frequency-domain noise constraints, procedure 3, is even a factor 5.7 smaller than the corresponding interval for procedure 1. The new parameter bounding identification methods 2, 3 and 4 appear to outperform the standard parameter bounding identification procedure 1. Much smaller parameter uncertainty intervals are estimated with these new methods compared to the result of identification with bounds on the amplitude of the noise in the time domain.

Remark 3.6.1 *For simplicity exact knowledge of the second order noise statistics has been used to establish the bounds on the noise. The procedure of Section 3.3 may be used to estimate these statistics. In particular the method indicated in Remark 3.3.4 is applicable as the input is periodic.*

3.7 Discussion

Several valuable alternatives have been presented for time-domain bounds on the noise in parameter bounding identification by linear programming. Especially cross-covariance bounds on the noise are powerful as consistency has been proven for an arbitrary persistently exciting input signal. Frequency-domain bounds on the noise are powerful if the input signal contains sinusoids, in which case consistency has been proven as well. Also, consistency has been proven in case time-domain bounds on the

noise are combined with periodic excitation and averaging. In all these cases the number of constraints does not increase with an increasing number of measurements, such that the identification problem remains tractable for large N. The example showed a considerable reduction of parameter uncertainty in case the new types of noise constraints are utilized.

A stochastic analysis of the new types of noise bounds has been presented. The basic assumptions needed to justify this analysis are that the number of samples is large enough, and that the noise is stationary, i.e. its stochastic properties do not change in time. The analysis showed that the bounds for the noise can be specified such that they are correct with a certain probability. This means that a parameter uncertainty region is calculated which contains the true parameter vector with a certain probability. It is emphasized that the probability density function of the noise process need not be known. Only knowledge of the second order statistics of the noise process is required. A procedure has been presented to estimate these statistics from data.

In the next chapter uncertainty bounding by linear programming is applied to the problem of identifying frequency and pulse response uncertainty regions. The alternative noise bounds presented in this chapter can fruitfully be applied there. In the next chapter also the problem of undermodelling is considered.

3.A Proofs

Proof of Theorem 3.2.5

The Lemma's 9A.1 and 9A.2 of Ljung (1987) are applied to

$$S_N = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} r_p(t) e_0(t).$$

Let

$$e_0(t) = e_0^M(t) + \tilde{e}_0^M(t),\ e_0^M(t) = \sum_{k=0}^{M} h_0(k) w_0(t-k),\ \tilde{e}_0^M(t) = \sum_{k=M+1}^{\infty} h_0(k) w_0(t-k),$$

$$Z_N(M) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} r_p(t) e_0^M(t),\ X_N(M) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} r_p(t) \tilde{e}_0^M(t),$$

such that $S_N = Z_N(M) + X_N(M)$. The signal $x_N(t) = \frac{1}{\sqrt{N}} r_p(t) e_0^M(t)$ is M-dependent as $e_0^M(t)$ is M-dependent, and $r_p(t)$ and $e_0^M(t)$ are independent. Also, $Ex_N(t) = 0$ as $Ee_0^M(t) = 0$, and

$$\lim_{N\to\infty} \sum_{t=1}^{N} Ex_N^2(t) = E(e_0^M(t))^2 \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Er_p^2(t) < \infty,$$

as both $r_p(t)$ and $e_0(t)$, and consequently $e_0^M(t)$, have finite variances, $r_p(t)$ being quasi-stationary and $e_0^M(t)$ stationary. Moreover,

$$\lim_{N\to\infty}\sum_{t=1}^{N} Ex_N^4(t) = E(e_0^M(t))^4 \lim_{N\to\infty}\frac{1}{N^2}\sum_{t=1}^{N} Er_p^4(t) = 0,$$

as both $r_p(t)$ and $e_0^M(t)$ have bounded fourth moments. Consequently, the signal $x_N(t)$ satisfies the conditions of Lemma 9A.1 in Ljung (1987), and hence,

$$Z_N(M) \overset{N\to\infty}{\longrightarrow} \mathcal{N}(0, \Lambda_p(M)),$$

with

$$\Lambda_p(M) = \lim_{N\to\infty} EZ_N^2(M).$$

The signals $r_p(t)$ and $e_0(t)$ are generated in a sufficiently stable manner. With Lemma 2B.1 in Ljung (1987) it can be established that $EX_N^2(M)$ is bounded by a bound which tends to zero as M tends to infinity. Consequently, Lemma 9A.2 in Ljung (1987) applies, and hence,

$$S_N \overset{N\to\infty}{\longrightarrow} \mathcal{N}(0, \Lambda_p),$$

with

$$\Lambda_p = \lim_{M\to\infty}\Lambda_p(M) = \lim_{N\to\infty}\Lambda_p^N,\ \Lambda_p^N = ES_N^2.$$

Again using the fact that $\{r_p(t)\}$ and $\{e_0(t)\}$ are independent, and using the substitution $\tau = t' - t$, it is possible to write for Λ_p^N,

$$\begin{aligned}
\Lambda_p^N &= E\frac{1}{N}\sum_{t=1}^{N}\sum_{t'=1}^{N} r_p(t)e_0(t)r_p(t')e_0(t') = \frac{1}{N}\sum_{t=1}^{N}\sum_{\tau=1-t}^{N-t} Er_p(t)e_0(t)r_p(t+\tau)e_0(t+\tau) \\
&= \frac{1}{N}\sum_{\tau=-N+1}^{0}\sum_{t=-\tau+1}^{N} Er_p(t)r_p(t+\tau)R_{e_0}(\tau) + \frac{1}{N}\sum_{\tau=1}^{N-1}\sum_{t=1}^{N-\tau} Er_p(t)r_p(t+\tau)R_{e_0}(\tau) \\
&= \frac{1}{N}\sum_{\tau=-N+1}^{N-1} R_{e_0}(\tau)\sum_{t=1}^{N-|\tau|} Er_p(t)r_p(t+|\tau|) = \frac{1}{N}\sum_{\tau=-N+1}^{N-1} R_{e_0}(\tau)(N-|\tau|)R_{r_p}^N(\tau).
\end{aligned}$$

Finally,

$$\begin{aligned}
\Lambda_p &= \lim_{M\to\infty}\Lambda_p(M) = \lim_{M\to\infty}\lim_{N\to\infty} EZ_N^2(M) = \\
&= \lim_{M\to\infty}\lim_{N\to\infty}\sum_{\tau=-N+1}^{N-1}\frac{(N-|\tau|)}{N} R_{e_0^M}(\tau)R_{r_p}^N(\tau) = \\
&= \lim_{M\to\infty}\sum_{\tau=-M}^{M} R_{e_0^M}(\tau)R_{r_p}(\tau) = \sum_{\tau=-\infty}^{\infty} R_{e_0}(\tau)R_{r_p}(\tau).
\end{aligned}$$

Proof of Theorem 3.2.7

Under the given assumptions on $\{r_p(t)\}$ and $\{\phi_k(t)\}$ Theorem 2B.1 in Ljung (1987) is applicable and, consequently, with probability 1,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} r_p(t)\phi_k(t) = R_{r_p\phi_k}(0),\ \forall p,k.$$

The combination of (3.1) and (3.2) gives that $\phi^T(t)(\theta_0-\theta) = e(t) - e_0(t)$ and, consequently,

$$\frac{1}{N}\sum_{t=1}^{N} r_p(t)\phi^T(t)(\theta_0-\theta) = \frac{1}{N}\sum_{t=1}^{N} r_p(t)e(t) - \frac{1}{N}\sum_{t=1}^{N} r_p(t)e_0(t),\ p=1,\ldots,s.$$

By assumption $\theta=\theta_0$ is a feasible solution of this set of constraints as $e(t)=e_0(t)$, $\forall t$ is feasible. Taking the limit $N\to\infty$ gives

$$R(\theta_0-\theta) = \lim_{N\to\infty}\frac{c}{\sqrt{N}},\ c_p = \frac{1}{\sqrt{N}}\sum_{t=1}^{N} r_p(t)e(t) - \frac{1}{\sqrt{N}}\sum_{t=1}^{N} r_p(t)e_0(t),\ p=1,\ldots,s.$$

Here c is an s-dimensional vector with elements c_p, $p=1,\ldots,s$. The first contribution to c_p is bounded by the parameter bounding identification constraint (3.5). The second contribution to c_p is bounded by assumption and, consequently, $c_l(p) - c_u(p) \le c_p \le c_u(p) - c_l(p)$. Finally, the matrix R has full column rank. Hence, there exists a nonsingular square matrix R_s, consisting of n rows of R, such that for any feasible θ and $N\to\infty$,

$$R_s(\theta_0-\theta) = \lim_{N\to\infty}\frac{c}{\sqrt{N}} \iff \theta_0-\theta = \lim_{N\to\infty}\frac{R_s^{-1}c}{\sqrt{N}} = 0,$$

as both R_s^{-1} and c are bounded.

Proof of Theorem 3.3.3

The facts that $\{\psi(t)\}$, given in (3.8), is uncorrelated to $\{e_0(t)\}$ and that the latter process is stationary, give

$$E\widehat{R}_{\hat{e}}^N(\tau) = R_{\hat{e}}^N(\tau) = R_{\psi}^N(\tau) + R_{e_0}(\tau),\ |\tau| \le N-1.$$

Consequently, with (3.9),

$$E\widehat{\Lambda}_p^N = \sum_{\tau=-w}^{w} c_w(\tau)\frac{N-|\tau|}{N}R_{r_p}^N(\tau)\left(R_{\psi}^N(\tau) + R_{e_0}(\tau)\right).$$

As $\{r_p(t)\}$ is quasi-stationary, $R_{r_p}^N(\tau)$ is bounded. Moreover, the auto-covariance function of the prediction error

$$R_{\hat{e}}(\tau) = \lim_{N\to\infty} R_{\hat{e}}^N(\tau) = R_{\psi}(\tau) + R_{e_0}(\tau)$$

shows exponential decay rate in τ as both $R_\psi(\tau)$ and $R_{e_0}(\tau)$ show exponential decay rate. Also, the following asymptotic relation holds for $R^N_{r_p}(\tau)$ defined in Theorem 3.2.5,

$$\lim_{N\to\infty} R^N_{r_p}(\tau) = R_{r_p}(\tau), \ \forall |\tau| \leq w,$$

where w is allowed to tend to infinity as long as $\frac{w}{N}$ tends to 0, which is assumed to be the case. In combination with the facts that

$$\lim_{N\to\infty} \frac{N - |\tau|}{N} = 1, \ \forall |\tau| \leq w,$$

that $R_{\hat{e}}(\tau)$ shows exponential decay rate and that the window $c_w(\tau)$ satisfies assumption 3.3.2, this gives that

$$\lim_{N\to\infty} E\widehat{\Lambda}^N_p = \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau)R_{e_0}(\tau) + \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau)R_\psi(\tau) = \Lambda_p + \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau)R_\psi(\tau).$$

Under the given assumptions on $\{e_0(t)\}$, $\{\psi(\tau)\}$ and the window $c_w(\tau)$, Theorem 3.1 in Hjalmarsson (1993a) can be applied, which yields that

$$\lim_{N\to\infty} \widehat{\Lambda}^N_p = \lim_{N\to\infty} E\widehat{\Lambda}^N_p, \text{ w.p. 1.}$$

Finally, due to the fact that the auto-covariance function of $\psi(t)$ is exponentially decaying, there exists a stationary stochastic process $\{s(t)\}$ with exponentially decaying auto-covariance function for which $R_s(\tau) = R_\psi(\tau)$, $\forall \tau$. Without loss of generality it may be assumed that $\{s(t)\}$ is independent of $\{r_p(t)\}$. Define the signal $p(t) = r_p(t)s(t)$, then

$$\begin{aligned}
\sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau)R_\psi(\tau) &= \sum_{\tau=-\infty}^{\infty} R_{r_p}(\tau)R_s(\tau) = \\
&= \sum_{\tau=-\infty}^{\infty} Es(t)s(t+\tau) \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Er_p(t)r_p(t+\tau) = \\
&= \sum_{\tau=-\infty}^{\infty} \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Es(t)r_p(t)s(t+\tau)r_p(t+\tau) = \\
&= \sum_{\tau=-\infty}^{\infty} R_p(\tau) = \Phi_p(0) \geq 0,
\end{aligned}$$

where $\Phi_p(\omega)$ denotes the auto-spectrum of $\{p(t)\}$. This completes the proof.

Proof of Lemma 3.4.1

For any complex-valued x,

$$\text{Re}\left(e^{i2\pi k'/m'}x\right) \leq \left|e^{i2\pi k'/m'}x\right| \leq \left|e^{i2\pi k'/m'}\right| |x| = |x|,$$

which proves the left-hand inequality of (i). Next, for any x there exist an integer l' and $\delta \in [-\frac{\pi}{m'}, \frac{\pi}{m'})$ such that $x = |x|e^{(2\pi l'/m'+\delta)i}$, yielding

$$\begin{aligned}\mathrm{Re}\left(e^{i2\pi k'/m'}x\right) &= \mathrm{Re}\left(e^{i2\pi k'/m'}|x|e^{(2\pi l'/m'+\delta)i}\right) = |x|\mathrm{Re}\left(e^{(2\pi(k'+l')/m'+\delta)i}\right) = \\ &= |x|\cos\left(2\pi(k'+l')/m'+\delta\right).\end{aligned}$$

If $k' = k^*$ is chosen such that $k^* + l' = n'm'$ for some integer n', this gives

$$\mathrm{Re}\left(e^{i2\pi k^*/m'}x\right) = |x|\cos(2\pi n'+\delta) = |x|\cos(\delta) \geq |x|\cos(\pi/m') \Leftrightarrow$$

$$\Leftrightarrow |x| \leq \frac{\mathrm{Re}\left(e^{i2\pi k^*/m'}x\right)}{\cos(\pi/m')},$$

which proves the right-hand inequality of part (i). Finally, part (ii) immediately follows from part (i) for $m' \to \infty$.

Proof of Theorem 3.4.3

Under the given assumptions on $\{\phi_k(t)\}$ Theorem 2B.1 in Ljung (1987) is applicable and, consequently, with probability 1,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\phi_k(t)e^{-i\omega_j t} = \Psi_k(\omega_j),\ \forall\, j,k.$$

The combination of (3.1) and (3.2) gives that $\phi^T(t)(\theta_0-\theta) = e(t) - e_0(t)$ and, consequently,

$$\frac{1}{N}\sum_{t=1}^{N}e^{-i\omega_j t}\phi^T(t)(\theta_0-\theta) = \frac{1}{N}\sum_{t=1}^{N}e(t)e^{-i\omega_j t} - \frac{1}{N}\sum_{t=1}^{N}e_0(t)e^{-i\omega_j t},\ j=1,\ldots,l.$$

By assumption $\theta = \theta_0$ is a feasible solution of this set of constraints, as $e(t) = e_0(t)$, $\forall\, t$ is feasible. Taking the limit $N \to \infty$ gives

$$\Psi(\theta_0-\theta) = \lim_{N\to\infty}\frac{f}{\sqrt{N}},$$

with the j-th element of vector f given by,

$$f_j = \mathrm{Re}\left(E(\omega_j) - E_0(\omega_j)\right),\ f_{l+j} = \mathrm{Im}\left(E(\omega_j) - E_0(\omega_j)\right),\ j=1,\ldots,l,$$

Applying the result of Proposition 3.4.2 it is concluded that the first contribution to f_j and f_{l+j} is bounded by the parameter bounding identification constraint (3.11). The second contribution to f_j and f_{l+j} is bounded by assumption. Finally, the matrix Ψ has full column rank. Consequently, there exists a nonsingular square matrix Ψ_s, consisting of n rows of Ψ, such that for any feasible θ and $N \to \infty$,

$$\Psi_s(\theta_0-\theta) = \lim_{N\to\infty}\frac{f}{\sqrt{N}} \Longleftrightarrow \theta_0 - \theta = \lim_{N\to\infty}\frac{\Psi_s^{-1}f}{\sqrt{N}} = 0,$$

as both Ψ_s^{-1} and f are bounded.

Proof of Theorem 3.5.1

Under the given assumptions on $\{\phi_k(t)\}$ Theorem 2B.1 in Ljung (1987) is applicable and, consequently, with probability 1,

$$\lim_{m\to\infty} \frac{1}{m+1} \sum_{k'=0}^{m} \phi_k(t+k'T) = M_{\phi_k}(t).$$

The combination of (3.1) and (3.2) gives that $\phi^T(t)(\theta_0 - \theta) = e(t) - e_0(t)$ and, consequently, for $t = 1, \ldots, T$,

$$\frac{1}{m+1} \sum_{k'=0}^{m} \phi^T(t+k'T)(\theta_0 - \theta) = \frac{1}{m+1} \sum_{k'=0}^{m} e(t+k'T) - \frac{1}{m+1} \sum_{k'=0}^{N} e_0(t+k'T).$$

By assumption $\theta = \theta_0$ is a feasible solution of this set of constraints, as $e(t) = e_0(t)$, $\forall t$ is feasible. Taking the limit $m \to \infty$ gives

$$M(\theta_0 - \theta) = \lim_{m\to\infty} \frac{c}{\sqrt{m+1}},$$

with the t-th element of vector c given by,

$$c_t = \frac{1}{\sqrt{m+1}} \sum_{k'=0}^{m} e(t+k'T) - \frac{1}{\sqrt{m+1}} \sum_{k'=0}^{m} e_0(t+k'T), \ t = 1, \ldots, T.$$

The first contribution to c_t is bounded by the set membership identification constraints (3.15). The second contribution to c_t is bounded by assumption. Finally, the matrix M has full column rank. Consequently, there exists a nonsingular square matrix M_s, consisting of n rows of M, such that for any feasible θ and $m \to \infty$,

$$M_s(\theta_0 - \theta) = \lim_{m\to\infty} \frac{c}{\sqrt{m+1}} \iff \theta_0 - \theta = \lim_{m\to\infty} \frac{M_s^{-1}c}{\sqrt{m+1}} = 0,$$

as both M_s^{-1} and c are bounded.

Proof of Theorem 3.5.3

Substitution of $r_p(t) = 1$, $\forall t$ in Theorem 3.2.5 immediately gives that $\forall t, T$,

$$\frac{1}{\sqrt{m+1}} \sum_{k'=0}^{m} e_0(t+k'T) \overset{m\to\infty}{\longrightarrow} \mathcal{N}(0, \Lambda_T),$$

with

$$\Lambda_T = \lim_{m\to\infty} E\left(\frac{1}{\sqrt{m+1}} \sum_{k'=0}^{m} e_0(t+k'T) \right)^2 =$$

$$= \lim_{m\to\infty} \frac{1}{m+1} \sum_{k'=0}^{m} \sum_{k''=0}^{m} Ee_0(t+k'T)e_0(t+k''T) =$$

$$= \lim_{m\to\infty} \frac{1}{m+1} \sum_{k'=0}^{m} \sum_{k''=0}^{m} R_{e_0}((k'-k'')T),$$

and due to the exponential decay rate of $R_{e_0}(\tau)$,

$$\Lambda = \lim_{T\to\infty} \Lambda_T = \lim_{m\to\infty} \frac{1}{m+1} \sum_{k'=0}^{m} \sum_{k''=0}^{m} \lim_{T\to\infty} R_{e_0}((k'-k'')T) =$$

$$= \lim_{m\to\infty} \frac{1}{m+1} \sum_{k'=0}^{m} R_{e_0}(0) = R_{e_0}(0).$$

Chapter 4

A Procedure for Deterministic Uncertainty Bounding Identification

4.1 Introduction

In this chapter a procedure is developed to identify deterministic frequency and pulse response uncertainty regions for some unknown system. In particular, the objective is to identify system uncertainty sets $\mathcal{S}$, of the format specified in (2.3) and (2.4) in Chapter 2, on the basis of data and deterministic prior information about the system and the noise.

A linear system parametrization is used in terms of general basis functions. The prior information assumed about the system is a bound on the generalized pulse response parameters. The uncertainty bounding identification procedure developed in this chapter can utilize all types of prior information about the noise as presented in the previous chapter. This means that the procedure can handle time-domain, cross-covariance and frequency-domain noise constraints. In this chapter mainly a deterministic interpretation is given to these noise bounds. As such the identification approach in this chapter is referred to as deterministic uncertainty bounding identification, despite of the fact that the noise bounds may have a probabilistic interpretation, as discussed in Chapter 3.

As discussed in Section 2.3, the data in combination with the prior information about the system and the noise defines a set of feasible systems, $\mathcal{F}$. The identification objective is to outerbound this highly unstructured and implicitly defined set by the explicit system uncertainty set $\mathcal{S}$. In order to achieve this, first the set $\mathcal{F}$ is outerbounded by a set $\mathcal{L}$, which consists of linear inequalities. Next, frequency and pulse response uncertainty regions are calculated by solving a set of linear programming problems. This then leads to system uncertainty sets $\mathcal{S}$ as specified in (2.3) and (2.4), which outerbound the set $\mathcal{L}$. The sets $\mathcal{S}$ will in turn be utilized to calculate an upper bound on the H_∞- or ℓ_1-norm of the model error with respect to any specified nominal model.

The outline of this chapter is as follows. In the next section the identification setting is described, leading to the precise definition of the set of feasible systems, $\mathcal{F}$.

Section 4.3 contains the technical details describing how the data and prior information are processed, leading to a set of linear inequalities, $\mathcal{L}$. Section 4.4 describes the actual computation of the frequency response uncertainty regions, where use is made of linear programming. Also, expressions are derived for an upper bound on the H_∞-norm of the model error. In Section 4.5 it is described how pulse response uncertainty regions can be computed, which can be used to establish an upper bound on the ℓ_1-norm of the model error. Section 4.6 describes extensions to the MIMO case. Section 4.7 contains simulation examples. It is also indicated how data can be used to reliably estimate the required prior information about the system and the noise. Finally, in Section 4.8 the results are discussed.

4.2 Identification Setting

Consider the linear, time-invariant, discrete-time, causal and ℓ_∞-stable SISO system $G_0(z)$ represented by

$$G_0(z) = \sum_{k=0}^{\infty} g_0(k) B_k(z),$$

where $\{B_k(z)\}_{k=0,\ldots,\infty}$ is some specified set of basis functions, see Subsection 2.4.3, given by

$$B_k(z) = \sum_{k'=0}^{\infty} b_k(k') z^{-k'}, \ k = 0, \ldots, \infty, \tag{4.1}$$

for given and known coefficients $b_k(k')$. In Appendix 4.B some theory is presented about generalized orthonormal basis functions.

Consider given input data $\{u(t)\}$ and measured output data $\{y(t)\}$ and the following input-output relation of the data generating system,

$$y(t) = G_0(q)u(t) + e_0(t), \ t = 1, \ldots, N,$$

where N denotes the measurement time and $\{e_0(t)\}$ is an unknown additive output noise. There are no restrictions on the input signal, basically it may be determined in open loop as well as in closed loop.

Next, consider the model parametrization

$$G(z) = \sum_{k=0}^{\infty} g(k) B_k(z), \tag{4.2}$$

where $g(k)$ are the model parameters, with the corresponding input-output relation

$$y(t) = G(q)u(t) + e(t), \ t = 1, \ldots, N. \tag{4.3}$$

A deterministic identification setting is adopted, which means that a system uncertainty set is derived using certain deterministic constraints on the model $G(z)$ and the noise $\{e(t)\}$. Preferably these constraints are chosen such that the true system $G_0(z)$

and noise process $e_0(t)$ satisfy the constraints. In that case the system uncertainty set contains the true system $G_0(z)$ with probability 1.

In order to cope with unknown initial conditions the input signal in the past is assumed to be bounded by

$$|u(t)| \leq \bar{u}, \ \forall t \leq 0, \tag{4.4}$$

for some given $\bar{u}$.

The parameters $g(k)$ are bounded by

$$|g(k)| \leq \bar{g}(k), \ k = 0, \ldots, \infty, \tag{4.5}$$

for a priori given $\bar{g}(k)$. Moreover, it is assumed that the bound $\bar{g}(k)$ shows exponential decay rate for k larger than some k^*, i.e.

$$\bar{g}(k) \leq M\rho^k, \ \forall k > k^*, \tag{4.6}$$

for some given $M \geq 0$ and $\rho < 1$. In Heuberger *et al.* (1994) and in Appendix 4.B, Fact 4.B.2, it is discussed that such a bound exists when an arbitrary LTI, ℓ_∞-stable system is expanded in a generalized orthonormal basis, at least if the system is finite dimensional. In Section 4.7 it is explained how this bound may be established, if it is not available a priori, i.e. how to choose the bounds such that the true system $G_0(z)$ satisfies them. At this point it is mentioned that the bound $\bar{g}(k)$ may be chosen conservative, it can be strengthened later on using the measurement data.

Next, the following noise bounds are considered. A time-domain bound on the noise,

$$|e(t)| \leq \bar{e}(t), \ t = 1, \ldots, N, \tag{4.7}$$

for given $\bar{e}(t)$. And a cross-covariance bound on the noise,

$$\left| \frac{1}{\sqrt{\widetilde{N}}} \sum_{t=t_s}^{N} r_p(t)e(t) \right| \leq \bar{c}(p), \ p = 1, \ldots, s, \tag{4.8}$$

for given $\bar{c}(p)$. Here t_s is some user-defined integer such that $1 \leq t_s \leq N$, and $\widetilde{N}$ is defined as

$$\widetilde{N} := N - t_s + 1.$$

The influence of the choice of t_s is discussed later on. In (4.8) $r_p(t)$ are some specified signals or instrumental variables, typically highly correlated with the input $\{u(t)\}$, but uncorrelated with the noise process $\{e_0(t)\}$. See the previous chapter for a detailed treatment of the use of these cross-covariance bounds.

Frequency-domain noise bounds may be incorporated as well, see Chapter 3. As their use is quite similar to that of the cross-covariance noise bounds, details are omitted in this chapter.

There is a set of systems $G(z)$ and corresponding noise realizations $\{e(t)\}$ that satisfy the specified relations and inequalities. This set of *feasible* systems is denoted by $\mathcal{F}$,

$$\mathcal{F} = \{G(z) \mid (4.2) \text{ till } (4.8) \text{ are satisfied}\}. \tag{4.9}$$

Obviously, if the system $G_0(z)$ and the noise $\{e_0(t)\}$ satisfy the prior assumptions (4.2) till (4.8), then

$$G_0(z) \in \mathcal{F}.$$

The identification objective is to find uncertainty regions for the frequency and (weighted) pulse response of $G(z)$ given that $G(z) \in \mathcal{F}$. As indicated in Chapter 2, the identification result should be in the format of (2.3) or (2.4) respectively. The idea is to construct the uncertainty regions by direct numerical optimization, by calculating upper and lower bounds for the frequency and pulse response under the constraints that (4.2) till (4.8) are satisfied. It appears not possible to solve such optimization problems directly, due to the complicated structure of the constraints, the set $\mathcal{F}$. First, it is necessary to transform the constraints (4.2) till (4.8) to simpler ones, to be specified in a set $\mathcal{L}$, which is tractable by numerical optimization software. This implies that unavoidably approximations are introduced. These approximations always have to be conservative, i.e. the resulting constraints should describe a larger class of systems than the original ones, or in other words, the set $\mathcal{L}$ should outerbound the set $\mathcal{F}$. This then guarantees correctness of the resulting frequency and pulse response uncertainty regions, in case the system satisfies the assumptions (4.2) till (4.8). On the other hand, obviously, the added conservatism should be as small as possible, such that tight system uncertainty regions are identified.

4.3 Construction of a Set of Linear Constraints

In this section it is described how the set $\mathcal{F}$ is outerbounded by a set of linear constraints, $\mathcal{L}$, which is easy to handle in numerical optimization techniques. The two items that require attention are the influence of the unknown initial conditions and the difficulty that in fact infinitely many unknowns, $g(k)$, $k = 0, \ldots, \infty$, are involved in the description of the set of feasible systems $\mathcal{F}$.

First, substitute (4.3) into (4.7) and (4.8),

$$|y(t) - G(q)u(t)| \leq \bar{e}(t), \ t = 1, \ldots, N, \tag{4.10}$$

$$\left| \frac{1}{\sqrt{N}} \sum_{t=t_s}^{N} r_p(t)(y(t) - G(q)u(t)) \right| \leq \bar{c}(p), \ p = 1, \ldots, s. \tag{4.11}$$

Next, choose a truncation value n and split $G(z)$ into two parts,

$$G(z) = \widetilde{G}(z) + \bar{G}(z), \ \widetilde{G}(z) = \sum_{k=0}^{n} g(k)B_k(z), \ \bar{G}(z) = \sum_{k=n+1}^{\infty} g(k)B_k(z). \tag{4.12}$$

The idea is to use the finite order transfer function $\widetilde{G}(z)$ rather than $G(z)$. The influence of $\bar{G}(z)$ will be estimated in a conservative way using the bounds (4.5). This already indicates that n should preferably be chosen as large as possible. On the other hand, a large value of n will make that much computation time is required, as will become clear later.

Now consider the term $G(q)u(t)$ appearing in (4.10) and (4.11). Using (4.1) and (4.12) this term can be written as

$$\begin{aligned} G(q)u(t) &= \sum_{k=0}^{\infty} g(k)B_k(q)u(t) = \sum_{k=0}^{\infty} g(k) \sum_{k'=0}^{\infty} b_k(k')u(t-k') = \\ &= \sum_{k=0}^{n} g(k)w_k(t) + a(t) + b(t), \end{aligned}$$

with

$$\begin{aligned} w_k(t) &:= \sum_{k'=0}^{t-1} b_k(k')u(t-k'), \\ a(t) &:= \sum_{k=n+1}^{\infty} g(k) \sum_{k'=0}^{t-1} b_k(k')u(t-k'), \\ b(t) &:= \sum_{k=0}^{\infty} g(k) \sum_{k'=t}^{\infty} b_k(k')u(t-k'). \end{aligned}$$

The signals $w_k(t)$ are computable for $k = 0, \ldots, n$ and $t = 1, \ldots, N$.

Using (4.4) and (4.5) this can be further elaborated for (4.10),

$$|a(t)| = \left| \sum_{k=n+1}^{\infty} g(k) \sum_{k'=0}^{t-1} b_k(k')u(t-k') \right| \leq \sum_{k=n+1}^{\infty} \bar{g}(k) \left| \sum_{k'=0}^{t-1} b_k(k')u(t-k') \right| =: \bar{a}(t), \tag{4.13}$$

which represents a computable bound for the influence of truncation, and

$$|b(t)| = \left| \sum_{k=0}^{\infty} g(k) \sum_{k'=t}^{\infty} b_k(k')u(t-k') \right| \leq \bar{u} \sum_{k=0}^{\infty} \bar{g}(k) \sum_{k'=t}^{\infty} |b_k(k')| =: \bar{b}(t), \tag{4.14}$$

which represents a computable bound for the influence of the unknown initial conditions. Basically the bounds $\bar{a}(t)$ and $\bar{b}(t)$ are finite due to the fact that $\bar{g}(k)$ shows exponential decay rate in k and $b_k(k')$ shows exponential decay rate in k'. The actual computation may give some difficulty due to the infinite sums appearing in the expressions. However, due to the exponential decay rate an arbitrarily good approximation can be obtained by truncating the infinite sums. It is then possible to derive an upper bound on this truncation error. Details are given in Appendix 4.C.

Applying the triangle inequality this yields for the constraint (4.10),

$$\begin{aligned} \bar{e}(t) &\geq |y(t) - G(q)u(t)| = \left| y(t) - a(t) - b(t) - \sum_{k=0}^{n} g(k)w_k(t) \right| \geq \\ &\geq \left| y(t) - \sum_{k=0}^{n} g(k)w_k(t) \right| - |a(t)| - |b(t)| \geq \left| y(t) - \sum_{k=0}^{n} g(k)w_k(t) \right| - \bar{a}(t) - \bar{b}(t), \end{aligned}$$

which finally gives

$$\left| y(t) - \sum_{k=0}^{n} g(k) w_k(t) \right| \leq \bar{e}(t) + \bar{a}(t) + \bar{b}(t) =: \bar{e}_e(t), \; t = 1, \ldots, N, \tag{4.15}$$

where $\bar{e}_e(t)$ is an *extended* noise bound. In general it is important to choose n large enough such that $\bar{a}(t)$ is small compared to $\bar{e}(t)$ in order to keep the added conservatism small. For increasing t the error term $\bar{b}(t)$ will automatically vanish.

Now consider the expression (4.11),

$$\begin{aligned} \sum_{t=t_s}^{N} r_p(t) G(q) u(t) &= \sum_{t=t_s}^{N} r_p(t) \left(a(t) + b(t) + \sum_{k=0}^{n} g(k) w_k(t) \right) = \\ &= \sum_{k=0}^{n} g(k) \sum_{t=t_s}^{N} r_p(t) w_k(t) + \sum_{t=t_s}^{N} r_p(t)(a(t) + b(t)). \end{aligned}$$

Using (4.4) and (4.5) the following bounds can be established,

$$\begin{aligned} \left| \sum_{t=t_s}^{N} r_p(t) a(t) \right| &= \left| \sum_{t=t_s}^{N} r_p(t) \sum_{k=n+1}^{\infty} g(k) \sum_{k'=0}^{t-1} b_k(k') u(t-k') \right| = \\ &= \left| \sum_{k=n+1}^{\infty} g(k) \sum_{t=t_s}^{N} r_p(t) \sum_{k'=0}^{t-1} b_k(k') u(t-k') \right| \leq \\ &\leq \sum_{k=n+1}^{\infty} \bar{g}(k) \left| \sum_{t=t_s}^{N} r_p(t) \sum_{k'=0}^{t-1} b_k(k') u(t-k') \right| =: \bar{d}(p), \end{aligned} \tag{4.16}$$

which represents a computable bound for the error due to truncation. And using the substitution $t' = k' - t$,

$$\begin{aligned} \left| \sum_{t=t_s}^{N} r_p(t) \bar{b}(t) \right| &= \left| \sum_{t=t_s}^{N} r_p(t) \sum_{k=0}^{\infty} g(k) \sum_{k'=t}^{\infty} b_k(k') u(t-k') \right| = \\ &= \left| \sum_{k=0}^{\infty} g(k) \sum_{t'=0}^{\infty} \sum_{t=t_s}^{N} r_p(t) b_k(t+t') u(-t') \right| \leq \\ &\leq \sum_{k=0}^{\infty} |g(k)| \sum_{t'=0}^{\infty} \left| \sum_{t=t_s}^{N} r_p(t) b_k(t+t') \right| |u(-t')| \leq \\ &\leq \sum_{k=0}^{\infty} \bar{g}(k) \sum_{t'=0}^{\infty} \left| \sum_{t=t_s}^{N} r_p(t) b_k(t+t') \right| \bar{u} =: \bar{f}(p), \end{aligned} \tag{4.17}$$

which represents a computable bound for the error due to unknown initial conditions. Again, due to exponential decay rates the bounds $\bar{d}(p)$ and $\bar{f}(p)$ will be finite. Computational aspects are considered in Appendix 4.C.

Applying the triangle inequality to (4.11) gives

$$\begin{aligned}\bar{c}(p) &\geq \left|\frac{1}{\sqrt{\tilde{\tilde{N}}}}\sum_{t=t_s}^{N} r_p(t)(y(t)-G(q)u(t))\right| = \\ &= \frac{1}{\sqrt{\tilde{\tilde{N}}}}\left|\sum_{t=t_s}^{N} r_p(t)y(t) - \sum_{k=0}^{n} g(k)\sum_{t=t_s}^{N} r_p(t)w_k(t) + \sum_{t=t_s}^{N} r_p(t)(a(t)+b(t))\right| \geq \\ &\geq \frac{1}{\sqrt{\tilde{\tilde{N}}}}\left|\sum_{t=t_s}^{N} r_p(t)y(t) - \sum_{k=0}^{n} g(k)\sum_{t=t_s}^{N} r_p(t)w_k(t)\right| - \frac{\bar{d}(p)+\bar{f}(p)}{\sqrt{\tilde{\tilde{N}}}},\end{aligned}$$

which finally gives, for $p = 1, \ldots, s$,

$$\frac{1}{\sqrt{\tilde{\tilde{N}}}}\left|\sum_{t=t_s}^{N} r_p(t)y(t) - \sum_{k=0}^{n} g(k)\sum_{t=t_s}^{N} r_p(t)w_k(t)\right| \leq \bar{c}(p) + \frac{\bar{d}(p)+\bar{f}(p)}{\sqrt{\tilde{\tilde{N}}}} =: \bar{c}_e(p), \quad (4.18)$$

where $\bar{c}_e(p)$ is an *extended* cross-covariance bound. In general it is important to choose the truncation value n large enough such that the error contribution $\bar{d}(p)/\sqrt{\tilde{\tilde{N}}}$ is small compared to $\bar{c}(p)$. On the other hand, a large value of n means that much computational effort is required later on. Also, t_s should be chosen sufficiently large such that the influence of the unknown initial conditions is small, $\bar{f}(p)/\sqrt{\tilde{\tilde{N}}}$ small compared to $\bar{c}(p)$. However, in case t_s is chosen too large the constraints (4.8) do not contribute much to the size of the feasible system set $\mathcal{F}$, which should be avoided.

Altogether the following set of *linear* constraints for the unknown parameters $g(k)$, $k = 0, \ldots, \infty$, is obtained,

$$\mathcal{L} = \{G(z) \mid (4.2), (4.5), (4.15) \text{ and } (4.18) \text{ are satisfied}\}, \quad (4.19)$$

which by construction has the property $\mathcal{F} \subseteq \mathcal{L}$. This set serves as basis for the construction of uncertainty regions for the frequency and (weighted) pulse response of $G(z)$. Although $\mathcal{L}$ is an outer approximation of $\mathcal{F}$ it can be a very close approximation as formulated in the next proposition.

Proposition 4.3.1 *Let $\mathcal{F}$, the set of feasible systems, be defined in (4.9) and let $\mathcal{L}$, the set of linear constraints, be defined in (4.19). Then,*

$$\begin{aligned}&\textbf{(i)} \qquad \mathcal{F} \subseteq \mathcal{L},\\ &\textbf{(ii)} \quad \lim_{\bar{u}\to 0}\lim_{n\to\infty} \mathcal{L} = \mathcal{F},\end{aligned}$$

where convergence of the set $\mathcal{L}$ to the set $\mathcal{F}$ means that, first, each element in $\mathcal{L}$ converges to an element in $\mathcal{F}$ and, second, for every element f in $\mathcal{F}$ there is an element in $\mathcal{L}$ which converges to f.

Proof: Part (i) holds by construction of the set $\mathcal{L}$. If $\bar{u} \to 0$, $n \to \infty$, then $\bar{a}(t)$, $\bar{b}(t)$, $\bar{d}(p)$ and $\bar{f}(p) \to 0$ and, consequently, the constraints in $\mathcal{L}$ become identical to those in $\mathcal{F}$, which proves part (ii). □

The proposition says that $\mathcal{L}$, the outerbounding set of linear constraints, is arbitrarily close to the exact feasible system set $\mathcal{F}$, provided the influence of unknown initial conditions and truncation is negligible.

4.4 Frequency Response Uncertainty Regions and H_∞-Model Error Bound

4.4.1 Introduction

In this section it is described how the constructed set $\mathcal{L}$ can be used to calculate frequency response uncertainty regions. The identification result will be in the format as specified by (2.3) in Chapter 2,

$$\mathcal{S} = \left\{ G(z) \in \mathcal{A} \;\middle|\; G(e^{i\omega_j}) \in \mathcal{P}(\omega_j),\ \forall \omega_j \in \Omega,\ \left| \frac{\mathrm{d}\, G(e^{i\omega})}{\mathrm{d}\,\omega} \right| \leq \beta_1,\ \forall \omega \in [0, \pi] \right\}, \quad (4.20)$$

where Ω is a user-defined set of frequencies,

$$\Omega = \{\omega_1, \ldots, \omega_l\},\ 0 \leq \omega_1 < \ldots < \omega_l \leq \pi.$$

This means that a finite number of frequencies ω_j is considered and that the inter-sample frequencies are taken account of by means of interpolation. Based on this set $\mathcal{S}$ expressions are given for an upper bound on the H_∞-norm of the model error with respect to any specified nominal model.

4.4.2 Frequency Response Uncertainty Regions

In this subsection it is described how to calculate frequency response uncertainty regions $\mathcal{P}(\omega_j)$ as they appear in (4.20), with ω_j an element of the set Ω. It is also described how the derivative bound β_1 can be determined.

Define for an arbitrary frequency ω the frequency response set $\mathcal{F}_\omega$ as

$$\mathcal{F}_\omega = \left\{ G(e^{i\omega}) \mid G(z) \in \mathcal{F} \right\}, \quad (4.21)$$

and the frequency response set $\mathcal{L}_\omega$ as

$$\mathcal{L}_\omega = \left\{ G(e^{i\omega}) \mid G(z) \in \mathcal{L} \right\}, \quad (4.22)$$

which is a convex region in the complex plane, as the constraints in the set $\mathcal{L}$ define a convex uncertainty region for the parameters $g(k)$, and $G(e^{i\omega})$ depends linearly on the parameters $g(k)$. Basically the objective is to find a explicit description of $\mathcal{F}_{\omega_j}$. However, as noticed before, this is not feasible, and therefore the objective is to find an explicit description of $\mathcal{L}_{\omega_j}$. The following corollary to Proposition 4.3.1 states that this then yields a tight outer-bounding description of the set $\mathcal{F}_{\omega_j}$.

Corollary 4.4.1 *Let $\mathcal{F}_\omega$ and $\mathcal{L}_\omega$ be as defined in (4.21) and (4.22) respectively. Then,*

$$\begin{aligned} &\textbf{(i)} \qquad \mathcal{F}_\omega \subseteq \mathcal{L}_\omega, \ \forall \omega \in [0,\pi], \\ &\textbf{(ii)} \quad \lim_{\bar{u}\to 0} \lim_{n\to\infty} \mathcal{L}_\omega = \mathcal{F}_\omega, \ \forall \omega \in [0,\pi]. \end{aligned}$$

The set $\mathcal{L}_{\omega_j}$ needs to be evaluated for each $\omega_j \in \Omega$. This is done by applying the mechanism of Lemma 3.4.1 in Chapter 3. For each frequency ω_j and for some user-defined $m \geq 3$, the following quantities are evaluated,

$$\max_{G(e^{i\omega_j})\in\mathcal{L}_{\omega_j}} \mathrm{Re}\left(e^{i2\pi k/m}G(e^{i\omega_j})\right) = \max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}G(e^{i\omega_j})\right), \ k = 1,\ldots,m.$$

This corresponds to evaluating the complex-valued set $\mathcal{L}_{\omega_j}$ in m different directions in the complex plane, which then yields an outer-bounding polytope for this convex set. In case $m = 4$ it simply means that the maximum and minimum real and imaginary parts of $G(e^{i\omega_j})$ are evaluated.

The problem remains how to evaluate these quantities. This is done as follows. Define the real numbers $\mu_k(\omega_j)$ as

$$\mu_k(\omega_j) := \tilde{\mu}_k(\omega_j) + \bar{\mu}_k(\omega_j), \ k = 1,\ldots,m, \ j = 1,\ldots,l,$$

where $\tilde{\mu}_k(\omega_j)$ is the solution to the following finite dimensional linear programming problem

$$\begin{aligned} \tilde{\mu}_k(\omega_j) = \max_{g(0),\ldots,g(n)} &\sum_{k'=0}^{n} g(k')\mathrm{Re}\left(e^{i2\pi k/m}B_{k'}(e^{i\omega_j})\right) \text{ s.t.} \\ -\bar{e}_e(t) \leq y(t) - &\sum_{k'=0}^{n} g(k')w_{k'}(t) \leq \bar{e}_e(t), \ t = 1,\ldots,N, \\ -\bar{c}_e(p) \leq \frac{1}{\sqrt{\tilde{N}}}&\left(\sum_{t=t_s}^{N} r_p(t)y(t) - \sum_{k'=0}^{n} g(k')\sum_{t=t_s}^{N} r_p(t)w_{k'}(t)\right) \leq \bar{c}_e(p), \ p = 1,\ldots,s, \\ -\bar{g}(k') \leq g(k') &\leq \bar{g}(k'), \ k' = 0,\ldots,n, \end{aligned}$$

and $\bar{\mu}_k(\omega_j)$ is given by

$$\bar{\mu}_k(\omega_j) = \sum_{k'=n+1}^{\infty} \bar{g}(k')\left|\mathrm{Re}\left(e^{i2\pi k/m}B_{k'}(e^{i\omega_j})\right)\right|. \tag{4.23}$$

Then the following theorem establishes the bound obtainable in this way.

Theorem 4.4.2 *Consider $\mu_k(\omega_j)$ as defined above, and the set $\mathcal{L}$ as defined in (4.19). Then,*

$$\max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}G(e^{i\omega_j})\right) = \mu_k(\omega_j), \ k = 1,\ldots,m, \ j = 1,\ldots,l.$$

Proof: See Appendix 4.A. □

This means that $\mu_k(\omega_j)$ is an upper bound for the frequency response of $G(z)$ for a certain frequency and direction in the complex plane. It consists of two contributions. The first contribution, $\tilde{\mu}_k(\omega_j)$, is due to $\tilde{G}(z)$. It can be calculated with linear programming. See Luenberger (1984) and Karloff (1991) for a treatment of the linear programming problem and algorithms to solve LP problems. For each frequency ω_j altogether m linear programming problems need to be solved, each one with $n+1$ unknowns subject to $2(N+s+n+1)$ linear inequality constraints. For the evaluation of this first contribution, data $\{u(t)\}$, $\{y(t)\}$ are used, besides prior information (4.4) till (4.8).

The second contribution, $\bar{\mu}_k(\omega_j)$, is due to $\bar{G}(z)$, the truncated part of $G(z)$. Its calculation requires the evalution of an infinite summation, which has a finite outcome due to the exponential decay rate of $\bar{g}(k')$. Computational aspects are considered in Appendix 4.C. In the evaluation of this second contribution only use is made of the prior bounds (4.5). This generally gives rise to conservatism, as no data are used in evaluating this contribution. However, its appearance is unavoidable, as truncation is necessary in order to obtain a finite dimensional linear programming problem for the evaluation of $\tilde{\mu}_k(\omega_j)$. Thus, there is a trade-off in the choice of truncation parameter n. A small value means much conservatism, due to large bounds $\mu_k(\omega_j)$. A large value means a high computational effort, as the size of the linear programming problems to evaluate $\tilde{\mu}_k(\omega_j)$ increases.

The procedure presented above gives deterministic frequency response uncertainty regions on the basis of time-domain data. The procedure is a direct one, the data are not first transformed into the frequency domain. The procedure yields convex polytopes $\mathcal{P}_m(\omega_j)$ in the complex plane,

$$\mathcal{P}_m(\omega_j) = \left\{ G(e^{i\omega_j}) \left| \mathrm{Re}\left(e^{i2\pi k/m} G(e^{i\omega_j}) \right) \leq \mu_k(\omega_j),\ k = 1, \ldots, m \right. \right\},\ j = 1, \ldots, l, \tag{4.24}$$

with vertices $v_k(\omega_j)$, $k = 1, \ldots, m$, which are characterized by

$$\begin{aligned} \mathrm{Re}\left(e^{i2\pi k/m} v_k(\omega_j) \right) &= \mu_k(\omega_j), \\ \mathrm{Re}\left(e^{i2\pi (k+1)/m} v_k(\omega_j) \right) &= \mu_{k+1}(\omega_j), \end{aligned}$$

where $\mu_{m+1}(\omega_j) = \mu_1(\omega_j)$. In Figure 4.1 the meaning of the quantities $\mu_k(\omega_j)$ and the vertices $v_k(\omega_j)$ is illustrated for the situation that a polytope $\mathcal{P}_4(\omega)$ outerbounds the set $\mathcal{L}_{\omega_j}$.

By construction the identified polytopes have the outer-bounding property $\mathcal{L}_{\omega_j} \subseteq \mathcal{P}_m(\omega_j)$. Morover the polytopes are arbitrarily close to the convex set $\mathcal{L}_{\omega_j}$ if enough directions in the complex plane are evaluated, m large enough. The following proposition summarizes the properties of the polytopes.

Proposition 4.4.3 *Consider $\mathcal{P}_m(\omega_j)$, defined in (4.24), and $\mathcal{L}_\omega$, defined in (4.22). Then,*

$$\textbf{(i)} \qquad \mathcal{L}_{\omega_j} \subseteq \mathcal{P}_m(\omega_j),\ j = 1, \ldots, l,$$

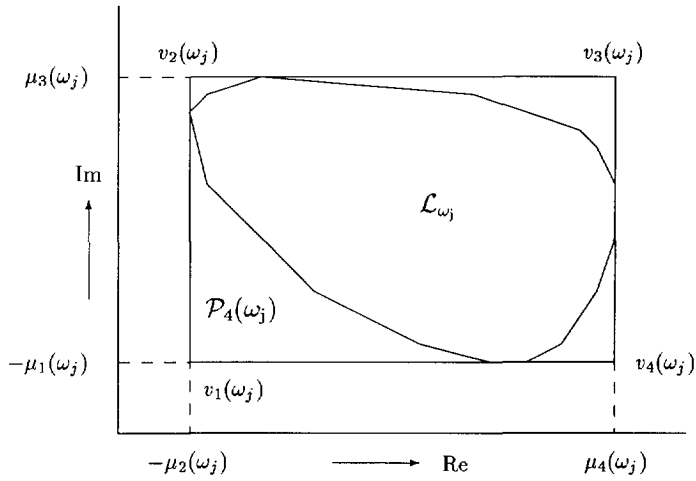

Fig. 4.1: Example of a polytope $\mathcal{P}_4(\omega_j)$ which outerbounds the convex set $\mathcal{L}_{\omega_j}$.

$$\textbf{(ii)} \quad \lim_{m\to\infty} \mathcal{P}_m(\omega_j) = \mathcal{L}_{\omega_j}, \ j = 1, \ldots, l.$$

Proof: Part (i) follows by construction of the polytopes $\mathcal{P}_m(\omega_j)$. Part (ii) follows from the fact that $\mathcal{L}_{\omega_j}$ is convex and $\mathcal{P}_m(\omega_j)$ is a tight convex approximation of this convex set with respect to an increasing number of different orientations in the complex plane. □

The combination of Proposition 4.4.3 and Corollary 4.4.1 lead to the observation that the identification procedure is non-conservative in the sense that, asymptotically in m, n and $\bar{u}$, no smaller uncertainty regions can be identified on the basis of the given data and prior information. In other words, this means that the procedure for deterministic uncertainty bounding identification as presented here, is optimal.

The polytopes $\mathcal{P}_m(\omega_j)$ bound the system's frequency response for frequencies in the set Ω. The other frequencies in the interval $[0, \pi]$ are dealt with by means of interpolation. For that purpose the following lemma establishes a bound β_1 on the first derivative of $G(e^{i\omega})$ to ω, based on the prior information (4.5).

Lemma 4.4.4 *Let $G(z)$ be any transfer function satisfying (4.2), with the coefficients $g(k)$ bounded as in (4.5). Let β_1 be defined as*

$$\beta_1 := \sum_{k=0}^{\infty} \bar{g}(k) \left\| \frac{\mathrm{d}\, B_k(z)}{\mathrm{d}\, z} \right\|_\infty .$$

Then,

$$\left| \frac{\mathrm{d}\, G(e^{i\omega})}{\mathrm{d}\, \omega} \right| \leq \beta_1, \ \forall \omega \in [0, \pi].$$

Proof: See Appendix 4.A. □

Computational aspects with respect to the actual calculation of β_1 are considered in Appendix 4.C. The uncertainty regions $\mathcal{P}_m(\omega_j)$ together with the bound β_1 yield a system uncertainty set $\mathcal{S}$ as specified by (4.20). By construction this set outerbounds the set $\mathcal{L}$, as stated in the next proposition.

Proposition 4.4.5 *Let $\mathcal{S}$, the system uncertainty set, be defined in (4.20), with $\mathcal{P}(\omega_j)$ equal to $\mathcal{P}_m(\omega_j)$, defined in (4.24), and with β_1 defined in Lemma 4.4.4. Moreover, let $\mathcal{L}$, the set of linear constraints, be defined in (4.19). Then,*

$$\mathcal{L} \subseteq \mathcal{S},$$

Proof: The proposition holds by construction of the set $\mathcal{S}$. □

4.4.3 An Upper Bound on the H_∞-Norm of the Model Error

In this subsection the problem is considered of deriving an H_∞-model error bound with respect to any user-defined nominal model. The set $\mathcal{S}$, as identified in the previous subsection, serves as the basis for this derivation.

Consider any stable, finite dimensional, LTI nominal model $\widehat{G}(z)$, and finite dimensional, LTI weighting $W(z)$ which is stable and stably invertible. The model error $\Delta(z)$ is defined as

$$\Delta(z) = W(z)\left(G(z) - \widehat{G}(z)\right). \tag{4.25}$$

An outer-bounding description for $\Delta(e^{i\omega_j})$ is given by the polytope with vertices $W(e^{i\omega_j})\left(v_k(\omega_j) - \widehat{G}(e^{i\omega_j})\right)$. Using this and the $\mu_k(\omega_j)$ as defined in the previous subsection, a frequency dependent upper bound on the weighted model error can be determined for the frequencies in the set Ω, as specified in the next proposition.

Proposition 4.4.6 *Consider $\Delta(z)$ defined in (4.25) for given $W(z)$ and $\widehat{G}(z)$. Let the sets $\mathcal{S}$ and $\mathcal{L}$ satisfy the conditions in Proposition 4.4.5. Define*

$$\delta_{\mathrm{l}}(\omega_j) := \left|W(e^{i\omega_j})\right| \max_{k=1,\ldots,m}\left(\mu_k(\omega_j) - \mathrm{Re}\left(e^{i2\pi k/m}\widehat{G}(e^{i\omega_j})\right)\right),\ j = 1,\ldots,l,$$

$$\delta_{\mathrm{u}}(\omega_j) := \left|W(e^{i\omega_j})\right| \max_{k=1,\ldots,m}\left|v_k(\omega_j) - \widehat{G}(e^{i\omega_j})\right|,\ j = 1,\ldots,l.$$

Then, for $j = 1,\ldots,l$,

(i) $$\max_{\Delta = W(G-\widehat{G}),\, G\in\mathcal{S}} \left|\Delta(e^{i\omega_j})\right| = \delta_{\mathrm{u}}(\omega_j),$$

(ii) $$\delta_{\mathrm{l}}(\omega_j) \leq \max_{\Delta = W(G-\widehat{G}),\, G\in\mathcal{L}} \left|\Delta(e^{i\omega_j})\right| \leq \delta_{\mathrm{u}}(\omega_j) \leq \frac{\delta_{\mathrm{l}}(\omega_j)}{\cos(\frac{\pi}{m})},$$

(iii) $$\lim_{m\to\infty} \delta_{\mathrm{l}}(\omega_j) = \lim_{m\to\infty} \delta_{\mathrm{u}}(\omega_j) = \max_{\Delta = W(G-\widehat{G}),\, G\in\mathcal{L}} \left|\Delta(e^{i\omega_j})\right|.$$

Proof: See Appendix 4.A. □

In this proposition a model error bound has been established for frequencies in the set Ω. Using an interpolation argument, based on the derivative bound β_1, a bound on the H_∞-norm of the model error can be established. This interpolation argument has previously been used by De Vries and Van den Hof (1992) and Bayard (1992). First, an auxiliary lemma is presented, which gives a bound on the H_∞-norm of any system $G(z)$ in the set $\mathcal{S}$.

Lemma 4.4.7 *Let $G(z)$ be any system in the set $\mathcal{S}$ defined by (4.20), where the uncertainty regions $\mathcal{P}(\omega_j)$ are the polytopes $\mathcal{P}_m(\omega_j)$ with vertices $v_k(\omega_j)$. Define λ_j, $j = 1, \ldots, l$, as the distance between two subsequent frequencies,*

$$\begin{aligned}
\lambda_1 &:= \max\{2\omega_1, \omega_2 - \omega_1\}, \\
\lambda_j &:= \max\{\omega_j - \omega_{j-1}, \omega_{j+1} - \omega_j\}, \ j = 2, \ldots, l-1, \\
\lambda_l &:= \max\{\omega_l - \omega_{l-1}, 2(\pi - \omega_l)\},
\end{aligned}$$

and define the bound β_2 as

$$\beta_2 := \max_{j=1,\ldots,l} \left[\max_{k=1,\ldots,m} |v_k(\omega_j)| + \frac{1}{2}\lambda_j \beta_1 \right],$$

with β_1 the derivative bound in (4.20). Then,

$$\|G(z)\|_\infty \leq \beta_2.$$

Proof: See Appendix 4.A. □

Using the Lemmas 4.4.4 and 4.4.7 and Proposition 4.4.6, an upper bound on the H_∞-norm of the model error can easily be derived. The result is given in the next theorem.

Theorem 4.4.8 *Consider $\Delta(z)$ defined in (4.25), where $W(z)$ is any given finite dimensional, LTI, stable and stably invertible weighting, and $\widehat{G}(z)$ is any given finite dimensional, LTI, stable nominal model. Let the sets $\mathcal{S}$ and $\mathcal{L}$ satisfy the conditions in Proposition 4.4.5. Moreover, let $\delta_u(\omega_j)$ be as defined in Proposition 4.4.6 and let β_2 and λ_j be as defined in Lemma 4.4.7. Finally, denote*

$$\beta := \left\| \frac{\mathrm{d}\, W(z)\widehat{G}(z)}{\mathrm{d}\, z} \right\|_\infty + \left\| \frac{\mathrm{d}\, W(z)}{\mathrm{d}\, z} \right\|_\infty \beta_2 + \|W(z)\|_\infty \beta_1.$$

Then an upper bound on the H_∞-norm of the model error is given by

$$\max_{\Delta = W(G-\widehat{G}), G \in \mathcal{L}} \|\Delta(z)\|_\infty \leq \max_{\Delta = W(G-\widehat{G}), G \in \mathcal{S}} \|\Delta(z)\|_\infty \leq \max_{j=1,\ldots,l} \left[\delta_u(\omega_j) + \frac{1}{2}\lambda_j \beta \right],$$

the second inequality becoming equality if $\Omega \to [0, \pi]$ and the first inequality becoming equality if in addition $m \to \infty$.

Proof: See Appendix 4.A. □

This theorem provides an H_∞-model error bound on the basis of the identified system uncertainty set $\mathcal{S}$. Asymptotically the given H_∞-model error bound is non-conservative. But in the non-asymptotic case the given upper bound need not be tight. In fact the contribution of the interpolation, $\lambda_j\beta/2$, is worst-case and in general overly pessimistic. In practical applications a more ad hoc approach is recommended. The following H_∞-model error bound will suffice in general,

$$\|\Delta(z)\|_\infty \leq \max_{j=1,\ldots,l} \delta_u(\omega_j),$$

provided sufficiently many frequencies have been used in the frequency grid Ω. In practice no problems are expected from the (small) error introduced by this approximation. This ad hoc approach has also been recommended in Bayard and Yam (1993).

4.5 Pulse Response Uncertainty Regions and ℓ_1-Model Error Bound

4.5.1 Introduction

In this section it is described how the constructed set $\mathcal{L}$ can be used to calculate (weighted) pulse response uncertainty regions. The identification result should be in the format as specified by (2.4) in Chapter 2,

$$\mathcal{S} = \left\{ G(z) \in \mathcal{A} \;\middle|\; G(z) = W^{-1}(z) \sum_{k=0}^{\infty} p(k) z^{-k},\ p(k) \in [p_l(k), p_u(k)],\ k = 0, \ldots, \infty \right\}, \tag{4.26}$$

for any user-defined, LTI, finite dimensional, stable and stably invertible weighting $W(z)$. The sequence $\{p(k)\}_{k=0,\ldots,\infty}$ is referred to as a pulse response sequence, keeping in mind that it is in fact the pulse response of $W(z)G(z)$ and not of $G(z)$. It will be shown that the identification result can straightforwardly be used to specify an upper bound on the ℓ_1-norm of the model error with respect to any specified nominal model.

4.5.2 Pulse Response Uncertainty Regions

In this subsection it is described how to calculate pulse response uncertainty regions $[p_l(k), p_u(k)]$ as they appear in (4.26).

Define, for $k = 0, \ldots, \infty$, the pulse response set $\mathcal{F}_k$ as

$$\mathcal{F}_k = \left\{ p(k) \;\middle|\; G(z) = W^{-1}(z) \sum_{k=0}^{\infty} p(k) z^{-k} \in \mathcal{F} \right\}, \tag{4.27}$$

and the pulse response set $\mathcal{L}_k$ as

$$\mathcal{L}_k = \left\{ p(k) \;\middle|\; G(z) = W^{-1}(z) \sum_{k=0}^{\infty} p(k) z^{-k} \in \mathcal{L} \right\}, \tag{4.28}$$

which is a convex region in the complex plane as $\mathcal{L}$ is a convex set, in the parameters $g(k)$. The following corollary to Proposition 4.3.1 states that the set $\mathcal{L}_k$ tightly outerbounds the set $\mathcal{F}_k$.

Corollary 4.5.1 *Let $\mathcal{F}_k$ be defined in (4.27) and $\mathcal{L}_k$ in (4.28). Then,*

$$\begin{aligned} &\textbf{(i)} && \mathcal{F}_k \subseteq \mathcal{L}_k, \ \forall k, \\ &\textbf{(ii)} && \lim_{\bar{u}\to 0} \lim_{n\to\infty} \mathcal{L}_k = \mathcal{F}_k, \ \forall k. \end{aligned}$$

Analogously to the H_∞-case treated in the previous section, the set $\mathcal{L}_k$ is used to identify the pulse response parameter bounds $p_{\mathrm{l}}(k)$ and $p_{\mathrm{u}}(k)$. This is done by evaluating the quantities,

$$\begin{aligned} \min_{p(k)\in\mathcal{L}_k} p(k) &= \min_{G(z)\in\mathcal{L}} p(k), \\ \max_{p(k)\in\mathcal{L}_k} p(k) &= \max_{G(z)\in\mathcal{L}} p(k). \end{aligned}$$

Because of convexity of the set $\mathcal{L}_k$ the following property holds,

$$\left[\min_{p(k)\in\mathcal{L}_k} p(k), \max_{p(k)\in\mathcal{L}_k} p(k)\right] = \mathcal{L}_k.$$

The only problem that remains is, how to evaluate these quantities. Analogously to the H_∞-case discussed in the previous section, this is done by means of linear programming. Basically $p(k)$ needs to be evaluated for $k = 0, \ldots, \infty$, but obviously, it is impossible to solve infinitely many linear programming problems. Therefore this is done only for $k = 0, \ldots, m$, with m some user-defined integer. For $k > m$ a simple upper bound will be given, which is conservative but computationally less demanding.

First, define $S_{k'}(z)$ as

$$S_{k'}(z) := \sum_{k=0}^{\infty} s_{k'}(k) z^{-k} := B_{k'}(z) W(z), \ k' = 0, \ldots, \infty,$$

then

$$\begin{aligned} \sum_{k=0}^{\infty} p(k) z^{-k} &= G(z)W(z) = \sum_{k'=0}^{\infty} g(k') B_{k'}(z) W(z) = \sum_{k'=0}^{\infty} g(k') S_{k'}(z) = \\ &= \sum_{k'=0}^{\infty} g(k') \sum_{k=0}^{\infty} s_{k'}(k) z^{-k} = \sum_{k=0}^{\infty} \left(\sum_{k'=0}^{\infty} g(k') s_{k'}(k) \right) z^{-k}, \end{aligned} \tag{4.29}$$

and hence,

$$p(k) = \sum_{k'=0}^{\infty} g(k') s_{k'}(k), \ k = 0, \ldots, \infty.$$

Apparently, each pulse response parameter $p(k)$ is a linear combination of infinitely many parameters $g(k')$. The objective is to evaluate lower and upper bounds for $p(k)$

by means of linear programming. But in the LP problem the number of unknowns $g(k')$ should be finite. Therefore each pulse response parameter $p(k)$ is split up into two parts, a first part $\widetilde{p}(k)$, and a tail $\bar{p}(k)$, defined as follows,

$$\widetilde{p}(k) := \sum_{k'=0}^{n} g(k')s_{k'}(k),\ k = 0,\ldots,m,$$

$$\bar{p}(k) := \sum_{k'=n+1}^{\infty} g(k')s_{k'}(k),\ k = 0,\ldots,m,$$

such that, obviously, $p(k) = \widetilde{p}(k) + \bar{p}(k)$. Next, lower and upper bounds $\widetilde{p}_{\mathrm{l}}(k)$ and $\widetilde{p}_{\mathrm{u}}(k)$ are derived for $\widetilde{p}(k)$ by means of linear programming. Also, lower and upper bounds $\bar{p}_{\mathrm{l}}(k)$ and $\bar{p}_{\mathrm{u}}(k)$ are derived for $\bar{p}(k)$ on the basis of the prior bounds $\bar{g}(k')$. The combination of these bounds then yields lower and upper bounds for $p(k)$.

Define the real numbers $p_{\mathrm{l}}(k)$ and $p_{\mathrm{u}}(k)$ as

$$p_{\mathrm{l}}(k) := \widetilde{p}_{\mathrm{l}}(k) + \bar{p}_{\mathrm{l}}(k),\ k = 0,\ldots,m,$$
$$p_{\mathrm{u}}(k) := \widetilde{p}_{\mathrm{u}}(k) + \bar{p}_{\mathrm{u}}(k),\ k = 0,\ldots,m,$$

where $\widetilde{p}_{\mathrm{l}}(k)$ is the solution to the following finite dimensional linear programming problem

$$\widetilde{p}_{\mathrm{l}}(k) = \min_{g(0),\ldots,g(n)} \sum_{k'=0}^{n} g(k')s_{k'}(k) \text{ s.t.}$$

$$-\bar{e}_{\mathrm{e}}(t) \le y(t) - \sum_{k'=0}^{n} g(k')w_{k'}(t) \le \bar{e}_{\mathrm{e}}(t),\ t = 1,\ldots,N,$$

$$-\bar{c}_{\mathrm{e}}(p) \le \frac{1}{\sqrt{\widetilde{N}}}\left(\sum_{t=t_s}^{N} r_p(t)y(t) - \sum_{k'=0}^{n} g(k')\sum_{t=t_s}^{N} r_p(t)w_{k'}(t)\right) \le \bar{c}_{\mathrm{e}}(p),\ p = 1,\ldots,s,$$

$$-\bar{g}(k') \le g(k') \le \bar{g}(k'),\ k' = 0,\ldots,n,$$

$\widetilde{p}_{\mathrm{u}}(k)$ is the solution to the following finite dimensional linear programming problem

$$\widetilde{p}_{\mathrm{u}}(k) = \max_{g(0),\ldots,g(n)} \sum_{k'=0}^{n} g(k')s_{k'}(k) \text{ s.t.}$$

$$-\bar{e}_{\mathrm{e}}(t) \le y(t) - \sum_{k'=0}^{n} g(k')w_{k'}(t) \le \bar{e}_{\mathrm{e}}(t),\ t = 1,\ldots,N,$$

$$-\bar{c}_{\mathrm{e}}(p) \le \frac{1}{\sqrt{\widetilde{N}}}\left(\sum_{t=t_s}^{N} r_p(t)y(t) - \sum_{k'=0}^{n} g(k')\sum_{t=t_s}^{N} r_p(t)w_{k'}(t)\right) \le \bar{c}_{\mathrm{e}}(p),\ p = 1,\ldots,s,$$

$$-\bar{g}(k') \le g(k') \le \bar{g}(k'),\ k' = 0,\ldots,n,$$

$\bar{p}_{\mathrm{l}}(k)$ is given by

$$\bar{p}_{\mathrm{l}}(k) = -\sum_{k'=n+1}^{\infty} \bar{g}(k')\,|s_{k'}(k)|\,, \tag{4.30}$$

and $\bar{p}_u(k)$ is given by

$$\bar{p}_u(k) = \sum_{k'=n+1}^{\infty} \bar{g}(k') \, |s_{k'}(k)| \, . \tag{4.31}$$

Finally, define

$$\begin{aligned} p_l(k) &:= -\sum_{k'=0}^{\infty} \bar{g}(k') \, |s_{k'}(k)| \, , \quad k = m+1, \ldots, \infty, \\ p_u(k) &:= \sum_{k'=0}^{\infty} \bar{g}(k') \, |s_{k'}(k)| \, , \quad k = m+1, \ldots, \infty. \end{aligned}$$

Then the following theorem establishes the pulse response parameter bounds obtainable in this way.

Theorem 4.5.2 *Consider the set $\mathcal{L}_k$ defined in (4.28) and $p_l(k)$ and $p_u(k)$ as defined above. Then,*

$$\begin{aligned} \min_{p(k) \in \mathcal{L}_k} p(k) &= p_l(k), \quad k = 0, \ldots, m, \\ \min_{p(k) \in \mathcal{L}_k} p(k) &\geq p_l(k), \quad k = m+1, \ldots, \infty, \\ \max_{p(k) \in \mathcal{L}_k} p(k) &= p_u(k), \quad k = 0, \ldots, m, \\ \max_{p(k) \in \mathcal{L}_k} p(k) &\leq p_u(k), \quad k = m+1, \ldots, \infty. \end{aligned}$$

Proof: See Appendix 4.A. □

This means that $p_l(k)$ and $p_u(k)$ as specified, are the desired lower and upper bound on the pulse response parameters $p(k)$. Completely similar to the quantities $\mu_k(\omega_j)$ in the H_∞-case (Theorem 4.4.2) the bounds on $p(k)$ for $k \leq m$, consist of two contributions. The first contribution, $\tilde{p}_l(k)$, resp. $\tilde{p}_u(k)$, is due to $\tilde{G}(z)$, and can be calculated by solving two linear programming problems with $n+1$ unknowns subject to $2(N+s+n+1)$ linear inequality constraints. The second contribution, $\bar{p}_l(k)$, resp. $\bar{p}_u(k)$, requires the evaluation of an infinite summation. How to actually compute this infinite summation is discussed in Appendix 4.C.

Obviously, the bounds $p_l(k)$ and $p_u(k)$, $k = 0, \ldots, \infty$, as specified in Theorem 4.5.2 yield a system uncertainty set $\mathcal{S}$ as specified by (4.26). By construction this set outerbounds the set $\mathcal{L}$, as stated in the next proposition.

Proposition 4.5.3 *Let $\mathcal{S}$, the system uncertainty set, be defined in (4.26), with $p_l(k)$ and $p_u(k)$ as defined above. Moreover, let $\mathcal{L}$, the set of linear constraints, be defined in (4.19). Then,*

$$\mathcal{L} \subseteq \mathcal{S},$$

Proof: The proposition holds by construction of the set $\mathcal{S}$. □

Finally, it is mentioned that $p_l(k)$ and $p_u(k)$ for $k > m$ actually need not be computed, as their individual values are not of importance. In the next subsection only a certain summation of $p_l(k)$ and $p_u(k)$ appears to be of interest.

4.5.3 An Upper Bound on the ℓ_1-Norm of the Model Error

In this subsection it is shown how the identified set $\mathcal{S}$ can be used to derive an upper bound on the ℓ_1-norm of the model error.

Consider any stable, finite dimensional, LTI nominal model $\widehat{G}(z)$ and the weighting $W(z)$ as specified in the previous subsection. The model error $\Delta(z)$ is defined as

$$\Delta(z) = W(z)\left(G(z) - \widehat{G}(z)\right). \tag{4.32}$$

In the following theorem an upper bound for the ℓ_1-norm of $\Delta(z)$ is specified.

Theorem 4.5.4 *Consider the model error $\Delta(z)$ as defined in (4.32) with $\widehat{G}(z)$ any given finite dimensional, LTI, stable nominal model. Denote*

$$\widehat{G}(z)W(z) = \sum_{k=0}^{\infty} \widehat{p}(k) z^{-k}.$$

Let the sets $\mathcal{S}$ and $\mathcal{L}$ satisfy the conditions of Proposition 4.5.3. Finally, denote

$$\gamma := \sum_{k=m+1}^{\infty} \left(p_u(k) + |\widehat{p}(k)|\right).$$

Then an upper bound on the ℓ_1-norm of $\Delta(z)$ is given by

$$\max_{\Delta = W(G-\widehat{G}), G\in\mathcal{L}} \|\Delta(z)\|_{\ell_1} \leq \max_{\Delta = W(G-\widehat{G}), G\in\mathcal{S}} \|\Delta(z)\|_{\ell_1} =$$

$$= \sum_{k=0}^{m} \frac{1}{2}\left(p_u(k) - p_l(k) + |p_u(k) + p_l(k) - 2\widehat{p}(k)|\right) + \gamma.$$

Proof: See Appendix 4.A. □

The upper bound on the ℓ_1-norm of the model error established in this way, requires the evaluation of an infinite summation, γ. In Appendix 4.C it is investigated how this sum can actually be computed.

Remark 4.5.5 *Note that the upper bound established is conservative in the sense that in general*

$$\max_{\Delta = W(G-\widehat{G}), G\in\mathcal{L}} \|\Delta(z)\|_{\ell_1} < \max_{\Delta = W(G-\widehat{G}), G\in\mathcal{S}} \|\Delta(z)\|_{\ell_1},$$

which is due to the fact that the maximum of a sum is generally less than the sum of the maxima of the individual terms. This problem did not occur in the H_∞-case, where (asymptotically) non-conservativeness has been shown, see Theorem 4.4.8. For the ℓ_1-case it means that, on the basis of the same data and prior information, basically a smaller upper bound on the ℓ_1-norm of the model error should be obtainable, by using another identification procedure. It deserves future research attention to develop an identification procedure which yields tighter ℓ_1-model error bounds. In Hakvoort (1991) such a procedure has been proposed, which reformulates the problem to a Generalized Linear Complementarity Problem (De Moor, 1988, Ch. 3), but in general this procedure seems computationally infeasible.

4.6 The Multivariable Case

In this section multivariable extensions to the procedure for deterministic uncertainty bounding identification are presented. Only the main differences with the SISO case are highlighted, most technical details are omitted.

4.6.1 Identification Setting

Consider a LTI, discrete-time, causal and ℓ_∞-stable MIMO system $G_0(z)$ with q' inputs and p' outputs. This system can be represented by

$$G_0(z) = \begin{bmatrix} G_0^{11}(z) & \cdots & G_0^{1q'}(z) \\ \vdots & \ddots & \vdots \\ G_0^{p'1}(z) & \cdots & G_0^{p'q'}(z) \end{bmatrix},$$

$$G_0^{i'j'}(z) = \sum_{k=0}^{\infty} g_0^{i'j'}(k) B_k^{i'j'}(z),$$

where $\{B_k^{i'j'}(z)\}_{k=0,\ldots,\infty}$ is some specified set of basis functions for the SISO transfer function $G_0^{i'j'}(z)$, completely similar to the representation used in Section 4.2.

Consider given vector-valued input data $\{u(t)\}$ and measured vector-valued output data $\{y(t)\}$ and the following input-output relation of the data generating system,

$$y(t) = G_0(q)u(t) + e_0(t), \ t = 1, \ldots, N,$$

where N is the measurement time and $\{e_0(t)\}$ is an unknown additive vector-valued output noise.

The model $G(z)$ is parametrized in the same way as $G_0(z)$, replacing $G_0^{i'j'}(z)$ by $G_{i'j'}(z)$ and $g_0^{i'j'}(k)$ by $g_{i'j'}(k)$, compare with (4.2). The model input-output relation is given by (4.3).

The prior information consists of a specified bound on the input signal in the past,

$$|u_{j'}(t)| \leq \bar{u}_{j'}, \ \forall t \leq 0, \ j' = 1, \ldots, q',$$

a specified bound on the parameters $g_{i'j'}(k)$,

$$|g_{i'j'}(k)| \leq \bar{g}_{i'j'}(k), \ k = 0, \ldots, \infty, \ i' = 1, \ldots, p', \ j' = 1, \ldots, q',$$

a time-domain bound on the noise,

$$|e_{i'}(t)| \leq \bar{e}_{i'}(t), \ t = 1, \ldots, N, \ i' = 1, \ldots, p',$$

and/or a cross-covariance bound on the noise,

$$\left| \frac{1}{\sqrt{\widetilde{N}}} \sum_{t=t_s}^{N} r_p(t) e_{i'}(t) \right| \leq \bar{c}_{i'}(p), \ p = 1, \ldots, s, \ i' = 1, \ldots, p'.$$

Similar to its definition in Section 4.2, the set $\mathcal{F}$ is defined as the set of all transfer functions $G(z)$ such that the system and noise bounds are satisfied, i.e. the set of all $G(z)$ which are consistent with data and prior information as specified above.

The identification objective is to find explicit and tight uncertainty regions for the frequency and (weighted) pulse response of each $G_{i'j'}(z)$, $i' = 1, \ldots, p'$, $j' = 1, \ldots, q'$, given that $G(z) \in \mathcal{F}$. Again, these identification problems are tackled with linear programming.

4.6.2 Construction of a Set of Linear Constraints

First, the data are processed, analogously to the procedure of Section 4.3. It is noticed that, instead of considering one big MIMO problem, the identification problem can be split up into p' independent MISO problems, one for each entry of the output signal. This is due to the fact that the transfer functions $G_{i'j'}(z)$, $i' = 1, \ldots, p'$, $j' = 1, \ldots, q'$, are parametrized independently from each other. The input-output relation for the i'th output signal entry is given by

$$y_{i'}(t) = \sum_{j'=1}^{q'} G_{i'j'}(q) u_{j'}(t), \ t = 1, \ldots, N.$$

The steps taken in Section 4.3 straightforwardly carry over to the MIMO, or in fact MISO, case. Each transfer function $G_{i'j'}(q)$ is truncated with a suitable choice of $n_{i'j'}$, and for each j' the signals $w_k(t)$, $\bar{a}(t)$, $\bar{b}(t)$, $\bar{d}(p)$ and $\bar{f}(p)$ are calculated. Adding the various contributions leads to extended noise bounds $\bar{e}_e(t)$ and $\bar{c}_e(p)$. Doing this for all i' ultimately leads to a set $\mathcal{L}$, consisting of linear constraints on the parameters $g_{i'j'}(k)$, and which by construction has the property $\mathcal{F} \subseteq \mathcal{L}$.

4.6.3 Frequency Response Uncertainty Regions

Frequency response uncertainty regions are identified for each transfer function $G_{i'j'}(z)$ separately. This is done by calculating the quantities,

$$\mu_k^{i'j'}(\omega_j) = \max_{G(z) \in \mathcal{L}} \operatorname{Re}\left(e^{i2\pi k/m} G_{i'j'}(e^{i\omega_j})\right), \ k = 1, \ldots, m, \ j = 1, \ldots, l.$$

Similar to the SISO case, Section 4.4, these quantities are calculated as follows,

$$\mu_k^{i'j'}(\omega_j) = \tilde{\mu}_k^{i'j'}(\omega_j) + \bar{\mu}_k^{i'j'}(\omega_j), \ k = 1, \ldots, m, \ j = 1, \ldots, l,$$

where $\tilde{\mu}_k^{i'j'}(\omega_j)$ is the solution to a finite dimensional linear programming problem, similar to the one considered in Section 4.4, and $\bar{\mu}_k^{i'j'}(\omega_j)$ is given by

$$\bar{\mu}_k^{i'j'}(\omega_j) = \sum_{k'=n_{i'j'}+1}^{\infty} \bar{g}_{i'j'}(k') \left| \operatorname{Re}\left(e^{i2\pi k/m} B_{k'}^{i'j'}(e^{i\omega_j})\right) \right|.$$

Completely analogously to the SISO case this leads to convex polytopes $\mathcal{P}_m^{i'j'}(\omega_j)$ in the complex plane, which constitute uncertainty regions for the system's frequency response, in the sense that

$$G_0^{i'j'}(e^{i\omega_j}) \in \mathcal{P}_m^{i'j'}(\omega_j),\ j = 1, \ldots, l,$$

provided the utilized prior information is correct.

The polytopes $\mathcal{P}_m^{i'j'}(\omega_j)$ can straightforwardly be applied to derive an upper bound on the H_∞-norm of the weighted model error for each transfer function $G_{i',j'}(z)$ separately. Let the model error $\Delta_{i'j'}(z)$ be defined as

$$\Delta_{i'j'}(z) = W_{i'j'}(z)\left(G_{i'j'}(z) - \widehat{G}_{i'j'}(z)\right),$$

for any stable, finite dimensional, LTI nominal model $\widehat{G}_{i'j'}(z)$ and weighting $W_{i'j'}(z)$. The results derived in Section 4.4 for the SISO case, carry over directly to the present situation. Application of Theorem 4.4.8 immediately yields an upper bound on the H_∞-norm of the model error, i.e. some bound $\delta_{i'j'}$ such that

$$\max_{G \in \mathcal{L}} \|\Delta_{i'j'}(z)\|_\infty \leq \delta_{i'j'},\ i' = 1, \ldots, p',\ j' = 1, \ldots, q',$$

the inequality becoming equality if $m \to \infty$ and $\Omega \to [0, \pi]$. This means that automatically a complex-valued, structured uncertainty representation is obtained. The identification procedure delivers altogether $p'q'$ scalar uncertainty blocks, which can be used in a μ-robust controller synthesis or analysis procedure, see Subsection 2.2.3 in Chapter 2. This structured uncertainty approach has the clear advantage to an unstructured one, that it is not sensitive to relative scaling. If, for example, one input signal entry $u_{j^*}(t)$ is relatively poorly excited compared to the other entries, this will become apparent in the size of the uncertainty regions $\mathcal{P}_m^{i'j^*}(\omega_j)$ and H_∞-model error bound $\delta_{i'j^*}$. Such information would not be obtained in case the identification delivers just one H_∞-bound on the multivariable model error $\Delta(z)$.

4.6.4 Pulse Response Uncertainty Regions

Pulse response uncertainty regions are identified for each (weighted) transfer function $G_{i'j'}(z)$ separately. Consider any finite dimensional, stable and stably invertible, LTI weighting $W_{i'j'}(z)$ and the relation

$$G_{i'j'}(z) = W_{i'j'}^{-1}(z) \sum_{k=0}^{\infty} p_{i'j'}(k) z^{-k}.$$

The identification objective is to find (tight) uncertainty intervals for the pulse response sequence $\{p_{i'j'}(k)\}_{k=0,\ldots,\infty}$,

$$p_{\mathrm{l}}^{i'j'}(k) \leq p_{i'j'}(k) \leq p_{\mathrm{u}}^{i'j'}(k),\ k = 0, \ldots, \infty.$$

These bounds can be calculated with linear programming by evaluating

$$\min_{G(z)\in\mathcal{L}} p_{i'j'}(k), \quad \max_{G(z)\in\mathcal{L}} p_{i'j'}(k).$$

Define

$$S_{k'}^{i'j'}(z) := \sum_{k=0}^{\infty} s_{k'}^{i'j'}(k) z^{-k} := B_{k'}^{i'j'}(z) W_{i'j'}(z), \; k' = 0, \ldots, \infty,$$

which implies that

$$p_{i'j'}(k) = \sum_{k'=0}^{\infty} g_{i'j'}(k') s_{k'}^{i'j'}(k), \; k = 0, \ldots, \infty.$$

Then bounds $p_{\mathrm{l}}^{i'j'}(k)$ and $p_{\mathrm{u}}^{i'j'}(k)$ are given by

$$\begin{aligned}
p_{\mathrm{l}}^{i'j'}(k) &:= \widetilde{p}_{\mathrm{l}}^{i'j'}(k) - \sum_{k'=n_{i'j'}+1}^{\infty} \bar{g}_{i'j'}(k') \left| s_{k'}^{i'j'}(k) \right|, \; k = 0, \ldots, m, \\
p_{\mathrm{l}}^{i'j'}(k) &:= -\sum_{k'=0}^{\infty} \bar{g}_{i'j'}(k') \left| s_{k'}^{i'j'}(k) \right|, \; k = m+1, \ldots, \infty, \\
p_{\mathrm{u}}^{i'j'}(k) &:= \widetilde{p}_{\mathrm{u}}^{i'j'}(k) + \sum_{k'=n_{i'j'}+1}^{\infty} \bar{g}_{i'j'}(k') \left| s_{k'}^{i'j'}(k) \right|, \; k = 0, \ldots, m, \\
p_{\mathrm{u}}^{i'j'}(k) &:= \sum_{k'=0}^{\infty} \bar{g}_{i'j'}(k') \left| s_{k'}^{i'j'}(k) \right|, \; k = m+1, \ldots, \infty,
\end{aligned}$$

where $\widetilde{p}_{\mathrm{u}}^{i'j'}(k)$ and $\widetilde{p}_{\mathrm{l}}^{i'j'}(k)$ are the solutions to finite dimensional linear programming problems which are similar to the ones formulated in Section 4.5.

The pulse response bounds $p_{\mathrm{l}}^{i'j'}(k)$ and $p_{\mathrm{u}}^{i'j'}(k)$ can straightforwardly be applied to derive an upper bound on the ℓ_1-norm of the weighted model error for each transfer function $G_{i',j'}(z)$ separately. Let the model error $\Delta_{i'j'}(z)$ be defined as

$$\Delta_{i'j'}(z) = W_{i'j'}(z) \left(G_{i'j'}(z) - \widehat{G}_{i'j'}(z) \right),$$

for any stable, finite dimensional, LTI nominal model $\widehat{G}_{i'j'}(z)$ and with the weighting $W_{i'j'}(z)$ as defined above. Denote,

$$\widehat{G}_{i'j'}(z) W_{i'j'}(z) = \sum_{k=0}^{\infty} \widehat{p}_{i'j'}(k) z^{-k},$$

and

$$\gamma_{i'j'} := \sum_{k=m+1}^{\infty} \left(p_{\mathrm{u}}^{i'j'}(k) + |\widehat{p}_{i'j'}(k)| \right).$$

Then an upper bound on the ℓ_1-norm of $\Delta_{i'j'}(z)$ is given by

$$\max_{G\in\mathcal{L}} \|\Delta_{i'j'}(z)\|_{\ell_1} \leq \sum_{k=0}^{m} \frac{1}{2}\left(p_{\mathrm{u}}^{i'j'}(k) - p_{\mathrm{l}}^{i'j'}(k) + \left|p_{\mathrm{u}}^{i'j'}(k) + p_{\mathrm{l}}^{i'j'}(k) - 2\widehat{p}_{i'j'}(k)\right|\right) + \gamma_{i'j'}.$$

Similar to the H_∞-case treated in the previous subsection, this provides a structured uncertainty description, consisting of $p'q'$ scalar uncertainty blocks. This information can straightforwardly be used in an ℓ_1-structured robustness analysis or controller synthesis procedure, see Subsection 2.2.3 in Chapter 2.

4.7 Examples

4.7.1 Example 1: A Simple Example

In this subsection some features of the procedure for deterministic uncertainty bounding identification are illustrated by means of a simple example. The example is taken from Wahlberg and Ljung (1992). The continuous-time system used in that paper is transformed to a discrete-time representation (under the zero-order hold assumption), which is given by

$$G_0(z) = \frac{0.1589z^2 + 0.3042z + 0.0323}{z^3 - 1.3219z^2 + 0.4618z - 0.0408}.$$

The input signal $u(t)$ is a uniformly [-1,1] distributed random signal. The noise $e_0(t)$ is a uniformly [-0.05,0.05] distributed random signal. This noise is 5 times as large as the noise used in the example in Wahlberg and Ljung (1992). The system is initiated at $t = 1$ and 200 input-output samples are taken, but the first 100 samples are not used for identification. Only the samples from $t = 101$ till $t = 200$ are used to construct frequency response uncertainty regions. The standard pulse functions are used to construct a basis (see the discussion in Subsection 2.4.3), so $B_k(z) = z^{-k}$. The prior information assumed available, see Section 4.2, is the noise bound $\bar{e}(t) = 0.05$, $\forall t$, the bound on the input signal in the past $\bar{u} = 1$ and the parameter bound $\bar{g}(k) = M\rho^k$, with $M = 2$ and $\rho = \frac{1}{1.2}$, which are conservative values. No cross-covariance noise bounds are used. The truncation value in (4.12) has been chosen $n = 49$.

For 40 frequencies logarithmically distributed between 0.01 and π the frequency response uncertainty regions $\mathcal{P}_m(\omega_j)$ have been calculated for $m = 4$. The result is shown in a Nyquist diagram, Figure 4.2.

The uncertainty regions appear to be slightly smaller than the uncertainty regions found by Wahlberg and Ljung (1992), whereas in the present case the signal-to-noise ratio is much worse (a factor 5)! This clearly demonstrates the good performance of the uncertainty bounding identification procedure developed in this chapter. However, note that the example is not very realistic, due to the high signal-to-noise ratio and the fact that the noise is uniformly distributed.

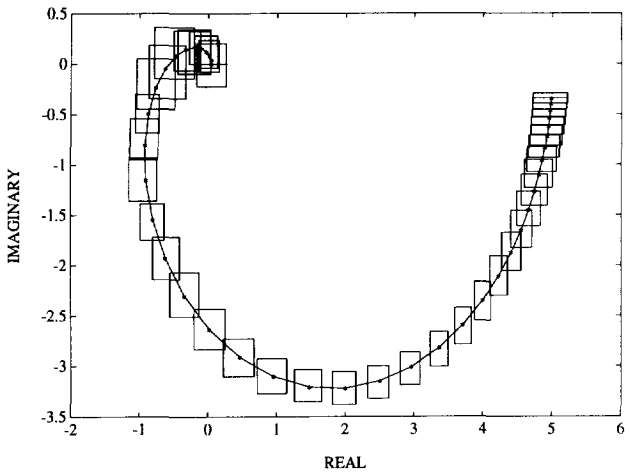

Fig. 4.2: Nyquist diagram identified frequency response uncertainty regions $\mathcal{P}_4(\omega_j)$ (rectangulars) and system's frequency response $G_0(e^{i\omega_j})$ (solid,*).

4.7.2 Example 2: A More Complicated Example

In this subsection it is shown by means of a more complicated example how the uncertainty bounding identification procedure can be used in practice. The resulting frequency response uncertainty regions are correct provided the prior information that is used is correct. However, often the required prior information is not available beforehand. Therefore special attention is paid to the problem of gathering the prior information from data. An attempt is made to develop a reliable procedure to estimate these priors.

Consider the data generating system

$$G_0(z) = \frac{0.21z^4 + 0.35z^3 - 0.12z^2 - 0.11z + 0.23}{z^5 - 2.5z^4 + 3.3z^3 - 2.5z^2 + 1.2z - 0.3}.$$

The output of the system is disturbed by low-pass noise,

$$y(t) = G_0(q)u(t) + e_0(t), \; e_0(t) = \frac{0.2q}{q - 0.5}w_0(t),$$

where $w_0(t)$ is gaussian white noise with variance 1.

The following experiment has been performed. First, 500 samples of the output have been measured while the input is 0. This is a so-called free-run experiment, the output consists of only the noise contribution. This gives insight in the (statistical) properties of the noise, which can be used to choose the noise bounds. After that 1000 samples of the output have been measured while the input has been excited with a Random Binary Sequence, i.e. a signal with amplitude switching between -1 and 1 at

random time instants. In Figure 4.3 a short interval of the measured output is shown, which clearly shows the start of the input excitation.

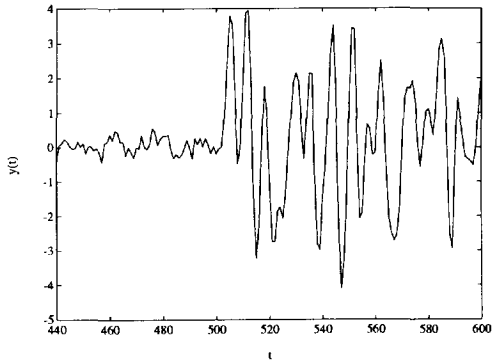

Fig. 4.3: Measured output signal.

It is assumed that there is some rough prior information of the system in the sense that it is assumed to be "close" to a nominal model given by

$$\widehat{G}(z) = \frac{-0.65z^3 + 3.25z^2 - 3.25z + 1.5}{z^3 - 1.79z^2 + 1.53z - 0.42}.$$

This "closeness" is formalized by using the model $\widehat{G}(z)$ as a basis-generating model and specifying bounds for the expansion coefficients $g(k)$, see Appendix 4.B. In first instance these bounds are chosen quite conservative,

$$\bar{g}(k) = M\rho^k, \ k = 0, \ldots, \infty,$$

with $M = 10$ and $\rho = 0.9$. Later on these bounds are tightened using insight obtained from the data.

By design, the initial conditions are 0, so $\bar{u} = 0$. The time-domain bound on the noise, $\bar{e}(t)$, is for all t chosen equal to 1.7652, which is 1.075 times the maximum measured output amplitude during the first 500 samples. The factor 1.075 is more or less ad hoc, added for safety in order to increase the chance that the bound is correct, although this can never be guaranteed without complete noise knowledge. In this case the upper bound appears to be correct for the next 1000 samples, i.e. the noise $e_0(t)$ satisfies the constraints.

Altogether 202 signals $r_p(t)$ have been used for the cross-covariance constraints on the noise. Part of them were chosen as delayed inputs,

$$r_p(t) = u(t - p + 1), \ p = 1, \ldots, 101,$$

and part of them as delayed filtered inputs,

$$r_p(t) = w(t - p + 102), \ w(t) = \widehat{G}(q)u(t), \ p = 102, \ldots, 202,$$

where $\widehat{G}(q)$ is the nominal model given above. The cross-covariance bounds on the noise, $\bar{c}(p)$, were determined with the statistical estimator given in Section 3.3, using the 500 samples of free-run data. They were chosen as 3.5σ-bounds corresponding to a normal distribution. Consequently, each cross-covariance constraint is correct with probability 0.9995 and all cross-covariance constraints together with at least probability $1 - 202(1 - 0.9995) = 0.9$. The resulting bounds are $\bar{c}(p) = 1.6008$, $p = 1, \ldots, 101$ and $\bar{c}(p) = 7.5935$, $p = 102, \ldots, 202$. No absolute guarantees can be given that the cross-covariance constraints are correct, only a probability. In this case they appear to be correct for the specific noise realization in the measured data, i.e. the noise $e_0(t)$ satisfies the constraints.

Using the prior information established above, and the measurement data, the set $\mathcal{L}$ has been calculated, where n has been chosen equal to 45. Next, the maximum and minimum of each $g(k)$ consistent with data and prior information have been calculated by solving linear programming problems similar to the ones described in the Sections 4.4 and 4.5,

$$g_u(k) = \max_{G(z)\in\mathcal{L}} g(k), \; g_l(k) = \min_{G(z)\in\mathcal{L}} g(k), \; k = 0, \ldots, 45.$$

The results are depicted in Figure 4.4.

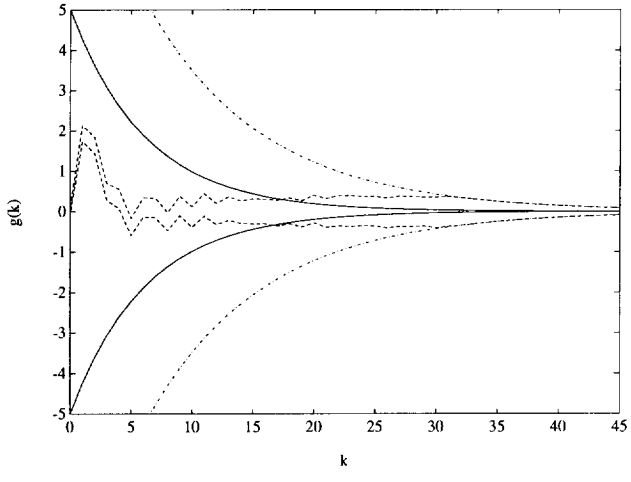

Fig. 4.4: Prior parameter bounds $\pm\bar{g}(k)$ (dash-dotted), identified bounds $g_l(k)$, $g_u(k)$ (dashed) and new prior bounds $\pm\bar{g}(k)$ (solid).

In this figure the bounds $\pm\bar{g}(k)$ are shown, as well as the calculated minima and maxima $g_l(k)$, $g_u(k)$. Next, new and less conservative "prior" bounds for $g(k)$ are established by choosing new values for M and ρ. These are chosen tight and such that the new bound $\bar{g}(k)$ is consistent with $g_u(k)$ and $g_l(k)$ for $k = 0, \ldots, 17$, but not necessarily consistent with $g_u(k)$ and $g_l(k)$ for $k > 17$. This is an ad-hoc choice, based

on the observation that the bounds $g_u(k)$, $g_l(k)$ for large k are much influenced by noise, and therefore often unnecessarily large. It is also based on the assumption that $g_0(k)$ shows a kind of smooth behaviour, also for $k > 45$, with an exponential decay rate which already becomes visible for low values of k. Then new bounds are likely to be correct as long as they are consistent with the first part. This will generally be the case if the basis-generating model $\widehat{G}(z)$ is chosen suitably, i.e. containing dynamics similar to those of $G_0(z)$. However, no guarantees, or even a probability, can be given that the new bounds are correct. The resulting new bounds, corresponding to $M = 5$ and $\rho = 0.85$, are also depicted in Figure 4.4. In this case they appear to be correct, i.e. they are satisfied by the system $G_0(z)$.

With the new prior bounds $\bar{g}(k)$, but the noise bounds $\bar{e}(t)$ and $\bar{c}(p)$ unchanged, a new set of linear constraints $\mathcal{L}$ has been calculated, with $n = 45$ again. With the procedure of Section 4.4 frequency response uncertainty regions are calculated for 32 frequencies between 0 and π, with $m = 8$, yielding polytopes $\mathcal{P}_8(\omega_j)$ in the complex plane. These uncertainty regions are depicted in a Nyquist diagram, Figure 4.5, together with the system's frequency response $G_0(e^{i\omega_j})$.

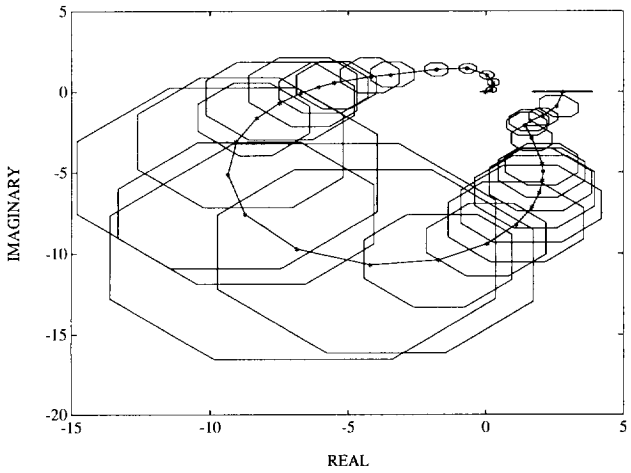

Fig. 4.5: Nyquist diagram identified frequency response uncertainty regions $\mathcal{P}_8(\omega_j)$ (polytopes) and system's frequency response $G_0(e^{i\omega_j})$ (solid,*).

Moreover, with the procedure of Section 4.5 uncertainty regions are calculated for the first 100 pulse response parameters of the system $G_0(z)$ (the weighting $W(z)$ is chosen equal to 1). These pulse response uncertainty regions are depicted in Figure 4.6, together with the system's pulse response $p_0(k)$.

Due to the fact that the prior information used in the identification procedure is correct, the system is within the identified uncertainty bounds, both in Figure 4.5 and in Figure 4.6. Of course in practice one does not know if this is the case, one can only say that the uncertainty regions are correct with high probability.

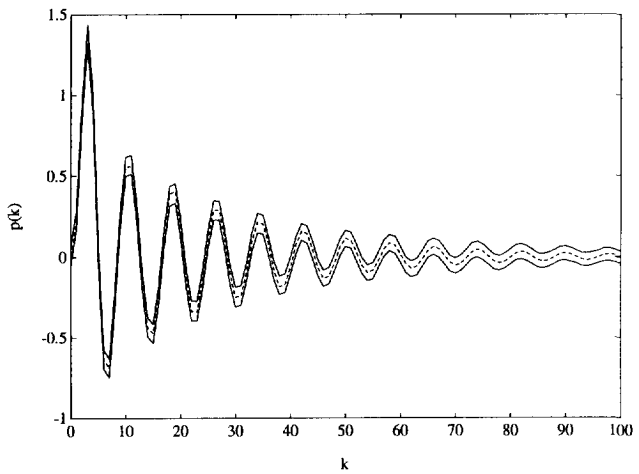

Fig. 4.6: Identified pulse response upper and lower bounds $p_{\mathrm{u}}(k)$, $p_{\mathrm{l}}(k)$ (solid) and system's pulse response $p_0(k)$ (dashed).

The identified frequency response uncertainty regions can be used to establish an upper bound on the H_∞-norm of the model error, using the expressions derived in Subsection 4.4.3. Suppose the additive error of the model $\widehat{G}(z)$ given above, is of interest. Figure 4.7 shows $\delta_{\mathrm{u}}(\omega_j)$ defined in Proposition 4.4.6, corresponding to $W(z) = 1$. Also, the model error $\left|\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right|$ is depicted in this figure. Neglecting the

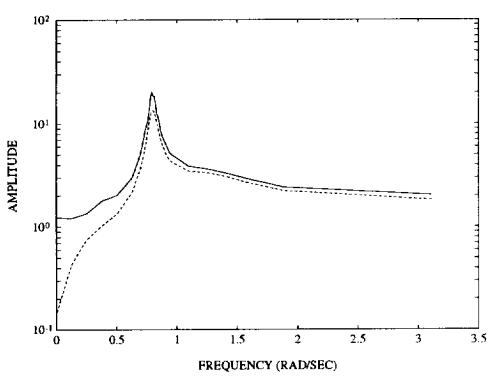

Fig. 4.7: Additive frequency response model error bound $\delta_{\mathrm{u}}(\omega_j)$ (solid) and additive model error $\left|\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right|$ (dashed).

contribution of the interpolation this yields an upper bound on the H_∞-norm of the additive model error equal to 20.3. This should be compared to the H_∞-norm of the model error, $\left\|\widehat{G}(z) - G_0(z)\right\|_\infty$, which equals 13.7. If desired, the influence of the interpolation between the discrete frequencies can be taken into account properly by applying Theorem 4.4.8.

Similarly, on the basis of the pulse response uncertainty regions shown in Figure 4.6 and the expressions derived in Subsection 4.5.3, an upper bound on the ℓ_1-norm of the additive model error can be derived. Figure 4.8 depicts the pulse response error bound $\max[p_u(k) - \widehat{p}(k), \widehat{p}(k) - p_l(k)]$, where $\widehat{p}(k)$ are the pulse response parameters of the model $\widehat{G}(z)$. Also, the pulse response error $|\widehat{p}(k) - p_0(k)|$ is shown, where $p_0(k)$ are the pulse response parameters of the system $G_0(z)$. This yields an upper bound on

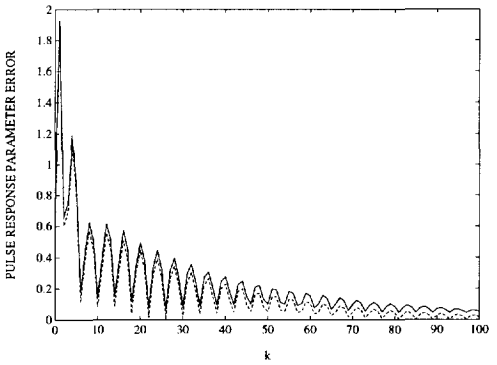

Fig. 4.8: Additive pulse response model error bound $\max[p_u(k) - \widehat{p}(k), \widehat{p}(k) - p_l(k)]$ (solid) and additive model error $|\widehat{p}(k) - p_0(k)|$ (dashed).

the ℓ_1-norm of the model error equal to 23.9. This should be compared to the ℓ_1-norm of the model error, $\left\|\widehat{G}(z) - G_0(z)\right\|_{\ell_1}$, which equals 18.9.

4.8 Discussion

In this chapter an identification procedure has been developed which yields deterministic uncertainty regions for the frequency or pulse response of some stable, LTI system. The procedure is a parametric one which directly operates on time-domain data, without first transforming them into the frequency domain. Important features of the identification procedure are:

- Essentially the identification approach is worst-case deterministic. Deterministic uncertainty regions and model error bounds are calculated based on data and deterministic prior information. The results are non-asymptotically correct.

- The actual computations can be performed quite efficiently. Use is made of linear programming, which can be solved fast and accurately. No nonlinear optimizations are involved and, consequently, there is no problem with local optima.
- The procedure is very flexible in the sense that various kinds of noise bounds can be incorporated: time-domain bounds on the noise, cross-covariance bounds on the noise, and even frequency-domain bounds on the noise if desired.
- There are no restrictions on the input signal, it need, for example, not be periodic. It is even not necessary that the input is generated in open loop.
- No system order assumption is made.
- The procedure is applicable to MIMO systems.
- Rough prior knowledge about the system, or more specifically pole-locations, can be incorporated by using generalized orthonormal basis functions.
- Unknown initial conditions are properly taken into account.
- Under certain asymptotic conditions the resulting frequency response uncertainty regions are non-conservative in the sense that, using the information available, no smaller regions can be identified for each frequency separately. For that reason the resulting upper bound on the H_∞-norm of the model error, for which expressions have been derived in Section 4.4, is non-conservative as well. The same holds for the bounds on each pulse response parameter separately, but unfortunately this does not imply non-conservativeness of the upper bound on the ℓ_1-norm of the model error.
- It is straightforward to incorporate more prior knowledge about the system in the set of linear constraints. If, for example, bounds for the static gain are known, inequalities of the form

$$s_l \leq \sum_{k=0}^{\infty} g(k) B_k(1) \leq s_u$$

  can be included in the linear programming problems.

On the other hand, some drawbacks of the uncertainty bounding identification procedure developed in this chapter, are:

- Basically the identification is deterministic, yielding frequency or pulse response uncertainty regions which are correct with probability 1. However, this point of view can no longer be taken, if it is considered that the prior information about system and noise has to be estimated from measurement data, as in the example in Subsection 4.7.2. This will be the case for any *"hard"* uncertainty bounding identification approach. It is not hard anymore, but *"soft"* or probabilistic, if the system is really black-box, and the engineer must take a so-called *"leap of faith"* to determine the priors.

- Although a linear programming problem can be solved fast and accurately, the identification procedure requires the solution of *many* LP problems. This can still be carried out with a reasonable computational effort. However, it means that, for example, on-line application of the procedure is, for the time being, infeasible.

- Basically the procedure is highly sensitive to the prior information and not robust to errors in the priors. For example, if there is an undetected noise outlier, no statement can be made about correctness of the resulting frequency or pulse response uncertainty regions. In many cases the identification results will still be correct in case of small error in the priors, simply due to the fact that the identified uncertainty regions are large enough to cope with these deficiencies. However, there are no guarantees that this will be the case. On the other hand, if the errors in the priors are large, this can often be detected by the fact that the linear programming problems do not have a feasible solution anymore.

4.A Proofs

Proof of Theorem 4.4.2

First, it is noticed that the constraints with respect to $g(0), \ldots, g(n)$ in the set $\mathcal{L}$ are independent of those with respect to $g(n+1), \ldots, g(\infty)$. Consequently,

$$\begin{aligned}
&\max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m} G(e^{i\omega_j})\right) = \\
&= \max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}\left(\widetilde{G}(e^{i\omega_j}) + \bar{G}(e^{i\omega_j})\right)\right) = \\
&= \max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}\widetilde{G}(e^{i\omega_j})\right) + \max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}\bar{G}(e^{i\omega_j})\right) = \\
&= \widetilde{\mu}_k(\omega_j) + \bar{\mu}_k(\omega_j) = \mu_k(\omega_j),
\end{aligned}$$

since

$$\begin{aligned}
&\max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}\widetilde{G}(e^{i\omega_j})\right) = \\
&= \max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}\sum_{k'=0}^{n} g(k')B_{k'}(e^{i\omega_j})\right) = \\
&= \max_{G(z)\in\mathcal{L}} \sum_{k'=0}^{n} g(k')\mathrm{Re}\left(e^{i2\pi k/m}B_{k'}(e^{i\omega_j})\right) = \widetilde{\mu}_k(\omega_j),
\end{aligned}$$

as defined by the solution of the linear programming problem given above, and

$$\max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}\bar{G}(e^{i\omega_j})\right) =$$

$$= \max_{G(z)\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m} \sum_{k'=n+1}^{\infty} g(k')B_{k'}(e^{i\omega_j})\right) =$$

$$= \max_{G(z)\in\mathcal{L}} \sum_{k'=n+1}^{\infty} g(k')\mathrm{Re}\left(e^{i2\pi k/m}B_{k'}(e^{i\omega_j})\right) = \bar{\mu}_k(\omega_j),$$

as defined above.

Proof of Lemma 4.4.4

For any LTI transfer function $G(z)$ the following relations hold,

$$\left|\frac{\mathrm{d}\,G(e^{i\omega})}{\mathrm{d}\,\omega}\right| = \left|\frac{\mathrm{d}\,G(e^{i\omega})}{\mathrm{d}\,e^{i\omega}}\frac{\mathrm{d}\,e^{i\omega}}{\mathrm{d}\,\omega}\right| = \left|\frac{\mathrm{d}\,G(e^{i\omega})}{\mathrm{d}\,e^{i\omega}}ie^{i\omega}\right| = \left|\frac{\mathrm{d}\,G(e^{i\omega})}{\mathrm{d}\,e^{i\omega}}\right| \leq \left\|\frac{\mathrm{d}\,G(z)}{\mathrm{d}\,z}\right\|_{\infty}, \quad \forall\omega,$$

where the given bound is finite if $G(z)$ is finite dimensional and stable. The proof of the lemma is completed by noting that,

$$\left\|\frac{\mathrm{d}\,G(z)}{\mathrm{d}\,z}\right\|_{\infty} = \left\|\sum_{k=0}^{\infty} g(k)\frac{\mathrm{d}\,B_k(z)}{\mathrm{d}\,z}\right\|_{\infty} \leq \beta_1.$$

Proof of Proposition 4.4.6

Part (i) holds by definition of $\delta_{\mathrm{u}}(\omega_j)$ and $\mathcal{S}$. Both the upper bound $\delta_{\mathrm{u}}(\omega_j)$ and the set $\mathcal{S}$ in (4.20) are namely based on the polytopes $\mathcal{P}_m(\omega_j)$.

Next, let ϕ_j be the argument of $W(e^{i\omega_j})$, i.e. $W(e^{i\omega_j}) = \left|W(e^{i\omega_j})\right| e^{i\phi_j}$. Also, define $\psi_{jk} := 2\pi k/m - \phi_j$. Then,

$$\max_{\Delta=W(G-\widehat{G}),G\in\mathcal{L}} \mathrm{Re}\left(e^{i\psi_{jk}}\Delta(e^{i\omega_j})\right) =$$

$$= \max_{G\in\mathcal{L}} \mathrm{Re}\left(e^{i\psi_{jk}}W(e^{i\omega_j})\left(G(e^{i\omega_j}) - \widehat{G}(e^{i\omega_j})\right)\right) =$$

$$= \max_{G\in\mathcal{L}} \mathrm{Re}\left(\left|W(e^{i\omega_j})\right| e^{i(\psi_{jk}+\phi_j)}G(e^{i\omega_j})\right) - \mathrm{Re}\left(\left|W(e^{i\omega_j})\right| e^{i(\psi_{jk}+\phi_j)}\widehat{G}(e^{i\omega_j})\right) =$$

$$= \left|W(e^{i\omega_j})\right|\left(\max_{G\in\mathcal{L}} \mathrm{Re}\left(e^{i2\pi k/m}G(e^{i\omega_j})\right) - \mathrm{Re}\left(e^{i2\pi k/m}\widehat{G}(e^{i\omega_j})\right)\right) =$$

$$= \left|W(e^{i\omega_j})\right|\left(\mu_k(\omega_j) - \mathrm{Re}\left(e^{i2\pi k/m}\widehat{G}(e^{i\omega_j})\right)\right).$$

Consequently,

$$\delta_{\mathrm{l}}(\omega_j) = \max_{k=1,\ldots,m} \max_{\Delta=W(G-\widehat{G}),G\in\mathcal{L}} \mathrm{Re}\left(e^{i\psi_{jk}}\Delta(e^{i\omega_j})\right) =$$

$$= \max_{\Delta=W(G-\widehat{G}),G\in\mathcal{L}} \max_{k=1,\ldots,m} \mathrm{Re}\left(e^{i\psi_{jk}}\Delta(e^{i\omega_j})\right).$$

Using Lemma 3.4.1 this yields the first inequality in part (ii). The second inequality follows from the fact that $v_k(\omega_j)$ are the vertices of the convex polytopes $\mathcal{P}_m(\omega_j)$. The bound given by this second inequality is the tightest bound that can be established on the basis of the region $\mathcal{P}_m(\omega_j)$. Noting that $\delta_1(\omega_j)/\cos(\pi/m)$ is also an upper bound, according to Lemma 3.4.1, it is concluded that the third inequality must hold. This completes the proof of part (ii).

Next, the statement in part (iii) is proven by noting that the left-hand side expression in part (ii) converges to the right-hand side expression for $m \to \infty$. Note that this result also follows straightforwardly from Proposition 4.4.3.

Proof of Lemma 4.4.7

First, it is noted that for any system $G(z)$ in the set $\mathcal{S}$,

$$\left|\frac{\mathrm{d}\,|G(e^{i\omega})|}{\mathrm{d}\,\omega}\right| \leq \left|\frac{\mathrm{d}\,G(e^{i\omega})}{\mathrm{d}\,\omega}\right| \leq \beta_1, \ \forall \omega \in [0,\pi].$$

Next, $\forall \omega$, ω_j, linear interpolation gives the following bound,

$$|G(e^{i\omega})| \leq |G(e^{i\omega_j})| + |\omega_j - \omega|\beta_1 \leq \max_{k=1,\ldots,m} |v_k(\omega_j)| + |\omega_j - \omega|\beta_1.$$

Performing this interpolation for every $\omega \in [\omega_j - \lambda_j/2, \omega_j + \lambda_j/2]$ and for every j gives the desired result.

Proof of Theorem 4.4.8

The first inequality follows directly from the fact that any system $G(z)$ which is an element of the set $\mathcal{L}$, is also an element of the set $\mathcal{S}$. Next, it is noted that, $\forall \omega \in [0,\pi]$,

$$\left|\frac{\mathrm{d}\,|\Delta(e^{i\omega})|}{\mathrm{d}\,\omega}\right| \leq \left|\frac{\mathrm{d}\,\Delta(e^{i\omega})}{\mathrm{d}\,\omega}\right| = \left|\frac{\mathrm{d}\,W(e^{i\omega})\left(G(e^{i\omega}) - \widehat{G}(e^{i\omega})\right)}{\mathrm{d}\,\omega}\right| \leq$$

$$\leq \left|\frac{\mathrm{d}\,W(e^{i\omega})\widehat{G}(e^{i\omega})}{\mathrm{d}\,\omega}\right| + \left|\frac{\mathrm{d}\,W(e^{i\omega})}{\mathrm{d}\,\omega}\right| |G(e^{i\omega})| + |W(e^{i\omega})| \left|\frac{\mathrm{d}\,G(e^{i\omega})}{\mathrm{d}\,\omega}\right| \leq \beta.$$

Consequently, $\forall \omega$, ω_j, linear interpolation gives the following bound,

$$|\Delta(e^{i\omega})| \leq |\Delta(e^{i\omega_j})| + |\omega_j - \omega|\beta.$$

Performing this interpolation for every $\omega \in [\omega_j - \lambda_j/2, \omega_j + \lambda_j/2]$ and for every j yields the H_∞-model error bound.

If Ω becomes dense in the interval $[0,\pi]$, the interpolation contribution $\lambda_j\beta/2$ vanishes, i.e. the upper bound on the H_∞-norm of Δ converges to $\max_{j=1,\ldots,l} \delta_u(\omega_j)$. With part (i) of Proposition 4.4.6 this proves that the second inequality then becomes equality. If in addition $m \to \infty$, application of part (iii) of Proposition 4.4.6 proves that the first inequality becomes equality as well, which completes the proof.

Proof of Theorem 4.5.2

Noting that the constraints with respect to $g(0), \ldots, g(n)$ in the set $\mathcal{L}$ are independent of those with respect to $g(n+1), \ldots, g(\infty)$, gives

$$\min_{g(k)\in\mathcal{L}_k} p(k) = \min_{G(z)\in\mathcal{L}} (\tilde{p}(k) + \bar{p}(k)) = \min_{G(z)\in\mathcal{L}} \tilde{p}(k) + \min_{G(z)\in\mathcal{L}} \bar{p}(k) =$$

$$= \min_{G(z)\in\mathcal{L}} \sum_{k'=0}^{n} g(k')s_{k'}(k) + \min_{G(z)\in\mathcal{L}} \sum_{k'=n+1}^{\infty} g(k')s_{k'}(k) = \tilde{p}_{\mathrm{l}}(k) + \bar{p}_{\mathrm{l}}(k) = p_{\mathrm{l}}(k), \ \forall k,$$

with $\tilde{p}_{\mathrm{l}}(k)$ as defined by the solution of the linear programming problem given above, and $\bar{p}_{\mathrm{l}}(k)$ defined in (4.30). This proves the equality-statement for $k \leq m$. The inequality-statement for $k > m$ is proven by noting that it is based on a subset of the complete set of constraints in $\mathcal{L}$, namely the upper bounds on $g(k)$. The statements with respect to $p_{\mathrm{u}}(k)$ can be proven completely similarly, replacing the minimizations by maximizations.

Proof of Theorem 4.5.4

$$\max_{\Delta=W(G-\widehat{G}),G\in\mathcal{L}} \|\Delta(z)\|_{\ell_1} = \max_{G\in\mathcal{L}} \left\| \sum_{k=0}^{\infty} p(k)z^{-k} - \sum_{k=0}^{\infty} \widehat{p}(k)z^{-k} \right\|_{\ell_1} =$$

$$= \max_{G\in\mathcal{L}} \sum_{k=0}^{\infty} |p(k) - \widehat{p}(k)| \leq \sum_{k=0}^{\infty} \max_{G\in\mathcal{L}} |p(k) - \widehat{p}(k)| =$$

$$= \sum_{k=0}^{\infty} \max_{p(k)\in\mathcal{L}_k} |p(k) - \widehat{p}(k)| \leq \sum_{k=0}^{\infty} \max\{p_{\mathrm{u}}(k) - \widehat{p}(k), \widehat{p}(k) - p_{\mathrm{l}}(k)\} =$$

$$= \sum_{k=0}^{\infty} \left(\frac{1}{2}(p_{\mathrm{u}}(k) - p_{\mathrm{l}}(k)) + \left| \frac{1}{2}(p_{\mathrm{l}}(k) + p_{\mathrm{u}}(k)) - \widehat{p}(k) \right| \right) =$$

$$= \sum_{k=0}^{m} \left(\frac{1}{2}(p_{\mathrm{u}}(k) - p_{\mathrm{l}}(k)) + \left| \frac{1}{2}(p_{\mathrm{l}}(k) + p_{\mathrm{u}}(k)) - \widehat{p}(k) \right| \right) +$$

$$+ \sum_{k=m+1}^{\infty} (p_{\mathrm{u}}(k) + |\widehat{p}(k)|),$$

where the last equality follows from the fact that $p_{\mathrm{u}}(k) = -p_{\mathrm{l}}(k)$, $\forall k > m$. The proof is completed by noting that

$$\max_{\Delta=W(G-\widehat{G}),G\in\mathcal{S}} \|\Delta(z)\|_{\ell_1} = \sum_{k=0}^{\infty} \max\{p_{\mathrm{u}}(k) - \widehat{p}(k), \widehat{p}(k) - p_{\mathrm{l}}(k)\}.$$

4.B Orthonormal Basis Functions

In this appendix some theory and results are presented concerning orthonormal basis functions. Generally in this thesis systems are assumed to belong to the set $\mathcal{A}$, i.e. the

set of discrete-time, ℓ_∞-stable, LTI, causal systems. The theory of orthonormal basis functions is generally developed with respect to the set $\mathcal{H}_2$, i.e. the space of discrete-time, LTI, causal systems, which are square integrable on the unit circle. This space is used because of the fact that it possesses an inner product. In discrete-time the set $\mathcal{H}_2$ is equal to the set $\mathcal{H}_\infty$, which is the set of discrete-time, ℓ_2-stable, LTI, causal systems. Consequently, as $\mathcal{A} \subset \mathcal{H}_\infty$, also $\mathcal{A} \subset \mathcal{H}_2$, and it appears that the theory of orthonormal basis functions is just slightly more general than needed in this thesis.

Consider a SISO system $G(z) \in \mathcal{H}_2$. A common representation of $G(z)$ is in terms of its Laurent expansion around $z = \infty$,

$$G(z) = \sum_{k=0}^{\infty} g(k) z^{-k}, \tag{4.33}$$

with $g(k)$ the Markov or pulse response parameters. Here the set $\{z^0, z^{-1}, z^{-2}, \ldots\}$ is a set of orthonormal basis functions with respect to the inner product in $\mathcal{H}_2$. From the work of Heuberger (Heuberger and Bosgra, 1990; Heuberger *et al.*, 1994) it is known that more general system expansions exist, in terms of generalized orthonormal basis functions,

$$G(z) = D + z^{-1} \sum_{k=0}^{\infty} L_k V_k(z), \tag{4.34}$$

where D, $\{L_k\}_{k=0,\ldots,\infty}$ are the orthonormal expansion coefficients of $G(z)$. Here D is a scalar and the coefficients L_k are vectors of dimension $1 \times n_b$. The sequence $\{V_k(z)\}_{k=0,\ldots,\infty}$ is a sequence of generating transfer functions. In Heuberger *et al.* (1994) it is shown that any finite dimensional and stable model $H(z)$, with state-space representation (A_h, B_h, C_h, D_h), gives rise to a set of orthonormal basis functions $\{V_k(z)\}$. Any system $G(z) \in \mathcal{H}_2$ can be represented in this basis with unique coefficients D, $\{L_k\}$. The basis functions $V_k(z)$ are given by

$$V_k(z) = V_0(z) \left(V_{\text{ap}}(z)\right)^k, \quad k = 0, \ldots, \infty,$$

for some $n_b \times 1$ stable transfer function $V_0(z)$ and inner (stable, all-pass) scalar transfer function $V_{\text{ap}}(z)$, so $V_{\text{ap}}(z)V_{\text{ap}}(z^{-1}) = 1$. It is not discussed here how to calculate $\{V_k(z)\}$ using a given basis generating model (A_h, B_h, C_h, D_h). For this the reader is referred to Heuberger *et al.* (1994).

The expansion in pulse functions (4.33) is a special case of (4.34) for $V_k(z) = z^{-k}$, generated by $H(z)$ with state-space representation $(0, 1, 1, 0)$. Moreover, the well-known Laguerre polynomials (Szegö, 1975) are a special case as well,

$$V_k(z) = \sqrt{1 - a^2}\, z \frac{(1 - az)^k}{(z - a)^{k+1}},$$

generated by $H(z)$ with state-space representation $(a, 1, 1, 0)$. The description in terms of generalized system-based orthonormal basis functions has the big advantage that it offers the possibility to give an arbitrarily good, linear-in-the-parameters approximation of any $G(z) \in \mathcal{H}_2$ with few parameters, by means of a suitable choice of basis generating model. This is due to the following two properties of the basis functions.

Fact 4.B.1 (Heuberger *et al.*, 1994) *If the n_bth order system $H(z)$, with state-space representation (A_h, B_h, C_h, D_h), is used to generate $\{V_k(z)\}$, then only 2 vector coefficients are needed to expand the system $H(z)$ in its own basis,*

$$H(z) = D_h + z^{-1} \sum_{k=0}^{\infty} L_{h,k} V_k(z), \; L_{h,0} = C_h, \; L_{h,k} = 0, \; k > 0.$$

Fact 4.B.2 (Heuberger *et al.*, 1994) *Denote the series-expansion of any other n_gth order dynamical system $G(z)$, with state-space representation (A_g, B_g, C_g, D_g), by*

$$G(z) = D + z^{-1} \sum_{k=0}^{\infty} L_k V_k(z).$$

Let A_g have eigenvalues μ_i, $i = 1, \ldots, n_g$, and A_h have eigenvalues ν_j, $j = 1, \ldots, n_b$. Define

$$\lambda := \max_{i=1,\ldots,n_g} \prod_{j=1}^{n_b} \left| \frac{\mu_i - \nu_j}{1 - \mu_i \nu_j} \right|.$$

Then for any $\rho \in (\lambda, 1)$ $\exists$ finite $M \in \mathbb{R}^{n_b}$ such that

$$-M_{i'} \rho^k \leq L_k^{i'} \leq M_{i'} \rho^k,$$

where the sub- or superscript i' denotes the i'th vector-element.

This means that the coefficients L_k converge rapidly to 0 if $G(z)$ is close to $H(z)$, in the sense that their poles need to be close. In fact the use of basis functions allows for utilizing (approximate) prior knowledge about pole locations.

Another important property of the basis functions is that, due to the orthonormality, they lead in general to well-scaled, numerically well-conditioned (identification) problems.

Some norm properties of the basis functions $V_k(z)$ are given. From the fact that $V_{\text{ap}}(z)$ is all-pass it follows that

Fact 4.B.3

$$\left| V_k^{i'}(e^{i\omega}) \right| = \left| V_0^{i'}(e^{i\omega}) \right|, \; \forall \omega, \; i' = 1, \ldots, n_b \Rightarrow \left\| V_k^{i'}(z) \right\|_\infty = \left\| V_0^{i'}(z) \right\|_\infty.$$

Also, extensive simulations have motivated the following conjecture.

Conjecture 4.B.4 *Consider the pulse response representation of any basis function, generated with the procedure described in Heuberger et al. (1994),*

$$V_k^{i'}(z) = \sum_{k'=0}^{\infty} v_k^{i'}(k') z^{-k'},$$

then it is conjectured that for any $p \geq 0$,

$$\sum_{k'=0}^{\infty} (k')^p \left| v_k^{i'}(k') \right| \leq c_1^{i'} + c_2^{i'} k^{p+0.5},$$

for some finite constants $c_1^{i'}$, $c_2^{i'}$ *which are independent of* k, *but possibly depend on* p. *This implies that*

$$\left\| V_k^{i'}(z) \right\|_{\ell_1} \leq c_1^{i'} + c_2^{i'} \sqrt{k}.$$

This conjecture is used in Appendix 4.C.

Elsewhere in this thesis a different notation is used for the basis functions, $B_k(z)$ instead of $V_k(z)$, see (4.1) and (4.2). This is done for the reason of notational simplicity, as $B_k(z)$ are scalars, whereas $V_k(z)$ are vectors. It is straightforward to obtain $B_k(z)$, $g(k)$ from given $V_k(z)$, L_k by renumbering as follows,

$$B_0(z) = 1,\ B_{n_b k + i'}(z) = z^{-1} V_k^{i'}(z),\ g(0) = D,\ g(n_b k + i') = L_k^{i'}.$$

4.C Computational Aspects

In this appendix it is discussed how various quantities, which need to be evaluated in the identification procedure and which consist of infinite sums, can be calculated. The idea always is to truncate the infinite sums and give a bound for the truncation error which vanishes for increasing truncation horizon. For this purpose use is made of the fact that the bound $\bar{g}(k)$ shows exponential decay rate for k larger than some value k^*, see (4.6). It is assumed that the transfer functions $\{B_k(z)\}$ as given in (4.1) are orthonormal basis functions of the kind described in Appendix 4.B.

Also, use is made of the fact that computation of the ℓ_1-norm and H_∞-norm of a given transfer function can be performed to an arbitrary accuracy, see Balakrishnan and Boyd (1992) and Boyd and Balakrishnan (1990). Moreover, use is made of the fact that summations of the form

$$\sum_{k=\tilde{k}+1}^{\infty} \rho^k k^p,\ p = 0.5,\ 1.5,\ 0 < \rho < 1$$

can be calculated to an arbitrary accuracy, or at least upper bounded, for example by calculating the upper bound

$$\sum_{k=\tilde{k}+1}^{\tilde{k}} \rho^k k^p + \frac{1}{(\tilde{k}+1)^{0.5}} \sum_{k=\tilde{k}+1}^{\infty} \rho^k k^{p+0.5},$$

where exact solutions can be given for the second part, which vanishes if $\tilde{k} \to \infty$.

Computation of $\bar{a}(t) + \bar{b}(t)$

Consider $\bar{a}(t)$ defined in (4.13) and $\bar{b}(t)$ defined in (4.14). As follows from (4.15) the value $\bar{a}(t) + \bar{b}(t)$ is of interest. Choose some truncation value $\tilde{k} \geq \max\{k^*, n\}$ and calculate

$$\hat{a}(t) = \sum_{k=n+1}^{\tilde{k}} \bar{g}(k) \left| \sum_{k'=0}^{t-1} b_k(k') u(t-k') \right|,$$

$$\widehat{b}(t) = \bar{u} \sum_{k=0}^{\bar{k}} \bar{g}(k) \sum_{k'=t}^{\infty} |b_k(k')| = \bar{u} \sum_{k=0}^{\bar{k}} \bar{g}(k) \left(\|B_k(z)\|_{\ell_1} - \sum_{k'=0}^{t-1} |b_k(k')| \right),$$

$$\delta = u_m \sum_{k=\bar{k}+1}^{\infty} M\rho^k(c_1 + c_2\sqrt{k}).$$

Here u_m is defined as $u_m = \max\{\bar{u}, |u(1)|, \ldots, |u(N)|\}$, and c_1 and c_2 are chosen such that

$$\|B_k(z)\|_{\ell_1} \leq c_1 + c_2\sqrt{k}, \ \forall\, k, \tag{4.35}$$

which is possible according to Conjecture 4.B.4. Since

$$\begin{aligned} \delta \ &\geq \ u_m \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \|B_k(z)\|_{\ell_1} \geq \\ &\geq \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \sum_{k'=0}^{t-1} |b_k(k')u(t-k')| + \bar{u} \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \sum_{k'=t}^{\infty} |b_k(k')|, \ \forall\, t, \end{aligned}$$

this gives

$$\widehat{a}(t) + \widehat{b}(t) \leq \bar{a}(t) + \bar{b}(t) \leq \widehat{a}(t) + \widehat{b}(t) + \delta.$$

Consequently, a conservative value for $\bar{a}(t) + \bar{b}(t)$ is given by $\widehat{a}(t) + \widehat{b}(t) + \delta$, where $\delta \to 0$ if $\bar{k} \to \infty$.

Computation of $\bar{d}(p) + \bar{f}(p)$

Consider $\bar{d}(p)$ defined in (4.16) and $\bar{f}(p)$ defined in (4.17). As follows from (4.18) the value $\bar{d}(p) + \bar{f}(p)$ is of interest. Choose truncation values $\bar{k} \geq \max\{k^*, n\}$ and $\bar{t}' \geq 0$ and calculate

$$\widehat{d}(p) = \sum_{k=n+1}^{\bar{k}} \bar{g}(k) \left| \sum_{t=t_s}^{N} r_p(t) \sum_{k'=0}^{t-1} b_k(k')u(t-k') \right|,$$

$$\widehat{f}(p) = \sum_{k=0}^{\bar{k}} \bar{g}(k) \sum_{t'=0}^{\bar{t}'} \left| \sum_{t=t_s}^{N} r_p(t) b_k(t+t') \right| \bar{u},$$

$$\delta_1(p) = \bar{u} \sum_{k=0}^{\bar{k}} \bar{g}(k) \sum_{t=t_s}^{N} |r_p(t)| \left(\|B_k(z)\|_{\ell_1} - \sum_{k'=0}^{\bar{t}'+t} |b_k(k')| \right),$$

$$\delta_2(p) = u_m \sum_{t=t_s}^{N} |r_p(t)| \sum_{k=\bar{k}+1}^{\infty} M\rho^k(c_1 + c_2\sqrt{k}).$$

Here u_m is defined as $u_m = \max\{\bar{u}, |u(1)|, \ldots, |u(N)|\}$, and c_1 and c_2 are chosen such that (4.35) holds. Since

$$\delta_1(p) = \bar{u} \sum_{k=0}^{\bar{k}} \bar{g}(k) \sum_{t=t_s}^{N} |r_p(t)| \sum_{t'=\bar{t}'+1}^{\infty} |b_k(t+t')| \geq \bar{u} \sum_{k=0}^{\bar{k}} \bar{g}(k) \sum_{t'=\bar{t}'+1}^{\infty} \left| \sum_{t=t_s}^{N} r_p(t) b_k(t+t') \right|$$

and

$$\delta_2(p) \geq u_m \sum_{t=t_s}^{N} |r_p(t)| \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \|B_k(z)\|_{\ell_1} =$$

$$= u_m \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \sum_{t=t_s}^{N} |r_p(t)| \sum_{k'=0}^{t-1} |b_k(k')| + u_m \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \sum_{t=t_s}^{N} |r_p(t)| \sum_{k'=t}^{\infty} |b_k(k')| \geq$$

$$\geq \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \left| \sum_{t=t_s}^{N} r_p(t) \sum_{k'=0}^{t-1} b_k(k') u(t-k') \right| + \bar{u} \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \sum_{t'=0}^{\infty} \left| \sum_{t=t_s}^{N} r_p(t) b_k(t+t') \right|,$$

this gives

$$\widehat{d}(p) + \widehat{f}(p) \leq \bar{d}(p) + \bar{f}(p) \leq \widehat{d}(p) + \widehat{f}(p) + \delta_1(p) + \delta_2(p).$$

Consequently, a conservative value for $\bar{d}(p) + \bar{f}(p)$ is given by $\widehat{d}(p) + \widehat{f}(p) + \delta_1(p) + \delta_2(p)$, where $\delta_1(p) \to 0$ if $\bar{t}' \to \infty$ and $\delta_2(p) \to 0$ if $\bar{k} \to \infty$.

Computation of $\bar{\mu}_k(\omega_j)$

Consider $\bar{\mu}_k(\omega_j)$ defined in (4.23). In order to calculate $\bar{\mu}_k(\omega_j)$ choose a truncation value $\bar{k}' \geq \max\{k^*, n\}$ and calculate

$$\widehat{\mu}_k(\omega_j) = \sum_{k'=n+1}^{\bar{k}'} \bar{g}(k') \left| \mathrm{Re}\left(e^{i2\pi k/m} B_{k'}(e^{i\omega_j}) \right) \right|,$$

$$\delta = \frac{M c \rho^{\bar{k}'+1}}{1-\rho},$$

where c is given by

$$c = \max_{k'=\bar{k}'+1,\ldots,\infty} \|B_{k'}(z)\|_{\infty},$$

which is easily computable according to Fact 4.B.3 in Appendix 4.B. Since

$$\delta = \sum_{k'=\bar{k}'+1}^{\infty} M \rho^{k'} c \geq \sum_{k'=\bar{k}'+1}^{\infty} M \rho^{k'} \|B_{k'}(z)\|_{\infty} \geq$$

$$\geq \sum_{k'=\bar{k}'+1}^{\infty} \bar{g}(k') \left| \mathrm{Re}\left(e^{i2\pi k/m} B_{k'}(e^{i\omega_j}) \right) \right|, \ \forall k, \omega_j,$$

this gives

$$\widehat{\mu}_k(\omega_j) \leq \bar{\mu}_k(\omega_j) \leq \widehat{\mu}_k(\omega_j) + \delta.$$

Consequently, a conservative value for $\bar{\mu}_k(\omega_j)$ is given by $\widehat{\mu}_k(\omega_j) + \delta$, where $\delta \to 0$ if $\bar{k}' \to \infty$.

Computation of β_1

Consider β_1 defined in Lemma 4.4.4. Choose a truncation value $\bar{k} \geq k^*$ and calculate

$$\widehat{\beta}_1 = \sum_{k=0}^{\bar{k}} \bar{g}(k) \left\| \frac{\mathrm{d}\, B_k(z)}{\mathrm{d}\, z} \right\|_\infty ,$$

$$\delta = \|W(z)\|_\infty \sum_{k=\bar{k}+1}^{\infty} M\rho^k \left(c_1 + c_2 k^{\frac{3}{2}} \right),$$

where c_1 and c_2 are chosen such that

$$\sum_{k'=0}^{\infty} k' \, |b_k(k')| \leq c_1 + c_2 k^{\frac{3}{2}},$$

which is possible according to Conjecture 4.B.4. Since

$$\begin{aligned} \delta &\geq \|W(z)\|_\infty \sum_{k=\bar{k}+1}^{\infty} M\rho^k \sum_{k'=0}^{\infty} k' \, |b_k(k')| = \\ &= \|W(z)\|_\infty \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \left\| \frac{\mathrm{d} \sum_{k'=0}^{\infty} b_k(k') z^{-k'}}{\mathrm{d}\, z} \right\|_{\ell_1} \geq \\ &\geq \|W(z)\|_\infty \sum_{k=\bar{k}+1}^{\infty} \bar{g}(k) \left\| \frac{\mathrm{d}\, B_k(z)}{\mathrm{d}\, z} \right\|_\infty , \end{aligned}$$

this gives

$$\widehat{\beta}_1 \leq \beta_1 \leq \widehat{\beta}_1 + \delta.$$

Consequently, a conservative value for β_1 is given by $\widehat{\beta}_1 + \delta$, where $\delta \to 0$ if $\bar{k} \to \infty$.

Computation of $\bar{p}_\mathrm{l}(k)$ and $\bar{p}_\mathrm{u}(k)$

Consider $\bar{p}_\mathrm{l}(k)$ and $\bar{p}_\mathrm{u}(k)$ defined in (4.30) and (4.31) respectively. Choose a truncation value $\bar{k}' \geq \max\{k^*, n\}$ and calculate

$$\begin{aligned} \widehat{p}_\mathrm{l}(k) &= -\sum_{k'=n+1}^{\bar{k}'} \bar{g}(k') \, |s_{k'}(k)| , \\ \widehat{p}_\mathrm{u}(k) &= \sum_{k'=n+1}^{\bar{k}'} \bar{g}(k') \, |s_{k'}(k)| , \end{aligned}$$

$$\delta = \frac{Mc \, \|W(z)\|_\infty \, \rho^{\bar{k}'+1}}{1-\rho},$$

where c is given by

$$c = \max_{k'=\bar{k}'+1,\ldots,\infty} \|B_{k'}(z)\|_\infty ,$$

which is easily computable according to Fact 4.B.3 in Appendix 4.B. The H_∞-norm of any ℓ_2-stable transfer function overbounds the absolute value of any of its pulse response parameters, in particular

$$\|B_{k'}(z)W(z)\|_\infty = \|S_{k'}(z)\|_\infty \geq |s_{k'}(k)|, \ \forall k.$$

Using this and the multiplicative property of the H_∞-norm, yields

$$\begin{aligned} \delta &= \sum_{k'=\bar{k}'+1}^{\infty} M\rho^{k'} c \|W(z)\|_\infty \geq \sum_{k'=\bar{k}'+1}^{\infty} M\rho^{k'} \|B_{k'}(z)\|_\infty \|W(z)\|_\infty \geq \\ &\geq \sum_{k'=\bar{k}'+1}^{\infty} \bar{g}(k') \|B_{k'}(z)W(z)\|_\infty \geq \sum_{k'=\bar{k}'+1}^{\infty} \bar{g}(k')|s_{k'}(k)|, \ \forall k, \end{aligned}$$

and consequently,

$$\begin{aligned} \widehat{p}_{\mathrm{l}}(k) - \delta \leq \bar{p}_{\mathrm{l}}(k) \leq \widehat{p}_{\mathrm{l}}(k), \\ \widehat{p}_{\mathrm{u}}(k) \leq \bar{p}_{\mathrm{u}}(k) \leq \widehat{p}_{\mathrm{u}}(k) + \delta. \end{aligned}$$

Consequently, a conservative value for $\bar{p}_{\mathrm{l}}(k)$ is given by $\widehat{p}_{\mathrm{l}}(k) - \delta$, and a conservative value for $\bar{p}_{\mathrm{u}}(k)$ is given by $\widehat{p}_{\mathrm{u}}(k) + \delta$, where $\delta \to 0$ if $\bar{k}' \to \infty$.

Completely similar to this $p_{\mathrm{l}}(k)$ and $p_{\mathrm{u}}(k)$ may be evaluated for $k > m$. However, their individual values are not relevant, only their summation γ in Theorem 4.5.4, is relevant.

Computation of γ

Consider γ defined in Theorem 4.5.4. Choose a truncation value $\bar{k}' \geq k^*$ and calculate

$$\widehat{\gamma} = \sum_{k'=0}^{\bar{k}'} \bar{g}(k') \left(\|S_{k'}(z)\|_{\ell_1} - \sum_{k=0}^{m} |s_{k'}(k)| \right) + \left\| \widehat{G}(z)W(z) \right\|_{\ell_1} - \sum_{k=0}^{m} |\widehat{p}(k)| ,$$

$$\delta = \sum_{k'=\bar{k}'+1}^{\infty} M\rho^{k'} \left(c_1 + c_2\sqrt{k'} \right) \|W(z)\|_{\ell_1} ,$$

where c_1 and c_2 are chosen such that (4.35) holds. Since

$$\begin{aligned} \gamma &= \sum_{k=m+1}^{\infty} (p_{\mathrm{u}}(k) + |\widehat{p}(k)|) = \sum_{k=m+1}^{\infty} \left(\sum_{k'=0}^{\infty} \bar{g}(k') |s_{k'}(k)| + |\widehat{p}(k)| \right) = \\ &= \sum_{k'=0}^{\infty} \bar{g}(k') \sum_{k=m+1}^{\infty} |s_{k'}(k)| + \sum_{k=m+1}^{\infty} |\widehat{p}(k)| = \\ &= \sum_{k'=0}^{\infty} \bar{g}(k') \left(\|S_{k'}(z)\|_{\ell_1} - \sum_{k=0}^{m} |s_{k'}(k)| \right) + \left\| \widehat{G}(z)W(z) \right\|_{\ell_1} - \sum_{k=0}^{m} |\widehat{p}(k)| , \end{aligned}$$

and

$$\delta \geq \sum_{k'=\bar{k}'+1}^{\infty} M\rho^{k'} \|B_{k'}(z)\|_{\ell_1} \|W(z)\|_{\ell_1} \geq \sum_{k'=\bar{k}'+1}^{\infty} \bar{g}(k') \|S_{k'}(z)\|_{\ell_1} \geq$$

$$\geq \sum_{k'=\bar{k}'+1}^{\infty} \bar{g}(k') \left(\|S_{k'}(z)\|_{\ell_1} - \sum_{k=0}^{m} |s_{k'}(k)| \right),$$

this gives

$$\widehat{\gamma} \leq \gamma \leq \widehat{\gamma} + \delta.$$

Consequently, a conservative value for γ is given by $\widehat{\gamma} + \delta$, where $\delta \to 0$ if $\bar{k}' \to \infty$.

Chapter 5

A Procedure for Probabilistic Uncertainty Bounding Identification

5.1 Introduction

In this chapter a procedure is developed to identify probabilistic uncertainty regions for the frequency and pulse response of some unknown system. In fact the procedure is the statistical counterpart of the procedure for deterministic uncertainty bounding identification described in the previous chapter. Similarly, the procedure leads to the construction of a system uncertainty set $\mathcal{S}$ of the format as specified in (2.3) or (2.4).

The prior information about the system is exactly identical, i.e. a bound is assumed on the system's generalized pulse response parameters. The prior assumptions about the noise are different from those in Chapter 4. Instead of using deterministic noise bounds, stochastic noise assumptions are made. The basic assumption about the noise process is that it is stationary and independent of the input signal in open loop, or an external reference signal in closed loop. The probability density function of the noise process is arbitrary and not assumed to be known. Instead asymptotic results are derived with a central limit theorem.

The procedure is based on an instrumental variable (IV) estimate, using a linear model parametrization in terms of general basis functions. The frequency and pulse response confidence regions are estimated by explicitly evaluating the bias and variance errors of this instrumental variable estimate. The bias error of the IV estimate is bounded by using the prior bound on the generalized pulse response parameters of the system. It is also indicated how data can be used to estimate this prior parameter bound.

In order to correctly quantify the variance error of the IV estimate, knowledge is required of the auto-covariance function of the noise process. A procedure, which makes use of classical spectral estimation, is presented to estimate this from data, such that asymptotically correct results are obtained. In fact it is shown in this chapter that classical identification techniques, in particular instrumental variable and spectral estimation techniques, can fruitfully be applied to identify probabilistic frequency and pulse response uncertainty regions.

The outline of the chapter is as follows. In the next section the identification setting is described. Section 5.3 presents the instrumental variable estimate. In Section 5.4 the frequency response error of the IV model is evaluated, which leads to probabilistic frequency response system uncertainty regions. In the subsequent section the pulse response error of the IV model is evaluated, leading to probabilistic pulse response uncertainty regions. In Section 5.6 it is explained how the required prior information about system and noise can be estimated from data. Next, in Section 5.7 a survey is given of the entire uncertainty bounding identification procedure. In Section 5.8 MIMO extensions are presented and in Section 5.9 an example is shown. Finally, in Section 5.10 the results are discussed.

5.2 Identification Setting

Consider the linear, time-invariant, discrete time, causal and ℓ_∞-stable SISO system $G_0(z)$ represented by

$$G_0(z) = \sum_{k=0}^{\infty} g_0(k) B_k(z),$$

where $\{B_k(z)\}_{k=0,\ldots,\infty}$, is some specified set of basis functions, see Subsection 2.4.3, given by

$$B_k(z) = \sum_{k'=0}^{\infty} b_k(k') z^{-k'}, \ k = 0, \ldots, \infty, \tag{5.1}$$

for given and known scalar pulse response parameters $b_k(k')$.

Consider given input data $\{u(t)\}_{t=1,\ldots,N}$ and measured output data $\{y(t)\}_{t=1,\ldots,N}$ and the following input-output relation of the data generating system,

$$y(t) = G_0(q)u(t) + e_0(t), \ t = 1, \ldots, N, \tag{5.2}$$

where N denotes the measurement time and $\{e_0(t)\}$ is an unknown additive output noise. There are no restrictions on the input signal, basically it may be determined in open loop as well as in closed loop.

It is assumed that a signal $\{r(t)\}_{t=1,\ldots,N}$ is available, which is highly correlated with the input signal $\{u(t)\}$, but independent of the noise process $\{e_0(t)\}$. Let by definition $r(t) = 0$ for $t \leq 0$. Typically in open loop operation the signal $\{r(t)\}$ is equal to the input $\{u(t)\}$. In a closed loop environment an external reference signal $\{\bar{r}(t)\}$ can be used, or a filtered version of this signal, $r(t) = F(q)\bar{r}(t)$. More is said about this later, in Remark 5.4.5.

The following assumptions are made about the noise process $\{e_0(t)\}$.

Assumption 5.2.1 *The noise process $\{e_0(t)\}$ is stationary with auto-covariance function $R_{e_0}(\tau) = Ee_0(t+\tau)e_0(t)$ and it satisfies $e_0(t) = H_0(q)w_0(t)$ for some ℓ_2-stable $H_0(q)$, and where $\{w_0(t)\}$ is a sequence of independent random variables with zero mean values, variances λ_0 and bounded fourth moments.*

Note that the distribution of the noise process is arbitrary and not assumed to be known. The following assumptions about $\{r(t)\}$ are made.

Assumption 5.2.2 *The signal $\{r(t)\}$ is a bounded deterministic quasi-stationary signal (Ljung, 1987, Ch. 2), so its auto-covariance function*

$$R_r(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} r(t+\tau)r(t)$$

exists $\forall\, \tau$.

In order to cope with unknown initial conditions the input signal in the past is assumed to be bounded by

$$|u(t)| \leq \bar{u}, \ \forall\, t \leq 0, \tag{5.3}$$

for some given $\bar{u}$.

The coefficients $g_0(k)$ are assumed to be bounded by

$$|g_0(k)| \leq \bar{g}(k), \ k = 0, \ldots, \infty, \tag{5.4}$$

for given $\bar{g}(k)$. Moreover, it is assumed that the bound $\bar{g}(k)$ shows exponential decay rate for k larger than some k^*, i.e.

$$\bar{g}(k) \leq M\rho^k, \ \forall\, k > k^*, \tag{5.5}$$

for some given $M \geq 0$ and $\rho < 1$. In Subsection 5.6.2 it is explained how this bound may be established, if it is not available a priori, i.e. how to choose the bounds such that the system $G_0(z)$ satisfies them.

The identification objective is to derive probabilistic uncertainty bounds for the system's frequency response,

$$G_0(e^{i\omega}) = \sum_{k=0}^{\infty} g_0(k)B_k(e^{i\omega}),$$

and also for the pulse response parameters $p_0(k)$, where

$$G_0(z) = W^{-1}(z) \sum_{k=0}^{\infty} p_0(k)z^{-k},$$

with $W(z)$ an arbitrary LTI, finite dimensional, stable and stably invertible weighting. The identification problem is tackled by splitting the transfer function $G_0(z)$ into two parts,

$$G_0(z) = \widetilde{G}_0(z) + \bar{G}_0(z), \ \widetilde{G}_0(z) = \sum_{k=0}^{n} g_0(k)B_k(z), \ \bar{G}_0(z) = \sum_{k=n+1}^{\infty} g_0(k)B_k(z), \tag{5.6}$$

for some user-defined truncation value n. Also, define $\widetilde{p}_0(k)$ and $\bar{p}_0(k)$ correspondingly,

$$\sum_{k=0}^{\infty} \widetilde{p}_0(k) z^{-k} = W(z)\widetilde{G}_0(z), \quad \sum_{k=0}^{\infty} \bar{p}_0(k) z^{-k} = W(z)\bar{G}_0(z),$$

such that, obviously, $p_0(k) = \widetilde{p}_0(k) + \bar{p}_0(k)$, $\forall\, k$.

Next, deterministic uncertainty bounds will be determined for the tails $\bar{G}_0(e^{i\omega})$ and $\bar{p}_0(k)$, using the deterministic prior bounds $\bar{g}(k)$ given in (5.4). And probabilistic uncertainty bounds will be derived for $\widetilde{G}_0(e^{i\omega})$ and $\widetilde{p}_0(k)$, using variance expressions of an instrumental variable estimate. These variance expressions are based on the stochastic noise assumption 5.2.1. In the variance expressions the influence of the undermodelling part $\bar{G}_0(z)$ is properly taken into account. The sum of the deterministic uncertainty bounds for $\bar{G}_0(z)$ and the probabilistic uncertainty bounds for $\widetilde{G}_0(z)$ provides probabilistic uncertainty regions for the system $G_0(z)$.

Note that generally there is an optimal value for n. If it is chosen too small, the resulting bounds will be completely determined by the prior information (5.4), which is generally conservative. If it is chosen too large, the confidence regions for $\widetilde{G}_0(z)$ will be large as the variance increases with the number of parameters to be estimated. More is said about this later, after the main Theorem 5.4.3 has been established.

5.3 The Instrumental Variable Estimate

Let the model $G(z)$ be parametrized as follows,

$$G(z) = \sum_{k=0}^{n} g(k) B_k(z),$$

where $\{g(k)\}_{k=0,\ldots,n}$ are the model parameters. Define the model input signal $\widetilde{u}(t)$ as

$$\widetilde{u}(t) := \begin{cases} 0, & t \leq 0, \\ u(t), & t = 1, \ldots, N. \end{cases} \tag{5.7}$$

The model input-output relation is given by

$$y(t) = G(q)\widetilde{u}(t) + e(t) = \sum_{k=0}^{n} g(k) B_k(q)\widetilde{u}(t) + e(t) = \sum_{k=0}^{n} g(k) w_k(t) + e(t),$$

where $e(t)$ is the output error, and

$$w_k(t) := \sum_{k'=0}^{t-1} b_k(k') u(t - k'), \;\; k = 0, \ldots, n. \tag{5.8}$$

Next, define the instrumental signals

$$v_k(t) := B_k(q) r(t) = \sum_{k'=0}^{t-1} b_k(k') r(t - k'), \;\; k = 0, \ldots, n, \tag{5.9}$$

and the column-vectors

$$U(t) := \begin{bmatrix} w_0(t) \\ \vdots \\ w_n(t) \end{bmatrix}, \; V(t) := \begin{bmatrix} v_0(t) \\ \vdots \\ v_n(t) \end{bmatrix}. \tag{5.10}$$

Also, introduce for notational convenience

$$\widetilde{N} = N - t_s + 1,$$

for some integer $t_s \in [1, N)$, which is user-defined. The integer t_s represents the starting sample used in the IV estimate and can be used to reduce the influence of the unknown initial conditions, as will become clear later.

Consider the basic IV estimate (Söderström and Stoica, 1989, p. 262; Ljung, 1987, p. 192/193),

$$\begin{aligned} \begin{bmatrix} \widehat{g}(0) \\ \vdots \\ \widehat{g}(n) \end{bmatrix} &= \text{sol}\left\{ \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)e(t) = 0 \right\} = \\ &= \text{sol}\left\{ \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t) \left(y(t) - \sum_{k=0}^{n} g(k)w_k(t) \right) = 0 \right\}, \end{aligned}$$

which is given by

$$\begin{bmatrix} \widehat{g}(0) \\ \vdots \\ \widehat{g}(n) \end{bmatrix} = \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t) \right]^{-1} \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)y(t). \tag{5.11}$$

Notice that in case of open loop operation, $r(t) = u(t)$, this is just a FIR least squares estimate for general basis functions. The estimated IV model is given by

$$\widehat{G}(z) = \sum_{k=0}^{n} \widehat{g}(k)B_k(z).$$

On the basis of this identified model, frequency and pulse response uncertainty regions are constructed. This is done by explicitly calculating the bias and variance errors of the IV estimate.

5.4 Frequency Response Uncertainty Regions

An analysis is made of the frequency response identification error of the instrumental variable estimate. This then leads to frequency response confidence regions for the system $G_0(z)$. These in turn specify a system uncertainty set $\mathcal{S}$ of the format specified in (2.3).

5.4.1 The Frequency Response Error of the IV Model

Consider some frequency ω_j chosen arbitrarily in the interval $[0, \pi]$. Substitution of the parameter estimate (5.11) yields the value of the frequency response of the IV estimate for the frequency ω_j,

$$\begin{aligned}\widehat{G}(e^{i\omega_j}) &= \sum_{k=0}^{n} \widehat{g}(k) B_k(e^{i\omega_j}) = \\ &= \left[B_0(e^{i\omega_j}) \cdots B_n(e^{i\omega_j})\right] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t)\right]^{-1} \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)y(t). \end{aligned} \tag{5.12}$$

Define for $t = t_s, \ldots, N$ the signals $r_1(t)$ and $r_2(t)$ as

$$r_1(t) := \left[\mathrm{Re}\left(B_0(e^{i\omega_j})\right) \cdots \mathrm{Re}\left(B_n(e^{i\omega_j})\right)\right] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t)\right]^{-1} V(t), \tag{5.13}$$

$$r_2(t) := \left[\mathrm{Im}\left(B_0(e^{i\omega_j})\right) \cdots \mathrm{Im}\left(B_n(e^{i\omega_j})\right)\right] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t)\right]^{-1} V(t). \tag{5.14}$$

These signals $r_p(t)$, $p = 1, 2$, are filtered versions of the signal $r(t)$ and they can be computed, as they only depend on known quantities. They play an essential role throughout the following derivation of IV model error bounds. Note that they depend on the frequency ω_j that has been chosen, but for notational convenience this dependency is not explicitly mentioned all the time. Finally, it is mentioned that the signals $r_p(t)$, $p = 1, 2$, in fact play a role identical to the signals $r_p(t)$ in Chapter 3 (Section 3.2) and Chapter 4.

Using (5.2) and (5.6) the output $y(t)$ can be written as

$$\begin{aligned} y(t) &= G_0(q)u(t) + e_0(t) = \widetilde{G}_0(q)u(t) + \bar{G}_0(q)u(t) + e_0(t) = \\ &= \sum_{k=0}^{n} g_0(k)B_k(q)u(t) + \sum_{k=n+1}^{\infty} g_0(k)B_k(q)u(t) + e_0(t) = \\ &= \sum_{k=0}^{n} g_0(k)w_k(t) + a(t) + b(t) + e_0(t), \end{aligned}$$

where $w_k(t)$ is defined in (5.8) and

$$a(t) := \sum_{k=n+1}^{\infty} g_0(k) \sum_{k'=0}^{t-1} b_k(k')u(t-k'), \tag{5.15}$$

$$b(t) := \sum_{k=0}^{\infty} g_0(k) \sum_{k'=t}^{\infty} b_k(k')u(t-k'). \tag{5.16}$$

The signal $a(t)$ represents the response of the tail $\bar{G}_0(q)$. The signal $b(t)$ represents the response due to past input signals, the initial conditions. Using this the following alternative expression can be given for $\widehat{G}(e^{i\omega_j})$ given by (5.12),

$$\begin{aligned}\widehat{G}(e^{i\omega_j}) &= \frac{1}{\widetilde{N}}\sum_{t=t_s}^{N}(r_1(t)+ir_2(t))y(t) = \\ &= \frac{1}{\widetilde{N}}\sum_{t=t_s}^{N}(r_1(t)+ir_2(t))\left(\sum_{k=0}^{n}g_0(k)w_k(t)+a(t)+b(t)+e_0(t)\right).\end{aligned} \tag{5.17}$$

The first term of this expression can be worked out as follows,

$$\begin{aligned}&\frac{1}{\widetilde{N}}\sum_{t=t_s}^{N}(r_1(t)+ir_2(t))\sum_{k=0}^{n}g_0(k)w_k(t) = \\ &= \left[B_0(e^{i\omega_j})\cdots B_n(e^{i\omega_j})\right]\left[\frac{1}{\widetilde{N}}\sum_{t=t_s}^{N}V(t)U^T(t)\right]^{-1}\frac{1}{\widetilde{N}}\sum_{t=t_s}^{N}V(t)U^T(t)\begin{bmatrix}g_0(0)\\ \vdots \\ g_0(n)\end{bmatrix} = \\ &= \left[B_0(e^{i\omega_j})\cdots B_n(e^{i\omega_j})\right]\begin{bmatrix}g_0(0)\\ \vdots \\ g_0(n)\end{bmatrix} = \widetilde{G}_0(e^{i\omega_j}).\end{aligned}$$

Next, define for $p = 1, 2$,

$$d(p) := \sum_{t=t_s}^{N} r_p(t)a(t), \tag{5.18}$$

$$f(p) := \sum_{t=t_s}^{N} r_p(t)b(t), \tag{5.19}$$

which depend on the frequency ω_j as $r_p(t)$, $p = 1, 2$, depends on the frequency ω_j. Again using (5.6), this finally yields the following expression for the identification error.

Proposition 5.4.1 *Let the IV model $\widehat{G}(z)$ be as defined in Section 5.3. Consider the signals $r_p(t)$, $p = 1, 2$, defined in (5.13) and (5.14), and the signals $d(p)$, $f(p)$, $p = 1, 2$, defined in (5.18) and (5.19), which all depend on the chosen frequency ω_j. Then, the frequency response model error $\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})$ is equal to*

$$\begin{aligned}&\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j}) = \widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j}) - \bar{G}_0(e^{i\omega_j}) = \\ &= \frac{1}{\widetilde{N}}\left(d(1)+id(2)+f(1)+if(2)+\sum_{t=t_s}^{N}(r_1(t)+ir_2(t))e_0(t)\right) - \bar{G}_0(e^{i\omega_j}).\end{aligned} \tag{5.20}$$

Proof: By construction. □

This proposition expresses the frequency response identification model error in terms of $d(p)$ (undermodelling), $f(p)$ (initial conditions), $e_0(t)$ (noise) and $\bar{G}_0(e^{i\omega_j})$ (undermodelling). Basically all terms at the right-hand side of the expression (5.20) are unknown. However, it appears possible to derive a probabilistic distribution for the term containing $e_0(t)$, using assumption 5.2.1. And the terms with $d(p)$, $f(p)$ and $\bar{G}_0(e^{i\omega_j})$ can be bounded using the prior information (5.3) and (5.4).

5.4.2 Auxiliary Results

In this subsection the various terms appearing in (5.20) are evaluated. Consider any bounded signal $\{r_p(t)\}$ and consider $\bar{d}(p)$, $\bar{f}(p)$ defined by (5.18), (5.19) respectively, with $a(t)$, $b(t)$ defined by (5.15), (5.16) respectively. Completely similar to (4.16) and (4.17) in Chapter 4, the following bounds can be derived, making use of (5.3) and (5.4),

$$|d(p)| \leq \bar{d}(p) := \sum_{k=n+1}^{\infty} \bar{g}(k) \left| \sum_{t=t_s}^{N} r_p(t) \sum_{k'=0}^{t-1} b_k(k') u(t-k') \right|, \tag{5.21}$$

which represents a computable bound for the tail contribution. And,

$$|f(p)| \leq \bar{f}(p) := \sum_{k=0}^{\infty} \bar{g}(k) \sum_{t'=0}^{\infty} \left| \sum_{t=t_s}^{N} r_p(t) b_k(t+t') \right| \bar{u}, \tag{5.22}$$

which represents a computable bound for the contribution of the unknown initial conditions. The actual computation of the expressions involve the evaluation of infinite sums. Due to the fact that $\bar{g}(k)$ shows exponential decay rate in k and $b_k(k')$ shows exponential decay rate in k', the outcomes are finite. Computational aspects are considered in Appendix 4.C of Chapter 4. Clearly $\bar{d}(p)$ will be small if n is chosen large and $\bar{f}(p)$ will be small if t_s is chosen large.

The real and imaginary part of the frequency response of the tail, $\bar{G}_0(e^{i\omega_j})$, can be bounded as follows,

$$\left|\mathrm{Re}\left(\bar{G}_0(e^{i\omega_j})\right)\right| = \left|\mathrm{Re}\left(\sum_{k=n+1}^{\infty} g_0(k) B_k(e^{i\omega_j})\right)\right| \leq \sum_{k=n+1}^{\infty} \bar{g}(k) \left|\mathrm{Re}\left(B_k(e^{i\omega_j})\right)\right| =: \delta(1), \tag{5.23}$$

$$\left|\mathrm{Im}\left(\bar{G}_0(e^{i\omega_j})\right)\right| \leq \sum_{k=n+1}^{\infty} \bar{g}(k) \left|\mathrm{Im}\left(B_k(e^{i\omega_j})\right)\right| =: \delta(2). \tag{5.24}$$

Note that $\delta(1)$ and $\delta(2)$ are finite due to the exponential decay rate of $\bar{g}(k)$. Computational aspects of the evaluation of these infinite sums are considered in Appendix 4.C of Chapter 4.

Next, a key lemma is established with respect to the asymptotic distribution of $\sum_{t=t_s}^{N} r_p(t) e_0(t)$. The lemma is in fact a multivariable extension of Theorem 3.2.5 in Chapter 3.

Lemma 5.4.2 *Suppose that* $\{e_0(t)\}$ *and* $\{r(t)\}$ *are independent and that they satisfy the assumptions 5.2.1 and 5.2.2 respectively. Consider the signals* $\{r_1(t)\}$ *and* $\{r_2(t)\}$, *given by* $r_1(t) = F_1(q)r(t)$, $r_2(t) = F_2(q)r(t)$ *for any* ℓ_∞*-stable linear filters* $F_1(q)$ *and* $F_2(q)$. *Denote*

$$\Lambda^N_{r_1r_2} := E\frac{1}{\widetilde{N}}\begin{bmatrix}\sum_{t=t_s}^{N} r_1(t)e_0(t)\\ \sum_{t=t_s}^{N} r_2(t)e_0(t)\end{bmatrix}\begin{bmatrix}\sum_{t=t_s}^{N} r_1(t)e_0(t) & \sum_{t=t_s}^{N} r_2(t)e_0(t)\end{bmatrix},$$

and

$$\Lambda_{r_1r_2} := \lim_{N\to\infty}\Lambda^N_{r_1r_2}.$$

Also, denote for $i, j = 1, 2$,

$$R^N_{r_ir_j}(\tau) := \frac{1}{\widetilde{N}+\tau}\sum_{t=t_s}^{N+\tau} r_i(t)r_j(t-\tau),\ \tau = -N+t_s,\dots,0,$$

$$R^N_{r_ir_j}(\tau) := \frac{1}{\widetilde{N}-\tau}\sum_{t=t_s}^{N-\tau} r_i(t+\tau)r_j(t),\ \tau = 1,\dots,N-t_s,$$

and $R^N_{r_i}(\tau) := R^N_{r_ir_i}(\tau)$, $i = 1, 2$. *Then,*

(i)
$$\Lambda^N_{r_1r_2} = \begin{bmatrix}\sum_{\tau=-N+t_s}^{N-t_s}\frac{\widetilde{N}-|\tau|}{\widetilde{N}}R^N_{r_1}(\tau)R_{e_0}(\tau) & \sum_{\tau=-N+t_s}^{N-t_s}\frac{\widetilde{N}-|\tau|}{\widetilde{N}}R^N_{r_1r_2}(\tau)R_{e_0}(\tau)\\ \sum_{\tau=-N+t_s}^{N-t_s}\frac{\widetilde{N}-|\tau|}{\widetilde{N}}R^N_{r_1r_2}(\tau)R_{e_0}(\tau) & \sum_{\tau=-N+t_s}^{N-t_s}\frac{\widetilde{N}-|\tau|}{\widetilde{N}}R^N_{r_2}(\tau)R_{e_0}(\tau)\end{bmatrix}$$

(ii)
$$\Lambda_{r_1r_2} = \begin{bmatrix}\sum_{\tau=-\infty}^{\infty}R_{r_1}(\tau)R_{e_0}(\tau) & \sum_{\tau=-\infty}^{\infty}R_{r_1r_2}(\tau)R_{e_0}(\tau)\\ \sum_{\tau=-\infty}^{\infty}R_{r_1r_2}(\tau)R_{e_0}(\tau) & \sum_{\tau=-\infty}^{\infty}R_{r_2}(\tau)R_{e_0}(\tau)\end{bmatrix}$$

(iii)
$$\frac{1}{\sqrt{\widetilde{N}}}\begin{bmatrix}\sum_{t=t_s}^{N} r_1(t)e_0(t)\\ \sum_{t=t_s}^{N} r_2(t)e_0(t)\end{bmatrix} \xrightarrow{N\to\infty} \mathcal{N}(0, \Lambda_{r_1r_2}),$$

where $\mathcal{N}(0, \Lambda_{r_1r_2})$ *denotes the Multivariate Normal distribution with mean 0 and co-*

variance matrix $\Lambda_{r_1r_2}$. *Moreover, if* $\Lambda^N_{r_1r_2}$ *is invertible,*

$$\textbf{(iv)}\quad \frac{1}{\widetilde{N}}\begin{bmatrix}\sum_{t=t_s}^{N} r_1(t)e_0(t) & \sum_{t=t_s}^{N} r_2(t)e_0(t)\end{bmatrix}\left(\Lambda^N_{r_1r_2}\right)^{-1}\begin{bmatrix}\sum_{t=t_s}^{N} r_1(t)e_0(t)\\ \sum_{t=t_s}^{N} r_2(t)e_0(t)\end{bmatrix}\xrightarrow{N\to\infty}\chi^2(2),$$

where $\chi^2(2)$ *denotes the Chi-square distribution with 2 degrees of freedom.*

Proof: See Appendix 5.A. □

The results given in (iii) and (iv) are asymptotic results, established using a central limit theorem. For finite N the given distributions are approximations of the true ones. However, extensive Monte Carlo simulations show that this approximation can be very good for small N already, see Appendix 5.B. Note that the expression for the covariance matrix in part (i) is a non-asymptotic result, it is correct for any N.

5.4.3 Frequency Response Confidence Regions

Using the results of the previous subsection a computable bound for the IV model error $\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})$ is straightforwardly obtained. And as such a confidence region for the system's frequency response $G_0(e^{i\omega_j})$ is obtained. The bound is given in the following main theorem.

Theorem 5.4.3 *Consider the IV estimate (5.11) with frequency response* $\widehat{G}(e^{i\omega_j})$ *given by (5.12). Suppose that* $\{e_0(t)\}$ *and* $\{r(t)\}$ *are independent and that they satisfy the assumptions 5.2.1 and 5.2.2 respectively. Let* $\bar{d}(p)$, $p = 1, 2$, *and* $\bar{f}(p)$, $p = 1, 2$, *be given by (5.21) and (5.22) respectively, with* $r_1(t)$ *and* $r_2(t)$ *given by (5.13) and (5.14) respectively. Moreover, let* $\delta(p)$, $p = 1, 2$, *be given by (5.23) and (5.24). Finally, denote matrix-element* (i, j) *of* $\Lambda^N_{r_1r_2}$ *as given in part (i) of Lemma 5.4.2 by* $\lambda^N_{r_ir_j}$. *Then, if* $N \to \infty$,

$$\textbf{(i)}\quad \left|\mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\right| \le c_{\mathcal{N},\alpha}\sqrt{\frac{\lambda^N_{r_1r_1}}{\widetilde{N}}} + \frac{\bar{d}(1)}{\widetilde{N}} + \frac{\bar{f}(1)}{\widetilde{N}} + \delta(1),\ \ w.p. \ge \alpha,$$

$$\textbf{(ii)}\quad \left|\mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\right| \le c_{\mathcal{N},\alpha}\sqrt{\frac{\lambda^N_{r_2r_2}}{\widetilde{N}}} + \frac{\bar{d}(2)}{\widetilde{N}} + \frac{\bar{f}(2)}{\widetilde{N}} + \delta(2),\ \ w.p. \ge \alpha,$$

where $c_{\mathcal{N},\alpha}$ *corresponds to a probability* α *in the standard Normal distribution, such that, if* $x \in \mathcal{N}(0,1) \Rightarrow \mathrm{prob}(|x| \le c_{\mathcal{N},\alpha}) = \alpha$.

Moreover, if $\Lambda^N_{r_1r_2}$ *is invertible, introduce* $\Gamma = \begin{bmatrix}\gamma_{11} & \gamma_{12}\\ \gamma_{21} & \gamma_{22}\end{bmatrix}$ *as the square-root of the inverse of* $\Lambda^N_{r_1r_2}$, *i.e.* $\Gamma^T\Gamma = \left(\Lambda^N_{r_1r_2}\right)^{-1}$ *and* $\Gamma^T = \Gamma$. *Then, if* $N \to \infty$,

$$\textbf{(iii)}\quad \begin{bmatrix}\mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\end{bmatrix}^T \Gamma^T\Gamma \begin{bmatrix}\mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\end{bmatrix} \le$$

$$\leq \left(\sqrt{\frac{c_{\chi,\alpha}}{\widetilde{N}}} + \sqrt{\gamma_{11}^2 + \gamma_{21}^2} \left(\frac{\bar{d}(1)}{\widetilde{N}} + \frac{\bar{f}(1)}{\widetilde{N}} + \delta(1) \right) + \right.$$
$$\left. + \sqrt{\gamma_{12}^2 + \gamma_{22}^2} \left(\frac{\bar{d}(2)}{\widetilde{N}} + \frac{\bar{f}(2)}{\widetilde{N}} + \delta(2) \right) \right)^2, \quad w.p. \geq \alpha,$$

where $c_{\chi,\alpha}$ corresponds to a probability α in the Chi-square distribution with 2 degrees of freedom, such that, if $x \in \chi^2(2) \Rightarrow \text{prob}(x \leq c_{\chi,\alpha}) = \alpha$.

Proof: See Appendix 5.A. □

The parts (i) and (ii) of this theorem provide probabilistic bounds for the real and imaginary parts of the IV model error, and as such for the frequency response of the system $G_0(z)$. These may be combined into rectangular system confidence regions in the complex plane using Bonferroni's inequality, Manoukian (1986, p. 49). In particular, if any complex-valued random variable x has the property that $\text{Re}(x) \leq a$, w.p. $\geq \alpha$, and $\text{Im}(x) \leq b$, w.p. $\geq \beta$, then $\text{Re}(x) \leq a \wedge \text{Im}(x) \leq b$, w.p. $\geq 1 - (1-\alpha) - (1-\beta)$.

Ellipsoidal system confidence regions are obtained with part (iii) of the above theorem, provided the matrix $\Lambda^N_{r_1 r_2}$ is invertible. Note that this is generally the case, except for frequencies $\omega_j = 0, \pi$. For these frequencies the signal $\{r_2(t)\}$ is identically zero, as $\text{Im}\left(B_k(e^{i\omega_j})\right)$ appearing in (5.14) is zero. This very naturally means that for frequencies 0 and π there is no imaginary system uncertainty.

The first contribution to the frequency response uncertainty regions as specified in Theorem 5.4.3, corresponds to the variance of the IV model, due to the noise $\{e_0(t)\}$. The second contribution, with $\bar{d}(p)$, is due to the response of the tail $\bar{G}_0(q)$ and represents a bias contribution. The third contribution, with $\bar{f}(p)$, is due to the unknown initial conditions. Finally, the fourth contribution, with $\delta(p)$, corresponds to the frequency response of the tail $\bar{G}_0(q)$ and also represents a bias contribution.

The different error sources in the IV estimate can be clearly distinguished and traded-off. In particular the truncation value n can be used to make a trade-off between bias and variance. A larger value n means a smaller bias, but a larger variance. By trying different values an optimal value can be determined. Similarly, the integer t_s offers the possibility to trade-off the influence of initial conditions to the variance. A larger value t_s means a decrease of the error contribution $\bar{f}(p)$, but an increase of the variance, due to a decreasing $\widetilde{N} = N - t_s + 1$.

It is emphasized that the identification of the IV model is not a goal as such, but serves as a basis for the construction of system uncertainty regions. The design variables in the IV identification, such as the IV model order n, should not be used to obtain a tractable (low-order) nominal model, but should be tuned in such a way that the uncertainty regions are as small as possible. The identification of a good nominal model, suited for use in control design, is not the issue here. This problem is considered in the next part of this thesis, in particular in Chapter 7.

Remark 5.4.4 *The probabilistic uncertainty regions given in Theorem 5.4.3 correspond to an explicit frequency response variance and bias expression for an instrumental variable estimate $\widehat{G}(e^{i\omega_j})$. In case of open loop identification, if $r(t) = u(t)$, the*

IV estimate is identical to a FIR least squares estimate. The expressions have been derived for any set of basis functions, $\{B_k(z)\}_{k=0,\ldots,\infty}$. Also, the contribution of the initial conditions and undermodelling are properly taken into account.

In literature variance expressions are given for IV and FIR estimates, however, mainly with respect to the parameter variance, assuming that the system is in the model set and neglecting the influence of the initial conditions, see e.g. Ljung (1987, Ch. 9) and Söderström and Stoica (1983, Ch. 5; 1989, Ch. 8). Some progress has been made in Hjalmarsson and Ljung (1992) and Hjalmarsson (1993b), where for a different identification setting a procedure is presented to incorporate the influence of the bias when computing the variance.

Another important aspect is that variance expressions found in literature are often derived for the asymptotic case $N \to \infty$, whereas here tractable non-asymptotic expressions have been derived. In particular, the expression for the covariance matrix, as given in part (i) of Lemma 5.4.2, is correct in the non-asymptotic case. Only the distributions in Theorem 5.4.3 are asymptotic.

Remark 5.4.5 *In Stoica and Söderström (1983) and Söderström and Stoica (1989, pp. 274,386,387) conditions are given that make an IV identification method optimal, by which is meant that the estimate has minimal variance (the model structure and model order being fixed). Translated to the current identification setting these conditions are that the output noise $\{e_0(t)\}$ should be white and that the instrumental signal $\{r(t)\}$ should be equal to the noise free part of the input $\{u(t)\}$. This implies that optimally the data $\{u(t)\}$ and $\{y(t)\}$ should be prefiltered with the inverse of (an estimate of) the noise generating filter $H_0(q)$. Moreover, in case of open loop identification the signal $\{r(t)\}$ should be chosen equal to the (prefiltered) input $\{u(t)\}$. In case of closed loop identification it should be chosen as $r(t) = F(q)\bar{r}(t)$, where $\{\bar{r}(t)\}$ is the external reference signal and $F(q)$ is (an estimate of) the transfer function from $\{\bar{r}(t)\}$ to $\{u(t)\}$. Note that imperfections in the estimates of $H_0(q)$ or $F(q)$ will not cause the identified frequency response confidence regions to be incorrect, but these will just be larger than the theoretically optimal ones.*

Finally, it is mentioned that, when prefiltering the data $\{u(t)\}$ and $\{y(t)\}$, the influence of initial conditions should be taken into account. It is possible to do this in a worst-case sense, by using an upper bound on the input and output signals in the past, similar to the calculation of $\bar{f}(p)$ in (5.22). However, in practice it will generally be sufficient to choose t_s large enough, which is dependent on the length of the pulse response of the prefilter, and after that neglect the contribution of the initial conditions.

With Theorem 5.4.3 rectangular or ellipsoidal frequency response uncertainty regions $\mathcal{P}(\omega_j)$ can be calculated, as appearing in (2.3) in Chapter 2. These confidence regions have the property that

$$G_0(e^{i\omega_j}) \in \mathcal{P}(\omega_j), \text{ w.p. } \geq \alpha,$$

for some user-defined confidence level α. The larger α is chosen, the larger $\mathcal{P}(\omega_j)$ will be. The sets can be calculated for all l frequencies ω_j in the user-defined frequency set Ω. Moreover, the derivative bound β_1 as appearing in (2.3) can be calculated

with Lemma 4.4.4, on the basis of the prior information (5.4). This leads to a system uncertainty set $\mathcal{S}$ as defined in (2.3). With Bonferroni's inequality the following simultaneous confidence can be established for this system uncertainty set,

$$G_0(z) \in \mathcal{S}, \text{ w.p. } \geq 1 - l(1-\alpha).$$

This means in practice, that a controller designed to work for all systems in the set $\mathcal{S}$, is guaranteed to work for the system $G_0(z)$ with probability larger than or equal to $1 - l(1-\alpha)$. Such a statement is useful only if l is not too large, or α is very close to 1. Otherwise the simultaneous probability could easily be unacceptable.

Remark 5.4.6 *It seems that, when deriving simultaneous system confidence regions, the application of Bonferroni's inequality is rather conservative. A much better simultaneous probability should be derivable, using the fact that there is one underlying $n+1$-dimensional multivariable normal distribution, the asymptotic distribution of the IV model parameters $\widehat{g}(0), \ldots, \widehat{g}(n)$. The derivation of better expressions for the simultaneous probability certainly deserves further research attention. The main difficulty is that the covariance matrix of the normal distribution can not be assumed completely known. In fact it is completely unknown, but it appears possible to estimate entries of the covariance matrix, as described in Subsection 5.6.1. But it is generally infeasible to estimate all $(n+1)^2$ entries, especially if n is chosen large.*

Similar to the derivation in Subsection 4.4.3, the identified set $\mathcal{S}$ can straightforwardly be applied to establish an upper bound on the H_∞-norm of the model error, with respect to an arbitrary user-defined nominal model. This upper bound will then be correct with a certain probability.

5.5 Pulse Response Uncertainty Regions

In this section an analysis is made of the identification error of the weighted pulse response of the IV estimate. This leads to pulse response uncertainty regions for the (weighted) system $G_0(z)$.

Recall that at the end of Section 5.2 the pulse response parameters $p_0(k)$ were defined by

$$\sum_{k=0}^{\infty} p_0(k) z^{-k} = W(z) G_0(z),$$

and the parameters $\widetilde{p}_0(k)$, $\bar{p}_0(k)$ by

$$\sum_{k=0}^{\infty} \widetilde{p}_0(k) z^{-k} = W(z)\widetilde{G}_0(z), \quad \sum_{k=0}^{\infty} \bar{p}_0(k) z^{-k} = W(z)\bar{G}_0(z),$$

such that $p_0(k) = \widetilde{p}_0(k) + \bar{p}_0(k)$. In addition define the the pulse response parameters $\widehat{p}(k)$ by the following relation,

$$\sum_{k=0}^{\infty} \widehat{p}(k) z^{-k} = W(z)\widehat{G}(z).$$

Also, define $S_{k'}(z)$ as

$$S_{k'}(z) := \sum_{k=0}^{\infty} s_{k'}(k) z^{-k} := B_{k'}(z) W(z), \; k' = 0, \ldots, \infty.$$

Analogously to the derivation in (4.29) in Section 4.5 it can be shown that, for $k = 0, \ldots, \infty$,

$$\widetilde{p}_0(k) = \sum_{k'=0}^{n} g_0(k') s_{k'}(k), \; \bar{p}_0(k) = \sum_{k'=n+1}^{\infty} g_0(k') s_{k'}(k), \; \widehat{p}(k) = \sum_{k'=0}^{n} \widehat{g}(k') s_{k'}(k).$$

Substitution of the parameter estimate (5.11) in this expression for $\widehat{p}(k)$ yields,

$$\widehat{p}(k) = \sum_{k'=0}^{n} \widehat{g}(k') s_{k'}(k) = [s_0(k) \cdots s_n(k)] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t) U^T(t) \right]^{-1} \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t) y(t).$$

Next, define for $t = t_s, \ldots, N$ the computable signal $r_k(t)$,

$$r_k(t) := [s_0(k) \cdots s_n(k)] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t) U^T(t) \right]^{-1} V(t). \tag{5.25}$$

Analogously to the derivation in Section 5.4 it can be shown that

$$\widehat{p}(k) = \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} r_k(t) y(t) = \widetilde{p}_0(k) + \frac{1}{\widetilde{N}} \left(d(k) + f(k) + \sum_{t=t_s}^{N} r_k(t) e_0(t) \right),$$

with

$$d(k) = \sum_{t=t_s}^{N} r_k(t) a(t), \; f(k) = \sum_{t=t_s}^{N} r_k(t) b(t),$$

and where the signals $a(t)$ and $b(t)$ are defined by (5.15), resp. (5.16). Bounds $\bar{d}(k)$ and $\bar{f}(k)$ are given by (5.21) and (5.22), replacing $r_p(t)$ by $r_k(t)$,

$$|d(k)| \leq \bar{d}(k) := \sum_{k'=n+1}^{\infty} \bar{g}(k') \left| \sum_{t=t_s}^{N} r_k(t) \sum_{k''=0}^{t-1} b_{k'}(k'') u(t - k'') \right|, \tag{5.26}$$

which represents a computable bound for the tail contribution. And,

$$|f(k)| \leq \bar{f}(k) := \sum_{k'=0}^{\infty} \bar{g}(k') \sum_{t'=0}^{\infty} \left| \sum_{t=t_s}^{N} r_k(t) b_{k'}(t + t') \right| \bar{u}, \tag{5.27}$$

which represents a computable bound for the contribution of the unknown initial conditions. Finally, define the bound $\delta(k)$ by,

$$|\bar{p}_0(k)| \leq \delta(k) := \sum_{k'=n+1}^{\infty} \bar{g}(k') \, |s_{k'}(k)| \, . \tag{5.28}$$

This leads to the following theorem, which provides probabilistic error bounds for the parameters $\widehat{p}(k)$, and as such confidence intervals for the system parameters $p_0(k)$.

Theorem 5.5.1 *Consider the IV estimate (5.11). Suppose that* $\{e_0(t)\}$ *and* $\{r(t)\}$ *are independent and that they satisfy the assumptions 5.2.1 and 5.2.2 respectively. Let* $\bar{d}(k)$ *and* $\bar{f}(k)$ *be given by (5.26) and (5.27) respectively, with* $r_k(t)$ *given by (5.25). Moreover, let* $\delta(k)$ *be given by (5.28).*

Let $c_{\mathcal{N},\alpha}$ *correspond to a probability* α *in the standard Normal distribution, such that, if* $x \in \mathcal{N}(0,1) \Rightarrow \text{prob}(|x| \leq c_{\mathcal{N},\alpha}) = \alpha$. *Denote,*

$$\lambda^N_{r_k r_k} := \sum_{\tau=-N+t_s}^{N-t_s} \frac{\widetilde{N} - |\tau|}{\widetilde{N}} R^N_{r_k}(\tau) R_{e_0}(\tau).$$

Then, if $N \to \infty$,

$$|\widehat{p}(k) - p_0(k)| \leq c_{\mathcal{N},\alpha} \sqrt{\frac{\lambda^N_{r_k r_k}}{\widetilde{N}}} + \frac{\bar{d}(k)}{\widetilde{N}} + \frac{\bar{f}(k)}{\widetilde{N}} + \delta(k), \;\; w.p. \geq \alpha,$$

Proof: Similar to the proof of part (i) of Theorem 5.4.3, replacing $r_1(t)$ by $r_k(t)$. □

Analogously to the procedure in the deterministic case as described in Section 4.5, the confidence intervals can be calculated for $k = 0, \ldots, m$, for some user-defined integer m, leading to parameter intervals,

$$p_l(k) \leq p_0(k) \leq p_u(k), \;\; \text{w.p.} \geq \alpha.$$

An upper bound can be specified for the remaining parameters, corresponding to $k = m+1, \ldots, \infty$, using the prior bounds in (5.4),

$$|p_0(k)| \leq \sum_{k'=0}^{\infty} \bar{g}(k') \, |s_{k'}(k)|, \;\; k = m+1, \ldots, \infty.$$

This leads to a system uncertainty set $\mathcal{S}$ of the format as given in equation (2.4) in Chapter 2, and which allows for the following statement,

$$G_0(z) \in \mathcal{S}, \;\; \text{w.p.} \geq 1 - m(1-\alpha),$$

where use is made of Bonferroni's inequality. Note that also in this situation Remark 5.4.6 applies.

Similar to the derivation in Subsection 4.5.3 the identified set $\mathcal{S}$ can straightforwardly be applied to establish an upper bound on the ℓ_1-norm of the model error, with respect to an arbitrary user-defined nominal model.

5.6 Estimation of the Prior Information from Data

To identify frequency or pulse response uncertainty regions some prior information about the noise and the system is needed, see Section 5.2. In this section it is shown that data can be used to reliably estimate the required prior information from measurement data. In particular this concerns the estimation of the noise auto-covariance function $R_{e_0}(\tau)$ and the parameter bounds $\bar{g}(k)$.

5.6.1 Estimation of the Noise Auto-Covariance Function

Theorem 5.4.3 provides frequency response confidence regions for the unknown system $G_0(z)$. However, it appears that these can only be calculated if the auto-covariance function of the noise process is known, as $\Lambda^N_{r_1r_2}$ given in part (i) of Lemma 5.4.2 contains $R_{e_0}(\tau)$, $\tau = -N + t_s, \ldots, N - t_s$. Similarly, Theorem 5.5.1 provides pulse response confidence intervals for $G_0(z)$, but $\lambda^N_{r_kr_k}$ in this theorem also contains $R_{e_0}(\tau)$. Generally this knowledge of the second order noise statistics is not available, and therefore in this subsection a procedure is given to reliably estimate the auto-covariance function of the noise from data. An attempt is made to make the procedure as general as possible, i.e. to impose as few restrictions on the data as possible, in order not to unnecessarily restrict the applicability of the uncertainty bounding identification procedure.

Notice that if an exact value for $\Lambda^N_{r_1r_2}$ is not available, a conservative estimate $\widehat{\Lambda}^N_{r_1r_2}$ is still of use. If $\widehat{\Lambda}^N_{r_1r_2} - \Lambda^N_{r_1r_2}$ is positive definite, the resulting probabilistic uncertainty regions as given by Theorem 5.4.3 will be correct, though conservative. This means that the uncertainty regions will then be larger than the uncertainty regions obtainable when using $\Lambda^N_{r_1r_2}$. For the parts (i) and (ii) of Theorem 5.4.3 this is obvious, as the estimated variance is larger than the true one, $\lambda^N_{r_ir_i}$, $i = 1, 2$. For the ellipsoids of part (iii) this can be seen as follows:

Consider Ellipsoid A given by $x^T\Lambda^{-1}x \leq c$ with $\Lambda > 0$ and Ellipsoid B given by $x^T\widehat{\Lambda}^{-1}x \leq c$ with $\widehat{\Lambda} \geq \Lambda$. From the last fact it follows that $\widehat{\Lambda}^{-1} \leq \Lambda^{-1}$, and therefore $x^T\widehat{\Lambda}^{-1}x \leq x^T\Lambda^{-1}x \leq c$, $\forall x \in$ Ellipsoid A. Consequently, any element of Ellipsoid A, the true one, is also an element of Ellipsoid B, the estimated one.

Completely similarly, an estimate $\widehat{\lambda}^N_{r_kr_k}$ which overestimates $\lambda^N_{r_kr_k}$, is a safe estimate, when using this in Theorem 5.5.1. As in fact $\lambda^N_{r_kr_k}$ corresponds to one entry of the matrix $\Lambda^N_{r_1r_2}$, attention will be restricted to the construction of an estimate $\widehat{\Lambda}^N_{r_1r_2}$ for the matrix $\Lambda^N_{r_1r_2}$. The estimation procedure is in fact a multivariable extension of the procedure presented in Section 3.3 in Chapter 3.

Let there be available a measurement sequence generated by (5.2). It is assumed that the measurement is carried out in open loop, i.e. that the input $\{u(t)\}$ is independent of the noise $\{e_0(t)\}$. Let there also be available an ℓ_∞-stable nominal model $\widehat{G}_N(q)$, which has been obtained independently of the given data set, but, for example, by identification based on another data set. Consider the output error $\{\widehat{e}(t)\}$ defined as

$$\widehat{e}(t) := y(t) - \widehat{G}_N(q)u(t) = \psi(t) + e_0(t), \tag{5.29}$$

with

$$\psi(t) := \left(G_0(q) - \widehat{G}_N(q)\right) u(t). \tag{5.30}$$

The idea is to use this output error in order to estimate the second order statistics of the noise process. First, some technical assumptions are made with respect to the input signal $\{u(t)\}$.

Assumption 5.6.1 *The signal $\{u(t)\}$ is a bounded deterministic quasi-stationary signal with an auto-covariance function which is exponentially decaying, i.e. $R_u(\tau) \leq M\rho^\tau$, $\forall \tau$, for certain finite M and $\rho < 1$.*

Next, denote

$$\widehat{R}^N_{\hat{e}}(\tau) := \frac{1}{\widetilde{N} - |\tau|} \sum_{t=t_s}^{N-|\tau|} \hat{e}(t)\hat{e}(t + |\tau|), \tag{5.31}$$

and consider the following estimate for $\Lambda^N_{r_1r_2}$,

$$\widehat{\Lambda}^N_{r_1r_2} = \sum_{\tau=-w(N)}^{w(N)} c_w(\tau) \frac{\widetilde{N} - |\tau|}{\widetilde{N}} \widehat{R}^N_{\hat{e}}(\tau) \begin{bmatrix} R^N_{r_1r_1}(\tau) & R^N_{r_1r_2}(\tau) \\ R^N_{r_1r_2}(\tau) & R^N_{r_2r_2}(\tau) \end{bmatrix}, \tag{5.32}$$

where $c_w(\tau)$ is a window-function, similar to the ones used in spectral analysis, see Ljung (1987, Ch. 6). In particular the window is assumed to satisfy the properties listed in assumption 3.3.2 in Chapter 3. The following theorem states that the given estimate $\widehat{\Lambda}^N_{r_1r_2}$ asymptotically overbounds $\Lambda^N_{r_1r_2}$.

Theorem 5.6.2 *Consider $\Lambda^N_{r_1r_2}$ as given in part (i) of Lemma 5.4.2, and the estimate $\widehat{\Lambda}^N_{r_1r_2}$ defined in (5.32) with a window which satisfies assumption 3.3.2. Suppose that $\{e_0(t)\}$ satisfies assumption 5.2.1, $\{r(t)\}$ assumption 5.2.2 and $\{u(t)\}$ assumption 5.6.1. Moreover, suppose that the input $\{u(t)\}$ is independent of the noise $\{e_0(t)\}$, and additionally satisfies certain weak conditions (in particular conditions (2.11) and (3.14) in Hjalmarsson, 1993a). Also, suppose that the nominal model $\widehat{G}_N(q)$ used in (5.29) has been established independently of the noise $\{e_0(t)\}_{t=t_s,\ldots,N}$. Denote,*

$$\Lambda_\Psi := \begin{bmatrix} \sum_{\tau=-\infty}^{\infty} R_{r_1r_1}(\tau)R_\psi(\tau) & \sum_{\tau=-\infty}^{\infty} R_{r_1r_2}(\tau)R_\psi(\tau) \\ \sum_{\tau=-\infty}^{\infty} R_{r_1r_2}(\tau)R_\psi(\tau) & \sum_{\tau=-\infty}^{\infty} R_{r_2r_2}(\tau)R_\psi(\tau) \end{bmatrix}.$$

Then,

$$\textbf{(i)} \quad \lim_{N\to\infty} \widehat{\Lambda}^N_{r_1r_2} = \Lambda_{r_1r_2} + \Lambda_\Psi \ \ w.p.\ 1,$$

$$\textbf{(ii)} \quad \Lambda_\Psi \geq 0.$$

Proof: See Appendix 5.A. □

The consequence of this result is that the given estimate $\widehat{\Lambda}^N_{r_1r_2}$ can be used to replace the true value $\Lambda^N_{r_1r_2}$ in Theorem 5.4.3 in order to determine frequency response uncertainty regions. Asymptotically correct results are obtained, with respect to both the ellipsoidal and the rectangular confidence regions. The result shows that classical spectral estimation techniques can fruitfully be applied to estimate the variance of an IV or FIR least squares estimate. This has also been shown in Hjalmarsson (1993a, Ch. 3).

Remark 5.6.3 *For finite N the estimate $\widehat{\Lambda}^N_{r_1r_2}$ has a nonzero variance. In fact it is a kind of spectral estimate, for which it has been shown in Ljung (1987, p. 160) that the variance is asymptotically linearly proportional to $\frac{w(N)}{N}$, which tends to zero. Consequently, the variance is negligible if N is large enough. Theoretically, for finite N the estimate $\widehat{\Lambda}^N_{r_1r_2}$ may become non-positive definite. This then indicates that the variance is too large and a better estimate is required.*

In Appendix 5.B Monte Carlo simulations are performed in order to investigate the finite-data behaviour of the estimate $\widehat{\Lambda}^N_{r_1r_2}$. It is also investigated how the accuracy of the identification results are influenced by the fact that the true variance $\Lambda^N_{r_1r_2}$ is replaced by an estimate $\widehat{\Lambda}^N_{r_1r_2}$.

In fact the nonparametric spectral analysis approach is not essential. One may also consider the use of a parametric approach to estimate $\Lambda^N_{r_1r_2}$, which possibly gives less variance, see e.g. Childers (1978). For any consistent identification method, results (i) and (ii) of Theorem 5.6.2 remain to hold.

Remark 5.6.4 *In assumption 5.6.1 it has been assumed that $R_u(\tau)$ is exponentially decaying. This is necessary in order that the auto-covariance function of the output error $\{\widehat{e}(t)\}$ is exponentially decaying. This means that, when determining the second order noise statistics with the procedure described in this subsection, $\{u(t)\}$ is not allowed to contain undecaying deterministic components such as sinusoids. If they are present in the input, they should be detected and removed from both the input signal $\{u(t)\}$ and the output signal $\{y(t)\}$, e.g. by taking the Discrete Fourier Transform, removing the peak-values and taking the inverse DFT of the remaining part. Notice that asymptotically this will not influence the contribution of $\{e_0(t)\}$ to $\{\widehat{e}(t)\}$. To obtain correct results for finite N, the estimate $\widehat{\Lambda}^N_{r_1r_2}$ should be multiplied with $\frac{\widetilde{N}}{N-N_0}$, where N_0 denotes the number of samples of the DFT of the signals $\{u(t)\}$ and $\{y(t)\}$ that have been set to zero.*

Finally, it is mentioned that, if the input signal is (approximately) periodic, the procedure mentioned in Remark 3.3.4 in Chapter 3 can be used to estimate the matrix $\Lambda^N_{r_1r_2}$.

5.6.2 Estimation of Parameter Bounds

In order to be able to derive frequency or pulse response uncertainty regions, prior parameter bounds $\bar{g}(k)$, $k = 0, \ldots, \infty$, should be available, see equation (5.4). In this subsection it is indicated how data may be used to derive these bounds.

Analogously to the derivation of frequency and pulse response confidence regions in the Sections 5.4 and 5.5 it is possible to derive parameter confidence intervals. Almost trivially the $(p+1)$-th estimated IV model parameter $\widehat{g}(p)$, $p \leq n$, is given by

$$\widehat{g}(p) = \sum_{k=0}^{n} \widehat{g}(k)\delta_p(k) = [\delta_p(0) \ \ldots \ \delta_p(n)] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t) \right]^{-1} \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)y(t),$$

where

$$\delta_p(k) := \begin{cases} 0, & k \neq p, \\ 1, & k = p. \end{cases}$$

Define the signal $r_p(t)$ as

$$r_p(t) := [\delta_p(0) \ \ldots \ \delta_p(n)] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t) \right]^{-1} V(t), \tag{5.33}$$

which depends on p, the parameter under consideration. Anologously to the derivation in Section 5.4 and with $a(t)$, $b(t)$, $d(p)$ and $f(p)$ defined by (5.15), (5.16), (5.18) and (5.19) respectively, it can be derived that

$$\widehat{g}(p) = \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} r_p(t)y(t) = g_0(p) + \frac{1}{\widetilde{N}} \left(d(p) + f(p) + \sum_{t=t_s}^{N} r_p(t)e_0(t) \right).$$

The following proposition, specifying parameter confidence regions, is straightforwardly obtained.

Proposition 5.6.5 *Consider the IV estimate (5.11). Suppose that $\{e_0(t)\}$ and $\{r(t)\}$ are independent and that they satisfy the assumptions 5.2.1 and 5.2.2 respectively. Let $\bar{d}(p)$ and $\bar{f}(p)$ be given by (5.21) and (5.22) respectively, with $r_p(t)$ given by (5.33).*

Let $c_{\mathcal{N},\alpha}$ correspond to a probability α in the standard Normal distribution, such that, if $x \in \mathcal{N}(0,1) \Rightarrow \text{prob}(|x| \leq c_{\mathcal{N},\alpha}) = \alpha$. Denote,

$$\lambda^N_{r_p r_p} := \sum_{\tau=-N+t_s}^{N-t_s} \frac{\widetilde{N} - |\tau|}{\widetilde{N}} R^N_{r_p}(\tau) R_{e_0}(\tau).$$

Then, if $N \to \infty$,

$$|\widehat{g}(p) - g_0(p)| \leq c_{\mathcal{N},\alpha} \sqrt{\frac{\lambda^N_{r_p r_p}}{\widetilde{N}}} + \frac{\bar{d}(p)}{\widetilde{N}} + \frac{\bar{f}(p)}{\widetilde{N}}, \ \ w.p. \geq \alpha.$$

Proof: Similar to the proof of part (i) of Theorem 5.4.3, replacing $r_1(t)$ by $r_p(t)$, and noting that the contribution $\delta(1)$ is not present. □

The procedure of Subsection 5.6.1 can be used to estimate the variance $\lambda^N_{r_p r_p}$, in order to be able to actually compute the parameter confidence intervals.

The result may seem to bite its own tail, as the parameter confidence intervals can only be derived if prior bounds $\bar{g}(k)$ are available. However, the result is still useful, if properly combined with an iterative procedure. The following procedure is proposed. First, choose conservative bounds $\bar{g}(k)$, with exponential decay rate according to requirement (5.5). Then compute parameter confidence intervals as given in the proposition above, where it is important to choose a relatively large value of the truncation parameter n. Typically these parameter bounds will be a lot smaller than the prior bounds, especially for low values of k.

Next, it is proposed to establish new prior bounds which are consistent with the confidence intervals for the first $n^* + 1$ parameters, for some $n^* < n$, but which are not necessarily larger than the identified bounds for $g_0(k)$, $k > n^*$. The motivation for this is that generally the tail parameters $g_0(k)$ are relatively small, especially if proper basis functions $B_k(z)$ have been chosen. At the same time this tail only contributes little to the input-output behaviour and, consequently, the tail parameter estimates have relatively large variances, compared to the true values. Therefore it is not considered to be necessary that the new prior parameter bounds for the tail are consistent with the identified probabilistic bounds.

There is still freedom in the choice of n^*. It is proposed to choose n^* such that the new prior bounds are only smaller than the identified bounds if 0 is contained in the identified confidence interval. In other words, it is proposed to choose n^* such that the new prior bounds are smaller than the identified bounds only for those values of k where the parameter uncertainty is larger than the true parameter value $g_0(k)$.

Although there is no guarantee that correct results will be obtained, the procedure described above appears to yield good results in practice. In De Vries (1994) basically the same procedure is used and good results are reported there as well. Note that the procedure may be carried out more than once in order to establish prior bounds $\bar{g}(k)$. The procedure then becomes an iteration of alternately identifying parameter confidence intervals and establishing new prior parameter bounds. The procedure is also reasonably robust in the sense that if the prior bounds are chosen too small, this may be detected by comparing with the parameter confidence intervals which are estimated next. In Section 5.9 the procedure to establish prior bounds $\bar{g}(k)$ is illustrated by means of an example.

5.7 Survey

In the previous sections a complete procedure has been developed to identify probabilistic frequency and pulse response uncertainty regions. In this section a survey is given of all the steps in this uncertainty bounding identification procedure. In the survey attention is restricted to the identification of frequency response uncertainty regions. With some minor changes the survey also applies to the identification of pulse response uncertainty regions, but this is left to the reader.

Altogether the procedure to identify frequency response confidence regions consists

of the following steps:

1. Obtain one or more measurement sequences $\{u(t)\}$, $\{y(t)\}$, $t = 1, \ldots, N$. If the input is generated in closed loop also an external reference signal $\{\bar{r}(t)\}$ should be available, which is independent of the noise process $\{e_0(t)\}$. However, at least one open loop measurement sequence should be available, in order to be able to carry out step 2. If possible, the measurements should be prefiltered in order to obtain (approximately) white output noise. This will give uncertainty regions of minimal size.

2. Use one or more open loop measurement sequences to determine an estimate of the noise auto-covariance function, $\widehat{R}_\varepsilon^N(\tau)$, $\tau = 0, \ldots, w$, given by (5.31). If the input $\{u(t)\}$ contains one or more sinusoids, the procedure mentioned in Remark 5.6.4 should be used. Note that here also a measurement sequence may be used where the input had not been excited, such that the measured output is identical to a noise realization.

3. Choose for each measurement sequence to be used in the identification a suitable signal $\{r(t)\}$. For an open loop measurement $r(t) = u(t)$ is suitable. In the closed loop situation $r(t) = F(q)\bar{r}(t)$ is a suitable choice, where $F(q)$ is equal to an estimate of the transfer function between $\{\bar{r}(t)\}$ and $\{u(t)\}$. Note that here only measurements should be used where the process had been excited. If more than one measurement sequence is available, the steps 4 till 14 should be carried out for each measurement sequence separately.

4. Choose a value for $\bar{u}$, an upper bound on the input signal in the past. This need not be a very tight bound, as its influence on the identification result is limited.

5. Choose a value for t_s. As t_s is chosen larger, the influence of the unknown initial conditions decreases, but also less data are used in the identification.

6. Choose a suitable basis, i.e. specify the basis functions $B_k(z)$, $k = 0, \ldots, \infty$, see (5.1). This actually corresponds to incorporating approximate knowledge about the pole-locations of the system $G_0(z)$ in the identificaton procedure. Note that the choice of basis functions does not determine whether or not the resulting confidence regions are correct. But a proper choice can mean a considerable reduction of the size of the uncertainty regions. In Chapter 8 of this thesis more is said about the choice of basis functions.

7. Choose bounds $\bar{g}(k)$, $k = 0, \ldots, \infty$, possibly with the procedure described in Subsection 5.6.2. Note that this may be done iteratively. In a first iteration step conservative bounds $\bar{g}(k)$ may be specified, and later they may be tightened by inspecting the identification result.

8. Choose a truncation value n, which is a trade-off between bias and variance. The optimal truncation value n can be determined by making several choices and evaluating which value for n yields the smallest uncertainty regions.

9. Calculate the signals $w_k(t)$ and $v_k(t)$, $k = 0, \ldots, n$, given by (5.8) and (5.9) respectively, and calculate the matrix $\frac{1}{N}\sum_{t=t_s}^{N} V(t)U^T(t)$ with $U(t)$ and $V(t)$ given by (5.10).

10. Choose frequencies ω_j, $j = 1, \ldots, l$, for which $G_0(e^{i\omega_j})$ is desired to be evaluated. The steps 11 till 14 need to be carried out for each frequency ω_j separately.

11. Calculate the signals $r_1(t)$ and $r_2(t)$ as given by (5.13) and (5.14). And calculate $\widehat{G}(e^{i\omega_j})$ as given by (5.17).

12. Calculate $\bar{d}(p)$ and $\bar{f}(p)$, $p = 1, 2$, given by (5.21) and (5.22) respectively. Also, calculate $\delta(p)$, $p = 1, 2$, given by (5.23) and (5.24).

13. Calculate $R^N_{r_i r_j}(\tau)$, $i, j = 1, 2$, $\tau = 0, \ldots, w$, defined in Lemma 5.4.2. Calculate the covariance matrix $\widehat{\Lambda}^N_{r_1 r_2}$ as defined in (5.32). The choice of window-parameters f and w is a trade-off between bias and variance of the estimate.

14. Determine a confidence level and calculate frequency response uncertainty regions as given by Theorem 5.4.3. If more than 1 measurement sequence is available, probably the easiest thing to do, is to estimate rectangular uncertainty regions for each measurement sequence separately, and then calculate the intersection of these regions. This automatically gives the nice possibility to check if the priors have been chosen suitably: If the bounds $\bar{g}(k)$ have been chosen too small, or the covariance matrix $\Lambda^N_{r_1 r_2}$ has not been estimated properly, the confidence regions (corresponding to a high confidence level) need not intersect. In that case the steps 7 and/or 13 need to be reconsidered.

5.8 The Multivariable Case

In this section multivariable extensions to the procedure for probabilistic uncertainty bounding identification are presented. Only the main differences with the SISO case are highlighted, most technical details are omitted. Note that the MIMO extensions are quite similar to those in Section 4.6.

5.8.1 Identification Setting

Consider a LTI, discrete-time, causal and ℓ_∞-stable MIMO system $G_0(z)$ with q' inputs and p' outputs. This system can be represented by

$$G_0(z) = \begin{bmatrix} G_0^{11}(z) & \cdots & G_0^{1q'}(z) \\ \vdots & \ddots & \vdots \\ G_0^{p'1}(z) & \cdots & G_0^{p'q'}(z) \end{bmatrix},$$

$$G_0^{i'j'}(z) = \sum_{k=0}^{\infty} g_0^{i'j'}(k) B_k^{i'j'}(z),$$

where $\{B_k^{i'j'}(z)\}_{k=0,\ldots,\infty}$ is some specified set of basis functions for the SISO transfer function $G_0^{i'j'}(z)$, completely similar to the representation used in Section 5.2.

Consider given vector-valued input data $\{u(t)\}$ and measured vector-valued output data $\{y(t)\}$. Let the input-output relation of the data generating system be given by (5.2), where $\{e_0(t)\}$ is an unknown additive vector-valued output noise.

A q'-dimensional signal $\{r(t)\}_{t=1,\ldots,N}$ is assumed available, which is highly correlated with the input signal $\{u(t)\}$, but independent of the noise process $\{e_0(t)\}$. In particular the j'th entry of the signal $\{r(t)\}$, i.e. $\{r_{j'}(t)\}$, should be highly correlated with $\{u_{j'}(t)\}$.

The assumptions about the noise process are still specified by assumption 5.2.1, but now for each entry of the noise signal separately, i.e. for each $\{e_0^{i'}(t)\}$, $i' = 1,\ldots,p'$. Similarly, assumption 5.2.2 applies to each signal entry $\{r_{j'}(t)\}$, $j' = 1,\ldots,q'$ separately.

The following bound is assumed for the input signal in the past,

$$|u_{j'}(t)| \leq \bar{u}_{j'}, \ \forall t \leq 0, \ j' = 1,\ldots,q',$$

and for the generalized pulse response parameters $g_0^{i'j'}(k)$,

$$|g_0^{i'j'}(k)| \leq \bar{g}_{i'j'}(k), \ k = 0,\ldots,\infty, \ i' = 1,\ldots,p', \ j' = 1,\ldots,q'.$$

The identification objective is to derive probabilistic uncertainty regions for the frequency and (weighted) pulse response of each $G_0^{i'j'}(z)$, $i' = 1,\ldots,p'$, $j' = 1,\ldots,q'$. Again this is done by explicitly calculating bias and variance expressions of an IV estimate.

5.8.2 The Instrumental Variable Estimate

The model $G(z)$ is parametrized as

$$G(z) = \begin{bmatrix} G_{11}(z) & \cdots & G_{1q'}(z) \\ \vdots & \ddots & \vdots \\ G_{p'1}(z) & \cdots & G_{p'q'}(z) \end{bmatrix},$$

$$G_{i'j'}(z) = \sum_{k=0}^{n_{i'j'}} g_{i'j'}(k) B_k^{i'j'}(z).$$

Due to this parametrization the MIMO IV identification problem can be split up into p' independent MISO identification problems.

Define $\widetilde{u}(t)$ as in (5.7) and define $\forall i', j'$,

$$w_k^{i'j'}(t) := B_k^{i'j'}(q)\widetilde{u}_{j'}(t), \ k = 0,\ldots,n_{i'j'},$$

and

$$v_k^{i'j'}(t) := B_k^{i'j'}(q) r_{j'}(t), \ k = 0,\ldots,n_{i'j'}.$$

The model input-output relation for the i'th output signal entry is given by

$$y_{i'}(t) = \sum_{j'=1}^{q'} G_{i'j'}(q)\widetilde{u}_{j'}(t) + e_{i'}(t) = \sum_{j'=1}^{q'} \sum_{k=0}^{n_{i'j'}} g_{i'j'}(k) B_k^{i'j'}(q)\widetilde{u}_{j'}(t) + e_{i'}(t) =$$
$$= \sum_{j'=1}^{q'} \sum_{k=0}^{n_{i'j'}} g_{i'j'}(k) w_k^{i'j'}(t) + e_{i'}(t),$$

where $e_{i'}(t)$ is the output error.

Define the matrices

$$U_{i'j'}(t) := \begin{bmatrix} w_0^{i'j'}(t) \\ \vdots \\ w_{n_{i'j'}}^{i'j'}(t) \end{bmatrix}, \quad V_{i'j'}(t) := \begin{bmatrix} v_0^{i'j'}(t) \\ \vdots \\ v_{n_{i'j'}}^{i'j'}(t) \end{bmatrix},$$

$$U_{i'}(t) := \begin{bmatrix} U_{i'1}(t) \\ \vdots \\ U_{i'q'}(t) \end{bmatrix}, \quad V_{i'}(t) := \begin{bmatrix} V_{i'1}(t) \\ \vdots \\ V_{i'q'}(t) \end{bmatrix}.$$

Then the basic IV estimate is given by

$$\begin{bmatrix} \widehat{g}_{i'1}(0) \\ \vdots \\ \widehat{g}_{i'1}(n_{i'1}) \\ \widehat{g}_{i'2}(0) \\ \vdots \\ \widehat{g}_{i'q'}(n_{i'q'}) \end{bmatrix} = \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V_{i'}(t) U_{i'}^T(t) \right]^{-1} \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V_{i'}(t) y_{i'}(t), \; i' = 1, \ldots, p'.$$

In case of open loop operation, $r(t) = u(t)$, this corresponds to a FIR least squares estimate for general basis functions. The estimated IV model is given by

$$\widehat{G}_{i'j'}(z) = \sum_{k=0}^{n_{i'j'}} \widehat{g}_{i'j'}(k) B_k^{i'j'}(z).$$

5.8.3 Frequency Response Uncertainty Regions

Similar to the SISO case an analysis can be made of the frequency response bias and variance errors of the IV estimate. This is done for each transfer function $G_0^{i'j'}(z)$ separately.

Consider a given frequency ω_j. The frequency response of the IV model is given by

$$\widehat{G}_{i'j'}(e^{i\omega_j}) = \sum_{k=0}^{n_{i'j'}} \widehat{g}_{i'j'}(k) B_k^{i'j'}(e^{i\omega_j}) =$$

$$= \left[0 \cdots 0\, B_0^{i'j'}(e^{i\omega_j}) \cdots B_{n_{i'j'}}^{i'j'}(e^{i\omega_j})\, 0 \cdots 0\right] \begin{bmatrix} \widehat{g}_{i'1}(0) \\ \vdots \\ \widehat{g}_{i'1}(n_{i'1}) \\ \widehat{g}_{i'2}(0) \\ \vdots \\ \widehat{g}_{i'q'}(n_{i'q'}) \end{bmatrix}.$$

Substitution of the IV parameter estimate and introduction of the scalar signals

$$\begin{aligned} r_1^{i'j'}(t) &:= \left[0 \cdots 0 \operatorname{Re}\left(B_0^{i'j'}(e^{i\omega_j})\right) \cdots \operatorname{Re}\left(B_{n_{i'j'}}^{i'j'}(e^{i\omega_j})\right) 0 \cdots 0\right] \cdot \\ &\quad \cdot \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V_{i'}(t) U_{i'}^T(t)\right]^{-1} V_{i'}(t), \\ r_2^{i'j'}(t) &:= \left[0 \cdots 0 \operatorname{Im}\left(B_0^{i'j'}(e^{i\omega_j})\right) \cdots \operatorname{Im}\left(B_{n_{i'j'}}^{i'j'}(e^{i\omega_j})\right) 0 \cdots 0\right] \cdot \\ &\quad \cdot \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V_{i'}(t) U_{i'}^T(t)\right]^{-1} V_{i'}(t), \end{aligned}$$

yields

$$\widehat{G}_{i'j'}(e^{i\omega_j}) = \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} \left(r_1^{i'j'}(t) + i r_2^{i'j'}(t)\right) y_{i'}(t).$$

The output $y_{i'}(t)$ can be written as

$$y_{i'}(t) = \sum_{j'=1}^{q'} G_0^{i'j'}(q) u_{j'}(t) + e_0^{i'}(t) = \sum_{j'=1}^{q'} \sum_{k=0}^{n_{i'j'}} g_0^{i'j'}(k) w_k^{i'j'}(t) + a_{i'}(t) + b_{i'}(t) + e_0^{i'}(t),$$

where $a_{i'}(t)$ represents the response of the tail of the system and $b_{i'}(t)$ represents the response to past input signals. They are defined similar to the signals $a(t)$ and $b(t)$ in (5.15) and (5.16) in the SISO case.

Using this the following expression is easily derived,

$$\widehat{G}_{i'j'}(e^{i\omega_j}) =$$

$$= \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} \left(r_1^{i'j'}(t) + i r_2^{i'j'}(t)\right) \left(\sum_{j'=1}^{q'} \sum_{k=0}^{n_{i'j'}} g_0^{i'j'}(k) w_k^{i'j'}(t) + a_{i'}(t) + b_{i'}(t) + e_0^{i'}(t)\right) =$$

$$= G_0^{i'j'}(e^{i\omega_j}) - \sum_{k=n_{i'j'}+1}^{\infty} \widehat{g}_0^{i'j'}(k) B_k^{i'j'}(e^{i\omega_j}) +$$

$$+ \frac{1}{\widetilde{N}} \left(d_{i'j'}(1) + i d_{i'j'}(2) + f_{i'j'}(1) + i f_{i'j'}(2) + \sum_{t=t_s}^{N} \left(r_1^{i'j'}(t) + i r_2^{i'j'}(t) \right) e_0^{i'}(t) \right),$$

where, for $p = 1, 2$,

$$d_{i'j'}(p) = \sum_{t=t_s}^{N} r_p^{i'j'}(t) a_{i'}(t), \;\; f_{i'j'}(p) = \sum_{t=t_s}^{N} r_p^{i'j'}(t) b_{i'}(t).$$

Similar to the derivation in Section 5.4 the terms $d_{i'j'}(p)$ and $f_{i'j'}(p)$ can be bounded, using the prior parameter bounds $\bar{g}_{i'j'}(k)$ and the input signal bound $\bar{u}_{i'}$. Moreover, Lemma 5.4.2 specifies a distribution for the term containing $e_0^{i'}(t)$. The combination of these results leads to a probabilistic error bound for the frequency response of the IV model, entirely similar to Theorem 5.4.3. This corresponds to a confidence region for the system's frequency response $G_0^{i'j'}(e^{i\omega_j})$. In this way uncertainty regions can be constructed for each i', j' and frequency ω_j. As explained in Section 4.6 this ultimately leads to a structured uncertainty description, suited for use in μ-robust controller analysis and synthesis procedures.

5.8.4 Pulse Response Uncertainty Regions

Pulse response parameter bounds can be identified for each transfer function $G_0^{i'j'}(z)$ separately.

Consider the IV model pulse response parameters $\widehat{p}_{i'j'}(k)$ defined by the relation,

$$\sum_{k=0}^{\infty} \widehat{p}_{i'j'}(k) z^{-k} = \widehat{G}_{i'j'}(z) W_{i'j'}(z),$$

and the system pulse response parameters $p_0^{i'j'}(k)$ defined by the relation,

$$\sum_{k=0}^{\infty} p_0^{i'j'}(k) z^{-k} = G_0^{i'j'}(z) W_{i'j'}(z),$$

for any user-defined finite dimensional, stable and stably invertible, LTI weighting $W_{i'j'}(z)$.

Define

$$S_{k'}^{i'j'}(z) := \sum_{k=0}^{\infty} s_{k'}^{i'j'}(k) z^{-k} := B_{k'}^{i'j'}(z) W_{i'j'}(z), \;\; k' = 0, \ldots, \infty,$$

which implies that, for $k = 0, \ldots, \infty$,

$$\widehat{p}_{i'j'}(k) = \sum_{k'=0}^{n_{i'j'}} \widehat{g}_{i'j'}(k') s_{k'}^{i'j'}(k) = \left[0 \cdots 0\, s_0^{i'j'}(k) \cdots s_{n_{i'j'}}^{i'j'}(k)\, 0 \cdots 0\right] \begin{bmatrix} \widehat{g}_{i'1}(0) \\ \vdots \\ \widehat{g}_{i'1}(n_{i'1}) \\ \widehat{g}_{i'2}(0) \\ \vdots \\ \widehat{g}_{i'q'}(n_{i'q'}) \end{bmatrix}.$$

Substitution of the IV parameter estimate in this expression yields, after some calculations,

$$\widehat{p}_{i'j'}(k) = \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} r_k^{i'j'}(t) y_{i'}(t) =$$

$$= p_0^{i'j'}(k) - \sum_{k'=n_{i'j'}+1}^{\infty} \widehat{g}_0^{i'j'}(k') s_{k'}^{i'j'}(k) + \frac{1}{\widetilde{N}} \left(d_{i'j'}(k) + f_{i'j'}(k) + \sum_{t=t_s}^{N} r_k^{i'j'}(t) e_0^{i'}(t) \right),$$

where

$$r_k^{i'j'}(t) := \left[0 \cdots 0\, s_0^{i'j'}(k) \cdots s_{n_{i'j'}}^{i'j'}(k)\, 0 \cdots 0\right] \left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V_{i'}(t) U_{i'}^T(t) \right]^{-1} V_{i'}(t),$$

and

$$d_{i'j'}(k) = \sum_{t=t_s}^{N} r_k^{i'j'}(t) a_{i'}(t), \quad f_{i'j'}(k) = \sum_{t=t_s}^{N} r_k^{i'j'}(t) b_{i'}(t),$$

with the signals $a_{i'}(t)$, $b_{i'}(t)$ as introduced in the previous subsection.

Next, probabilistic error bounds for the parameters $\widehat{p}_{i'j'}(k)$ are straightforwardly obtained along the lines of Theorem 5.5.1. This then provides confidence intervals for the system parameters $\widehat{p}_0^{i'j'}(k)$. As explained in Section 4.6 these parameter bounds can be used to obtain a structured uncertainty description, suited for use in ℓ_1-robust controller analysis or design schemes.

5.9 Examples

5.9.1 Example 1: A Simple Example

In this subsection the performance of the procedure for probabilistic uncertainty bounding identification is illustrated by applying it to the example in Subsection 4.7.1. This then allows for a comparison with the uncertainty bounding procedure of Chapter 4.

The system $G_0(z)$ and the data $u(t)$ and $y(t)$ are as specified in Subsection 4.7.1. Again, a system representation in terms of the standard pulse functions is used. For

a fair comparison also in this case exact noise knowledge is used to construct the frequency response uncertainty regions. This means that the auto-covariance function of the noise process is assumed to be known. As the noise process is white, only the variance of the noise need to be specified. This variance is equal to $\frac{1}{1200}$. The other prior information assumed available, is the bound on the input signal in the past $\bar{u} = 1$ and the parameter bound $\bar{g}(k) = M\rho^k$, with $M = 2$ and $\rho = \frac{1}{1.2}$.

For 40 frequencies logarithmically distributed between 0.01 and π rectangular frequency response 3σ-confidence regions $\mathcal{P}(\omega_j)$ have been calculated. This has been done for various values of t_s and n in order to investigate which values are optimal, i.e. yield the smallest confidence regions. The values $n = 39$ and $t_s = 40$ finally appear to be optimal. The resulting confidence regions are shown in a Nyquist diagram, Figure 5.1.

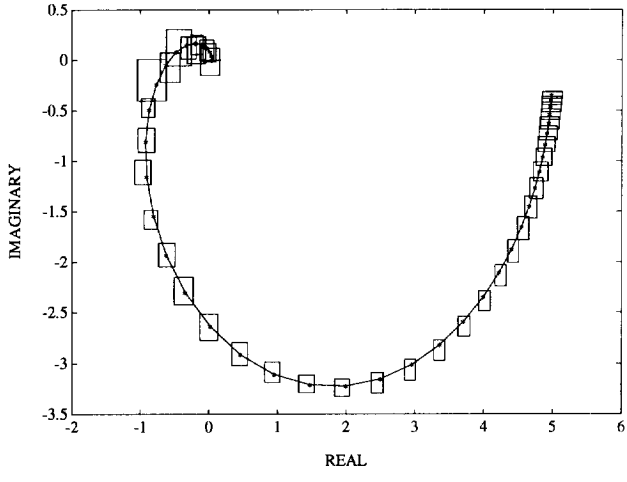

Fig. 5.1: Nyquist diagram frequency response 3σ-confidence regions $\mathcal{P}(\omega_j)$ (rectangulars) and system's frequency response $G_0(e^{i\omega_j})$ (solid,$*$).

A comparison with Figure 4.2 shows that the confidence regions identified here are approximately a factor 1.7 smaller than the uncertainty regions identified with the procedure of Chapter 4, with respect to both the real and the imaginary part of the frequency response. It appears that the probabilistic approach in this chapter performs better than the deterministic approach in chapter 4. This despite of the fact that the procedure for deterministic identification is optimal, i.e. yields the smallest possible uncertainty regions on the basis of the measured data and specified deterministic prior information. And also despite of the fact that the simulation example is more or less optimally tuned towards the procedure for deterministic identification, as a uniformly distributed noise has been used. Apparently even then a stochastic noise description is more favourable than a deterministic noise description. In case the noise process satisfies both types of noise priors, information about averaging properties of the noise

appears to be much more powerful than an amplitude bound on the noise.

5.9.2 Example 2: A More Complicated Example

In this subsection it is shown by means of a more complicated example how the procedure for probabilistic uncertainty bounding identification can be used in practice. The example is taken from Subsection 4.7.2 in the previous chapter. Again, the results will be compared to the results of the procedure for deterministic identification.

The system $G_0(z)$, data $u(t)$, $y(t)$ and basis-generating model $\widehat{G}(z)$ are as specified in Subsection 4.7.2. Again, in first instance the prior parameter bounds are chosen conservatively,

$$\bar{g}(k) = M\rho^k, \ k = 0, \ldots, \infty,$$

with $M = 10$ and $\rho = 0.9$. Later on these bounds are tightened using insight obtained from the data.

By design the initial conditions are 0, so $\bar{u} = 0$, and t_s is chosen equal to 1. As the data are generated in open loop, the signal $\{r(t)\}$ is chosen to be equal to $\{u(t)\}$. The first 500 measured output samples are used to estimate the auto-covariance function of the noise process. In the sequel the variance estimates make use of a Tukey-window, as specified by (5.34) in Appendix 5.B, with window-parameters $f = 0.6$ and $w = 100$.

With Proposition 5.6.5 stochastic bounds for $g_0(k)$ have been computed, with $n = 45$ and $c_{\mathcal{N},0.99} = 2.58$. The results are depicted in Figure 5.2. The bounds $\pm\bar{g}(k)$ are

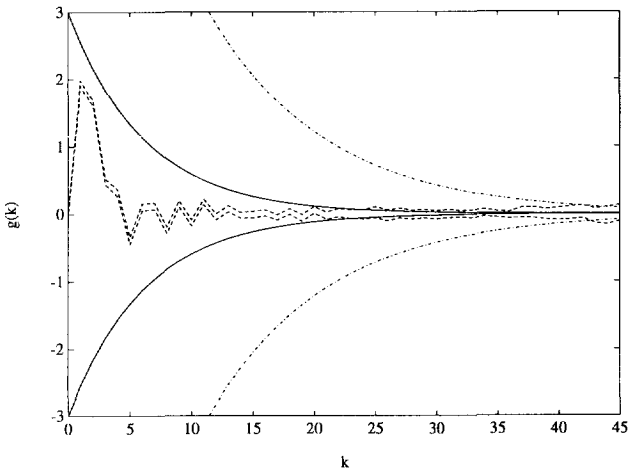

Fig. 5.2: Prior parameter bounds $\pm\bar{g}(k)$ (dash-dotted), identified bounds $g_{\mathrm{l}}(k)$, $g_{\mathrm{u}}(k)$ (dashed) and new prior bounds $\pm\bar{g}(k)$ (solid).

shown, as well as the calculated confidence intervals. Following the procedure proposed in Subsection 5.6.2, new and tighter prior bounds $\bar{g}(k)$ have been chosen, such that

they are consistent with the confidence intervals for the first 24 parameters. The new prior bounds are smaller than the confidence intervals of $g_0(k)$ for $k \geq 24$, where 0 is in the confidence intervals. The resulting new bounds, corresponding to $M = 3$ and $\rho = 0.85$, are also depicted in Figure 5.2. In fact they appear to be correct, the new prior parameter bounds are satisfied by the system $G_0(z)$.

With Theorem 5.4.3 probabilistic frequency response uncertainty regions are calculated for 32 frequencies between 0 and π. Both ellipsoidal confidence regions, corresponding to $c_{\chi,0.999} = 13.815$, and rectangular confidence regions, corresponding to $c_{\mathcal{N},0.9995} = 3.5$, are calculated. Consequently, for each frequency both the ellipsoidal and the rectangular confidence regions are correct with probability larger than 0.999 (for the rectangulars this follows by application of Bonferroni's inequality). And the probability that the confidence regions are correct for all 32 frequencies at the same time is larger than $1 - 32 \cdot 0.001 = 0.97$. The choice $n = 39$ appears to yield the smallest uncertainty regions. The identified confidence regions are depicted in Figure 5.3, together with the system's frequency response $G_0(e^{i\omega_j})$.

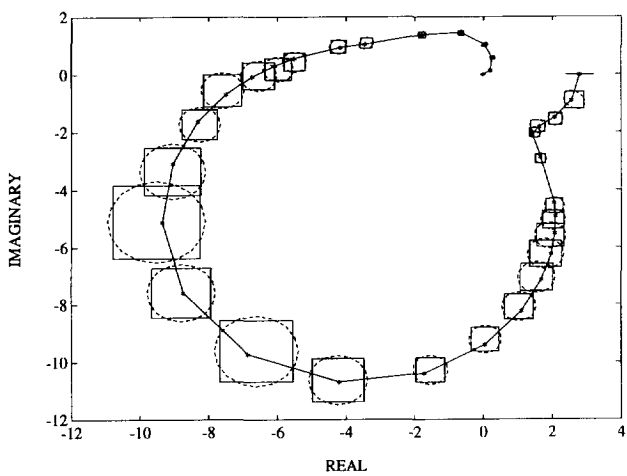

Fig. 5.3: Nyquist diagram identified 99.9%-confidence regions for the system's frequency response (rectangulars, ellipsoids) and system's frequency response $G_0(e^{i\omega_j})$ (solid,$*$).

Moreover, with the procedure of Section 5.5, 3.5σ-confidence regions are calculated for the first 100 pulse response parameters of the system $G_0(z)$ (the weighting $W(z)$ is chosen equal to 1). These pulse response uncertainty regions are depicted in Figure 5.4, together with the system's pulse response.

It appears that in this case the confidence regions are correct, i.e. the system is within the identified bounds, both in Figure 5.3 and in Figure 5.4. Moreover, a comparison with Figure 4.5 and Figure 4.6 shows that the deterministic uncertainty regions, identified with the method of Chapter 4, are much larger (about a factor

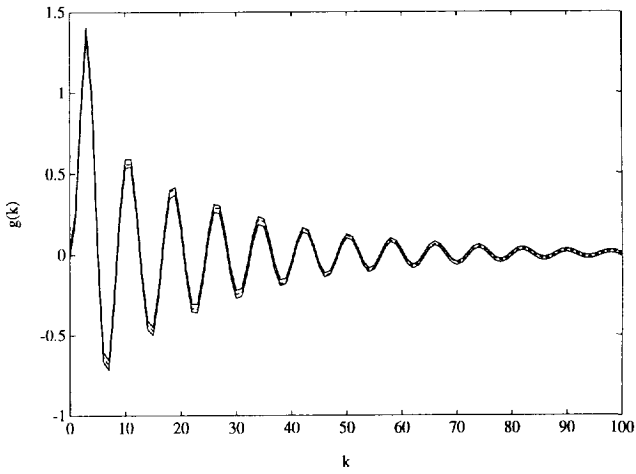

Fig. 5.4: Identified 99.9%-confidence regions $[p_l(k), p_u(k)]$ for the system's pulse response (solid) and system's pulse response $p_0(k)$ (dashed).

4) than the ones established here. The probabilistic identification approach of this chapter simply outperforms the deterministic approach of Chapter 4.

In a similar way as in Subsection 4.7.2, the identified frequency response confidence regions can be used to establish an upper bound on the H_∞-norm of the model error. On the basis of the rectangular uncertainty regions, the additive model error bound $\delta_u(\omega_j)$ is calculated with respect to the model $\widehat{G}(z)$, which is specified in Subsection 4.7.2. The result is depicted in Figure 5.5. Neglecting the influence of the intersample frequencies, this yields an H_∞-model error bound equal to 15.0. This should be compared to the H_∞-norm of the model error, which is 13.7, and the bound found in Subsection 4.7.2, which equals 20.3.

Also, in a similar way as in Subsection 4.7.2 the identified pulse response uncertainty regions can be used to establish an upper bound on the ℓ_1-norm of the additive model error. Figure 5.6 depicts the pulse response error bound $\max[p_u(k) - \widehat{p}(k), \widehat{p}(k) - p_l(k)]$. This yields an ℓ_1-model error bound equal to 20.4. This should be compared to the ℓ_1-norm of the model error, which is 18.9, and the bound found in Subsection 4.7.2, which equals 23.9.

Obviously, the model $\widehat{G}(z)$ is a bad nominal model, the greater part of the H_∞- and ℓ_1-model error bound is due to undermodelling or bias. Only a smaller portion is due to noise, the variance error. If only this variance error is taken into consideration, the uncertainty bounding identification procedure developed in this chapter performs about a factor 4 better than the procedure for deterministic uncertainty bounding identification as described in the previous chapter.

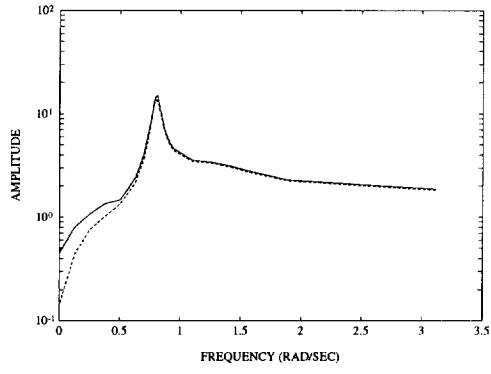

Fig. 5.5: Additive frequency response model error bound $\delta_u(\omega_j)$ (solid) and additive model error $\left|\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right|$ (dashed).

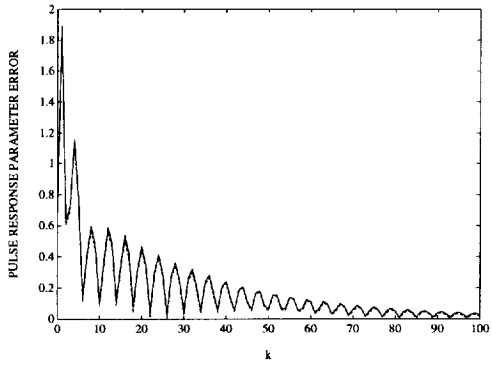

Fig. 5.6: Additive pulse response model error bound $\max[p_u(k) - \widehat{p}(k), \widehat{p}(k) - p_l(k)]$ (solid) and additive model error $|\widehat{p}(k) - p_0(k)|$ (dashed).

5.10 Discussion

In this chapter an identification procedure has been developed which yields confidence regions for the frequency or pulse response of some stable LTI system. The procedure involves the explicit calculation of bias and variance errors of an IV or FIR least squares estimate. Important features of the identification procedure are:

- Essentially the identification approach is stochastic. Probabilistic uncertainty regions and model error bounds are calculated based on data, deterministic system priors and stochastic noise priors.

- The actual computations can be performed quite efficiently. No nonlinear optimizations are involved, as use is made of a linear system parametrization and, consequently, there is no problem with local optima.

- The required prior information can be reliably estimated from data.

- There are no restrictions on the input signal, it need, for example, not be periodic. It is even not necessary that the input is generated in open loop.

- No order assumption about the system is made.

- The procedure is applicable to MIMO systems.

- Rough prior knowledge about the system, or more specifically pole-locations, can be incorporated by using generalized orthonormal basis functions.

- Unknown initial conditions are properly taken into account.

- The resulting uncertainty regions appear to be much smaller than the ones identified with the uncertainty bounding identification procedure developed in Chapter 4. This is not only true for the specific simulation examples considered in this and the previous chapter, but it is a typical behaviour. It is due to the fact that the probabilistic identification approach much better utilizes the basic property of the noise process, the fact that noise is uncorrelated to the input signal (or external reference signal).

  The incorporation of cross-covariance noise constraints in the deterministic identification approach certainly means an improvement compared to the use of time-domain amplitude noise constraints. Nevertheless, similar to the time-domain noise bounds, the cross-covariance noise bounds have to be chosen conservative in order to guarantee their correctness. Therefore the procedure for deterministic identification yields larger system uncertainty regions than the procedure outlined in this chapter, in which there is no need to specify conservative bounds.

- Under certain asymptotic conditions (persistent excitation; $N, n \to \infty$) the identification procedure is consistent, i.e. the resulting uncertainty regions converge to one point, the system $G_0(z)$, w.p. 1. This is due to the fact that a least squares FIR estimate is consistent, in the sense that its bias and variance errors vanish if $N, n \to \infty$, see for example Wahlberg (1991) and Ljung and Wahlberg (1992).

- The identification procedure is robust for noise outliers and small errors in the prior information. This means, for example, that if the system has a small nonlinearity (measured in terms of its ℓ_∞-induced norm), the resulting uncertainty regions are just slightly erratic and, consequently, are still (approximately) valid.

On the other hand some drawbacks of the uncertainty bounding identification procedure developed in this chapter, are:

- Although all computations can be carried out efficiently and accurately, the identification procedure requires a lot of computations. This means that, for the time being, on-line application of the procedure seems infeasible.
- The procedure makes use of results which are asymptotic in the number of data. As in applications there are always finite-data records, the results might not be valid in practice. On the other hand, Monte Carlo simulations (Appendix 5.B) show that the error caused by the finiteness of the number of data can be very small, even for small values of N. The acccuracy of the finite-data approximation depends on several factors, such as the length of the pulse response of the noise generating filter and the actual distribution of the noise process.
- The two most severe prior assumptions made, are linearity of the system, and stationarity of the noise process. In practical applications both are not satisfied. On the other hand, the first assumption is often approximately satisfied if the input excitation is such that the system operates in one working point. The second assumption is approximately satisfied if the data are pre-processed such that non-stationary disturbances (such as offsets) are removed beforehand. Moreover, the identification procedure is robust for small errors that are left.

5.A Proofs

Proof of Lemma 5.4.2

The Lemma's 9A.1 and 9A.2 of Ljung (1987) are applied to

$$S_N = \frac{1}{\sqrt{\widetilde{N}}} \sum_{t=t_s}^{N} \begin{bmatrix} r_1(t)e_0(t) \\ r_2(t)e_0(t) \end{bmatrix}.$$

First, it is noted that $\{r_1(t)\}$ and $\{r_2(t)\}$ are bounded deterministic quasi-stationary, as $\{r(t)\}$ is bounded determistic quasi-stationary, and the linear filters $F_1(q)$ and $F_2(q)$ are ℓ_∞-stable. Let the generating filter of the noise process $\{e_0(t)\}$ be given by $H_0(q) = \sum_{k=0}^{\infty} h_0(k)q^{-k}$, and let

$$e_0(t) = e_0^M(t) + \widetilde{e}_0^M(t),\ e_0^M(t) = \sum_{k=0}^{M} h_0(k)w_0(t-k),\ \widetilde{e}_0^M(t) = \sum_{k=M+1}^{\infty} h_0(k)w_0(t-k),$$

$$Z_N(M) = \frac{1}{\sqrt{\widetilde{N}}} \sum_{t=t_s}^{N} \begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} e_0^M(t),\ X_N(M) = \frac{1}{\sqrt{\widetilde{N}}} \sum_{t=t_s}^{N} \begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} \widetilde{e}_0^M(t),$$

such that $S_N = Z_N(M) + X_N(M)$. The signal $x_N(t) = \frac{1}{\sqrt{\widetilde{N}}} \begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} e_0^M(t)$ is M-dependent as $\{e_0^M(t)\}$ is M-dependent, and $\{e_0^M(t)\}$ is independent of $\{r_1(t)\}$ and

$\{r_2(t)\}$. Also, $Ex_N(t) = 0$ as $Ee_0^M(t) = 0$, and

$$\lim_{N\to\infty} \sum_{t=t_s}^{N} E|x_N(t)|^2 = E(e_0^M(t))^2 \lim_{N\to\infty} \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} \left(r_1^2(t) + r_2^2(t)\right) < \infty,$$

as $\{r_1(t)\}$, $\{r_2(t)\}$ and $\{e_0(t)\}$, and consequently $\{e_0^M(t)\}$, have finite variances, $\{r_1(t)\}$, $\{r_2(t)\}$ being quasi-stationary and $\{e_0^M(t)\}$ being stationary. Moreover,

$$\lim_{N\to\infty} \sum_{t=t_s}^{N} E|x_N(t)|^4 = E(e_0^M(t))^4 \lim_{N\to\infty} \frac{1}{\widetilde{N}^2} \sum_{t=t_s}^{N} \left(r_1^4(t) + 2r_1^2(t)r_2^2(t) + r_2^4(t)\right) = 0,$$

as $\{r_1(t)\}$, $\{r_2(t)\}$ are bounded signals and $\{e_0^M(t)\}$ has bounded fourth moments. Consequently, the signal $\{x_N(t)\}$ satisfies the conditions of Lemma 9A.1 in Ljung (1987), and hence,

$$Z_N(M) \xrightarrow{N\to\infty} \mathcal{N}(0, \Lambda_{r_1r_2}(M)),$$

with

$$\Lambda_{r_1r_2}(M) = \lim_{N\to\infty} EZ_N(M)Z_N^T(M).$$

The signals $\{r_1(t)\}$, $\{r_2(t)\}$ and $\{e_0(t)\}$ are generated in a sufficiently stable manner. With Lemma 2B.1 in Ljung (1987) it can then be established that $EX_N^2(M)$ is bounded by a bound which tends to zero as M tends to infinity. Consequently, Lemma 9A.2 in Ljung (1987) applies, and

$$S_N \xrightarrow{N\to\infty} \mathcal{N}(0, \Lambda_{r_1r_2}),$$

with

$$\Lambda_{r_1r_2} = \lim_{M\to\infty} \Lambda_{r_1r_2}(M) = \lim_{N\to\infty} \Lambda_{r_1r_2}^N, \quad \Lambda_{r_1r_2}^N = ES_NS_N^T.$$

The Continuous Mapping Theorem (Pollard, 1984, p. 46) states that if S_N converges in distribution to S, then $f(S_N)$ converges in distribution to $f(S)$, for any continuous function f. Therefore it follows that

$$S_N^T \left(\Lambda_{r_1r_2}^N\right)^{-1} S_N \xrightarrow{N\to\infty} S_N^T \Lambda_{r_1r_2}^{-1} S_N \xrightarrow{N\to\infty} \chi^2(2),$$

provided $\Lambda_{r_1r_2}^N$ is invertible.

Denote matrix-element (i,j) of $\Lambda_{r_1r_2}^N$ by $\lambda_{r_ir_j}^N$. Again, using the fact that the signals $\{r_1(t)\}$ and $\{r_2(t)\}$ are deterministic and independent of $\{e_0(t)\}$, it is possible to write for $\lambda_{r_ir_j}^N$,

$$\begin{aligned}\lambda_{r_jr_i}^N &= \lambda_{r_ir_j}^N := E\frac{1}{\widetilde{N}} \left(\sum_{t=t_s}^{N} r_i(t)e_0(t)\right)\left(\sum_{t=t_s}^{N} r_j(t)e_0(t)\right) = \\ &= \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} \sum_{t'=t_s}^{N} Er_i(t)e_0(t)r_j(t')e_0(t') =\end{aligned}$$

$$= \frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} \sum_{\tau=t_s-t}^{N-t} r_i(t) r_j(t+\tau) E e_0(t) e_0(t+\tau) =$$

$$= \frac{1}{\widetilde{N}} \sum_{\tau=-N+t_s}^{0} \sum_{t=-\tau+t_s}^{N} r_i(t) r_j(t+\tau) R_{e_0}(\tau) +$$

$$+ \frac{1}{\widetilde{N}} \sum_{\tau=1}^{N-t_s} \sum_{t=t_s}^{N-\tau} r_i(t) r_j(t+\tau) R_{e_0}(\tau) =$$

$$= \frac{1}{\widetilde{N}} \sum_{\tau=-N+t_s}^{0} R_{e_0}(\tau) \sum_{t=t_s}^{N+\tau} r_i(t-\tau) r_j(t) + \frac{1}{\widetilde{N}} \sum_{\tau=1}^{N-t_s} R_{e_0}(\tau) \sum_{t=t_s}^{N-\tau} r_i(t) r_j(t+\tau) =$$

$$= \frac{1}{\widetilde{N}} \sum_{\tau=-N+t_s}^{N-t_s} R_{e_0}(\tau)(\widetilde{N} - |\tau|) R^N_{r_j r_i}(\tau), \; i = 1,2, \; j = 1,2.$$

where the substitution $\tau = t' - t$ has been used. Finally,

$$\Lambda_{r_1 r_2} = \lim_{M \to \infty} \Lambda_{r_1 r_2}(M) = \lim_{M \to \infty} \lim_{N \to \infty} E Z_N(M) Z_N^T(M) =$$

$$= \lim_{M \to \infty} \lim_{N \to \infty} \begin{bmatrix} \lambda^N_{r_1 r_1}(M) & \lambda^N_{r_1 r_2}(M) \\ \lambda^N_{r_2 r_1}(M) & \lambda^N_{r_2 r_2}(M) \end{bmatrix},$$

where

$$\lambda^N_{r_i r_j}(M) := \sum_{\tau=-N+t_s}^{N-t_s} \frac{(\widetilde{N} - |\tau|)}{\widetilde{N}} R_{e_0^M}(\tau) R^N_{r_i r_j}(\tau), \; i = 1,2, \; j = 1,2,$$

and, for $i, j = 1, 2$,

$$\lim_{M \to \infty} \lim_{N \to \infty} \lambda^N_{r_i r_j}(M) = \lim_{M \to \infty} \sum_{\tau=-M}^{M} R_{e_0^M}(\tau) R_{r_i r_j}(\tau) = \sum_{\tau=-\infty}^{\infty} R_{e_0}(\tau) R_{r_i r_j}(\tau),$$

which completes the proof.

Proof of Theorem 5.4.3

Combining (5.20) with Lemma 5.4.2 yields

$$\sqrt{\widetilde{N}} \begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j})\right) - \dfrac{d(1)}{\widetilde{N}} - \dfrac{f(1)}{\widetilde{N}} \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j})\right) - \dfrac{d(2)}{\widetilde{N}} - \dfrac{f(2)}{\widetilde{N}} \end{bmatrix} =$$

$$= \frac{1}{\sqrt{\widetilde{N}}} \begin{bmatrix} \sum_{t=t_s}^{N} r_1(t) e_0(t) \\ \sum_{t=t_s}^{N} r_2(t) e_0(t) \end{bmatrix} \xrightarrow{N \to \infty} \mathcal{N}(0, \Lambda_{r_1 r_2}),$$

as the conditions of Lemma 5.4.2 are satisfied. In particular $\{r_1(t)\}$ and $\{r_2(t)\}$ are filtered versions of the signal $\{r(t)\}$, which is independent of $\{e_0(t)\}$. Also, note that $\Lambda^N_{r_1r_2}$ equals $\Lambda_{r_1r_2}$, if $N \to \infty$. Using $c_{\mathcal{N},\alpha}$ taken from the Standard Normal distribution, gives, for $N \to \infty$,

$$\sqrt{\widetilde{N}}\left|\mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) - \frac{d(1)}{\widetilde{N}} - \frac{f(1)}{\widetilde{N}} + \mathrm{Re}\left(\bar{G}_0(e^{i\omega_j})\right)\right| \leq c_{\mathcal{N},\alpha}\sqrt{\lambda^N_{r_1r_1}}, \text{ w.p. } \alpha.$$

Dividing the left- and right-hand side by $\sqrt{\widetilde{N}}$, using the bounds derived in (5.21), (5.22) and (5.23), and applying the triangle inequality, gives that, w.p. $\geq \alpha$,

$$\begin{aligned} c_{\mathcal{N},\alpha}\sqrt{\frac{\lambda^N_{r_1r_1}}{\widetilde{N}}} &\geq \left|\mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\right| - \frac{|d(1)|}{\widetilde{N}} - \frac{|f(1)|}{\widetilde{N}} - \left|\mathrm{Re}\left(\bar{G}_0(e^{i\omega_j})\right)\right| \geq \\ &\geq \left|\mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right)\right| - \frac{\bar{d}(1)}{\widetilde{N}} - \frac{\bar{f}(1)}{\widetilde{N}} - \delta(1), \end{aligned}$$

which proves part (i). Part (ii) can be proven completely similarly to this.

If $\Lambda^N_{r_1r_2}$ is invertible, part (iv) of Lemma 5.4.2 is applicable. Using $c_{\chi,\alpha}$ taken from the Chi-square distribution with 2 degrees of freedom, this gives for $N \to \infty$,

$$\widetilde{N}\begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j})\right) - \frac{d(1)}{\widetilde{N}} - \frac{f(1)}{\widetilde{N}} \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j})\right) - \frac{d(2)}{\widetilde{N}} - \frac{f(2)}{\widetilde{N}} \end{bmatrix}^T \Gamma^T \cdot$$

$$\cdot\Gamma\begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j})\right) - \frac{d(1)}{\widetilde{N}} - \frac{f(1)}{\widetilde{N}} \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - \widetilde{G}_0(e^{i\omega_j})\right) - \frac{d(2)}{\widetilde{N}} - \frac{f(2)}{\widetilde{N}} \end{bmatrix} \leq c_{\chi,\alpha}, \text{ w.p. } \alpha.$$

Hence, w.p. $\geq \alpha$,

$$\sqrt{\widetilde{N}}\left|\Gamma\begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) - \frac{d(1)}{\widetilde{N}} - \frac{f(1)}{\widetilde{N}} + \mathrm{Re}\left(\bar{G}_0(e^{i\omega_j})\right) \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) - \frac{d(2)}{\widetilde{N}} - \frac{f(2)}{\widetilde{N}} + \mathrm{Im}\left(\bar{G}_0(e^{i\omega_j})\right) \end{bmatrix}\right| \leq \sqrt{c_{\chi,\alpha}}.$$

Dividing the left- and right-hand side by $\sqrt{\widetilde{N}}$ and applying the triangle inequality, gives that, w.p. $\geq \alpha$,

$$\begin{aligned} \sqrt{\frac{c_{\chi,\alpha}}{\widetilde{N}}} \geq &\left|\Gamma\begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) \end{bmatrix}\right| - \frac{1}{\widetilde{N}}\left|\Gamma\begin{bmatrix} d(1) \\ d(2) \end{bmatrix}\right| + \\ &- \frac{1}{\widetilde{N}}\left|\Gamma\begin{bmatrix} f(1) \\ f(2) \end{bmatrix}\right| - \left|\Gamma\begin{bmatrix} \mathrm{Re}\left(\bar{G}_0(e^{i\omega_j})\right) \\ \mathrm{Im}\left(\bar{G}_0(e^{i\omega_j})\right) \end{bmatrix}\right| \geq \end{aligned}$$

$$\geq \left| \Gamma \begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) \end{bmatrix} \right| +$$

$$- \left| \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \end{bmatrix} \right| \left(\left| \mathrm{Re}\left(\bar{G}_0(e^{i\omega_j})\right) \right| + \frac{|d(1)|}{\widetilde{N}} + \frac{|f(1)|}{\widetilde{N}} \right) +$$

$$- \left| \begin{bmatrix} \gamma_{12} \\ \gamma_{22} \end{bmatrix} \right| \left(\left| \mathrm{Im}\left(\bar{G}_0(e^{i\omega_j})\right) \right| + \frac{|d(2)|}{\widetilde{N}} + \frac{|f(2)|}{\widetilde{N}} \right) \geq$$

$$\geq \left| \Gamma \begin{bmatrix} \mathrm{Re}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) \\ \mathrm{Im}\left(\widehat{G}(e^{i\omega_j}) - G_0(e^{i\omega_j})\right) \end{bmatrix} \right| - \sqrt{\gamma_{11}^2 + \gamma_{21}^2} \left(\delta(1) + \frac{\bar{d}(1)}{\widetilde{N}} + \frac{\bar{f}(1)}{\widetilde{N}} \right) +$$

$$- \sqrt{\gamma_{12}^2 + \gamma_{22}^2} \left(\delta(2) + \frac{\bar{d}(2)}{\widetilde{N}} + \frac{\bar{f}(2)}{\widetilde{N}} \right),$$

from which result (iii) follows by taking the second and third term to the left and squaring left- and right-hand side.

Proof of Theorem 5.6.2

The facts that $\{\psi(t)\}$, given in (5.30), is independent of $\{e_0(t)\}$ and that the latter process is stationary, give

$$R^N_{\hat{e}}(\tau) := E\widehat{R}^N_{\hat{e}}(\tau) = R^N_{\psi}(\tau) + R_{e_0}(\tau), \ |\tau| \leq \widetilde{N} - 1.$$

Denote matrix-element (i,j) of $\widehat{\Lambda}^N_{r_1 r_2}$ defined in (5.32) by $\widehat{\lambda}^N_{r_i r_j}$. Then,

$$E\widehat{\lambda}^N_{r_i r_j} = \sum_{\tau=-w}^{w} c_w(\tau) \frac{\widetilde{N} - |\tau|}{\widetilde{N}} R^N_{r_i r_j}(\tau) \left(R^N_{\psi}(\tau) + R_{e_0}(\tau) \right), \ i,j = 1,2.$$

Due to the fact that $\{r(t)\}$ is bounded deterministic quasi-stationary, $\{r_1(t)\}$ and $\{r_2(t)\}$ are jointly quasi-stationary and bounded deterministic, and therefore $R^N_{r_i r_j}(\tau)$ is bounded, $i,j = 1,2$. Moreover, the auto-covariance function of the output error

$$R_{\hat{e}}(\tau) = \lim_{N\to\infty} R^N_{\hat{e}}(\tau) = R_{\psi}(\tau) + R_{e_0}(\tau)$$

shows exponential decay rate in τ as both $R_{\psi}(\tau)$ and $R_{e_0}(\tau)$ show exponential decay rate. Also, the following asymptotic relation holds for $R^N_{r_i r_j}(\tau)$ defined in Lemma 5.4.2,

$$\lim_{N\to\infty} R^N_{r_i r_j}(\tau) = R_{r_i r_j}(\tau), \ \forall\, |\tau| \leq w,$$

where w is allowed to tend to infinity as long as $\frac{w}{N}$ tends to 0, which is assumed to be the case. The facts that $\frac{N-|\tau|}{N}$ tends to 1 for $N \to \infty$ and $|\tau| \leq w$, that $R_{\hat{e}}(\tau)$ shows

exponential decay rate and that the window $c_w(\tau)$ satisfies assumption 3.3.2, give that

$$\lim_{N\to\infty} E\widehat{\lambda}^N_{r_ir_j} = \sum_{\tau=-\infty}^{\infty} R_{r_ir_j}(\tau)R_{e_0}(\tau) + \sum_{\tau=-\infty}^{\infty} R_{r_ir_j}(\tau)R_{\psi}(\tau) =$$

$$= \lambda_{r_ir_j} + \sum_{\tau=-\infty}^{\infty} R_{r_ir_j}(\tau)R_{\psi}(\tau),$$

where $\lambda_{r_ir_j}$ denotes element (i,j) of matrix $\Lambda_{r_1r_2}$. Under the given assumptions on $\{e_0(t)\}$, $\{\psi(\tau)\}$ and the window $c_w(\tau)$, Theorem 3.1 in Hjalmarsson (1993a) can be applied, which yields that with probability 1,

$$\lim_{N\to\infty} \widehat{\Lambda}^N_{r_1r_2} = \lim_{N\to\infty} E\widehat{\Lambda}^N_{r_1r_2}.$$

Finally, due to the fact that the auto-covariance function of $\psi(t)$ is exponentially decaying, there exists a stationary stochastic process $\{s(t)\}$ with exponentially decaying auto-covariance function for which $R_s(\tau) = R_\psi(\tau)$, $\forall\tau$. Without loss of generality it may be assumed that $\{s(t)\}$ is independent of $\{r_i(t)\}$, $i = 1,2$. Define the signals $p_i(t) = r_i(t)s(t)$, $i = 1,2$, then

$$\sum_{\tau=-\infty}^{\infty} R_{r_ir_j}(\tau)R_{\psi}(\tau) = \sum_{\tau=-\infty}^{\infty} R_{r_ir_j}(\tau)R_s(\tau) =$$

$$= \sum_{\tau=-\infty}^{\infty} Es(t)s(t+\tau) \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} Er_i(t)r_j(t+\tau) =$$

$$= \sum_{\tau=-\infty}^{\infty} \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} Es(t)r_i(t)s(t+\tau)r_j(t+\tau) =$$

$$= \sum_{\tau=-\infty}^{\infty} R_{p_jp_i}(\tau) = \Phi_{p_jp_i}(0) = \Phi_{p_ip_j}(0),\ i,j = 1,2,$$

where $\Phi_{p_ip_j}(\omega)$ denotes the cross-spectrum between $\{p_i(t)\}$ and $\{p_j(t)\}$. Consequently,

$$\Lambda_\Psi = \begin{bmatrix} \Phi_{p_1p_1}(0) & \Phi_{p_1p_2}(0) \\ \Phi_{p_1p_2}(0) & \Phi_{p_2p_2}(0) \end{bmatrix} \geq 0,$$

which completes the proof.

5.B Monte Carlo Simulations

At various places the uncertainty bounding identification procedure developed in this chapter relies on asymptotic results, asymptotic in the number of data. However, in practical situations always a finite number of measurement data is available. In order to check the correctness of the various asymptotic results for finite N, Monte Carlo

simulations have been carried out. The results of these simulations are reported in this appendix. In particular the correctness for finite N of the asymptotic normality result, part (iii) of Lemma 5.4.2, Theorem 5.4.3, Theorem 5.5.1 and Proposition 5.6.5, is investigated. And it is investigated how much the accuracy of the identification result is influenced by the replacement of $\Lambda^N_{r_1r_2}$ by $\widehat{\Lambda}^N_{r_1r_2}$, as given by Theorem 5.6.2.

It is recognized that simulations for one specific situation do not necessarily have general validity for any situation. Nevertheless it is expected that a certain typical behaviour can be detected in the simulations, which then leads to conclusions valid for a wider class of the systems than just the specific simulation model considered here.

In this appendix attention is mainly restricted to the identification of parameter confidence regions using Proposition 5.6.5, but the conclusions straightforwardly carry over to the identification of frequency or pulse response confidence regions.

Consider the system

$$G_0(q) = g_0(0) + g_0(1)q^{-1} + g_0(2)q^{-2},\ g_0(0) = 2,\ g_0(1) = 1, g_0(2) = 0.6,$$

which generates data

$$y(t) = G_0(q)u(t) + e_0(t),\ t = 1, \ldots, N,$$

with the noise process given by

$$e_0(t) = H_0(q)n(t),\ H_0(q) = 1 + 0.8q^{-1} + 0.2q^{-2} + 0.1q^{-3},$$

where $\{n(t)\}$ is a white noise signal, *uniformly* distributed between -0.5 and 0.5. The input signal $\{u(t)\}$ is chosen to be

$$u(t) = \sin(\frac{\pi}{5}t) + \sin(\frac{\pi}{2}t),\ t = -\infty, \ldots, N.$$

Different values for N are used, $N = 40,\ 80,\ 200,\ 400,\ 800$.

The standard FIR basis is used to parametrize the model set. The following prior parameter bounds are used,

$$|g_0(k)| \leq \infty,\ k = 0, 1, 2,\ |g_0(k)| \leq 0,\ k = 3, \ldots, \infty.$$

The truncation value n is chosen $n = 2$ and the starting sample $t_s = 3$. The consequence of these choices is that $\bar{d}(p)$ and $\bar{f}(p)$ appearing in Proposition 5.6.5 are equal to 0. The motivation to do this is that in general $\bar{d}(p)$ and $\bar{f}(p)$ have a deterministic, conservative contribution to the resulting parameter confidence regions. Such a conservative contribution would only positively influence the reliability of the parameter confidence regions. However, here the main interest is to check the correctness of the probabilistic contribution $c_{\mathcal{N},\alpha}\sqrt{\lambda^N_{r_pr_p}/\widetilde{N}}$ or $c_{\mathcal{N},\alpha}\sqrt{\widehat{\lambda}^N_{r_pr_p}/\widetilde{N}}$.

First, the validity of assuming normality for finite N is checked. For this purpose exact knowledge of the auto-covariance function of the noise process is used, which is given by

$$R_{e_0}(\tau) = \lambda_0 \sum_{k=0}^{\infty} h_0(k)h_0(\tau + k),$$

with λ_0 the variance of $\{n(t)\}$, in this case $\lambda_0 = \frac{1}{12}$ and $\{h_0(k)\}$ is the pulse response sequence of $H_0(q)$. Parameter bounds are calculated with Proposition 5.6.5 for various values of $c_{\mathcal{N},\alpha}$ and N. This has been done 10000 times for different noise realizations. For each parameter separately the percentage of correct parameter confidence regions, which contain the true parameter value, is calculated and compared to the expected percentage α. The results are presented in Table 5.1. The numerically calculated probabilities appear not to differ significantly from the theoretical ones, even for small values of N. It can be concluded that the Monte Carlo simulations show that the normal distribution derived for the asymptotic case $N \to \infty$ can very well be used to approximate the distribution for finite N. A similar conclusion has been drawn in Ninness (1993, Section 4.2).

Next, it is investigated how accurate the identification results are if the true variance is replaced by the estimated one. For this purpose the estimation procedure mentioned in Remark 3.3.4 in Chapter 3 is used. This is possible as the input signal is periodic. The variance estimate is given by

$$\widehat{\lambda}^N_{r_p r_p} = \sum_{\tau=-w}^{w} c_w(\tau) \frac{N-|\tau|}{N} R^N_{r_p r_p}(\tau) \widehat{R}^N_{\hat{e}}(\tau),$$

with

$$\widehat{R}^N_{\hat{e}}(\tau) := \frac{1}{N/2-|\tau|} \sum_{t=t_s}^{N/2-|\tau|} \widehat{e}(t)\widehat{e}(t+|\tau|),$$

where

$$\widehat{e}(t) := \frac{y(t+N/2)-y(t)}{\sqrt{2}}, \quad t = 1, \ldots, N/2.$$

Notice that the signal $\{\widehat{e}(t)\}$ has second order statistics identical to those of the noise process $\{e_0(t)\}$, due to the fact that the deterministic sinusoidal components cancel, provided N is a multiple of 40. Note that this specific procedure of constructing $\{\widehat{e}(t)\}$ implies that effectively only $\frac{N}{2}$ data samples are used to estimate the noise statistics.

The window $c_w(\tau)$ is chosen to be the Tukey-window (Brillinger, 1981, p. 55),

$$c_w(\tau) = \begin{cases} 1, & 0 \le |\tau| \le fw(N) \\ \frac{1}{2} + \frac{1}{2}\cos\left(\frac{\pi}{1-f}\left(\frac{|\tau|}{w(N)} - f\right)\right), & fw(N) < |\tau| < w(N) \\ 0, & |\tau| \ge w(N) \end{cases}, \tag{5.34}$$

where f and $w(N)$ are some user-defined constants. Here the choice $f = 0.5$ is made and various choices for w are made.

Again, 10000 simulations have been performed, each time with a different noise realization. For each data set the variance estimates $\widehat{\lambda}^N_{r_p r_p}$, $p = 1, 2, 3$, have been calculated. If a negative estimate resulted, which can sometimes happen if w is large in comparison to N, it has been replaced by the value 0.

Let μ^N_p be the calculated mean value of the estimates $\widehat{\lambda}^N_{r_p r_p}$ and let σ^N_p be the calculated standard deviation of these estimates. In Table 5.2 the normalized mean

values

$$\widetilde{\mu}_p^N := \frac{\mu_p^N}{\lambda_{r_p r_p}^N},$$

and the normalized standard deviations

$$\widetilde{\sigma}_p^N := \frac{\sigma_p^N}{\lambda_{r_p r_p}^N}$$

are given for various values of N and window parameter w. Note that the variance estimate is unbiased, except for $w = 2$, which is due to undermodelling of the noise process. Even in the case $w = 2$ correct results are obtained here, as the estimated variances are overbiased.

It appears that the standard deviation of the variance estimates is linearly proportional to $\sqrt{w}$ for $w > 2$, and inversely linearly proportional to $\sqrt{N}$ for all w and N. So, the outcomes of the Monte Carlo simulations are consistent with Remark 5.6.3.

Especially for large values of w and small values of N the standard deviations seem large, and the question is how this influences the accuracy of the identification results. For that purpose the estimated variances have been used to construct parameter confidence intervals, and the designed probability α is compared to the numerically calculated probability, for various values of α and N. The outcomes, based on 10000 simulations, are shown in Table 5.3 for the choice $w = 6$ and in Table 5.4 for $w = 12$.

It appears that a (small) error is made when the true variance is replaced by the estimated one. This error is larger than the error caused by assuming a normal distribution for finite N, as becomes clear from a comparison with Table 5.1. However, still the actual probabilities using the estimated variances appear to be very close to the designed probabilities if $\frac{w}{N}$ is small. And even for small values of N the actual probability does not deviate much from α.

It can be concluded that the Monte Carlo simulations justify both the use of the asymptotic normal distribution for finite N, and the replacement of the true variance $\lambda_{r_p r_p}^N$ by $\widehat{\lambda}_{r_p r_p}^N$, especially if $\frac{w}{N}$ is relatively small, say smaller than 0.1.

	$N = 40$	$N = 80$	$N = 200$	$N = 400$	$N = 800$
$p = 0,\ c_{\mathcal{N},0.6} = 0.842$	0.602	0.590	0.602	0.604	0.591
$p = 0,\ c_{\mathcal{N},0.8} = 1.281$	0.794	0.797	0.797	0.808	0.793
$p = 0,\ c_{\mathcal{N},0.95} = 1.960$	0.952	0.949	0.949	0.949	0.950
$p = 0,\ c_{\mathcal{N},0.99} = 2.576$	0.992	0.992	0.990	0.989	0.991
$p = 1,\ c_{\mathcal{N},0.6} = 0.842$	0.602	0.605	0.602	0.602	0.600
$p = 1,\ c_{\mathcal{N},0.8} = 1.281$	0.799	0.805	0.797	0.798	0.800
$p = 1,\ c_{\mathcal{N},0.95} = 1.960$	0.955	0.952	0.949	0.951	0.953
$p = 1,\ c_{\mathcal{N},0.99} = 2.576$	0.991	0.990	0.990	0.989	0.990
$p = 2,\ c_{\mathcal{N},0.6} = 0.842$	0.593	0.600	0.596	0.603	0.596
$p = 2,\ c_{\mathcal{N},0.8} = 1.281$	0.794	0.797	0.800	0.799	0.803
$p = 2,\ c_{\mathcal{N},0.95} = 1.960$	0.952	0.950	0.952	0.947	0.955
$p = 2,\ c_{\mathcal{N},0.99} = 2.576$	0.990	0.990	0.991	0.991	0.991

Table 5.1: Numerically calculated confidence levels for parameters $g_0(p)$ using true $\lambda^N_{r_p r_p}$.

	$N = 40$	$N = 80$	$N = 200$	$N = 400$	$N = 800$
$p = 0,\ w = 2$	1.11, 0.48	1.11, 0.34	1.12, 0.22	1.12, 0.15	1.12, 0.11
$p = 0,\ w = 6$	0.99, 0.38	0.99, 0.27	1.00, 0.17	1.00, 0.12	1.00, 0.08
$p = 0,\ w = 12$	1.00, 0.69	1.00, 0.49	1.00, 0.31	1.00, 0.22	1.00, 0.16
$p = 0,\ w = 24$	x	1.00, 0.69	1.00, 0.44	1.00, 0.31	1.00, 0.22
$p = 1,\ w = 2$	1.45, 0.55	1.47, 0.39	1.49, 0.25	1.49, 0.18	1.49, 0.13
$p = 1,\ w = 6$	1.00, 0.69	1.00, 0.50	0.99, 0.32	1.00, 0.23	1.00, 0.16
$p = 1,\ w = 12$	1.02, 0.92	1.00, 0.65	0.99, 0.42	1.00, 0.30	1.00, 0.21
$p = 1,\ w = 24$	x	1.00, 0.88	1.00, 0.58	1.00, 0.41	1.00, 0.29
$p = 2,\ w = 2$	1.11, 0.48	1.11, 0.34	1.12, 0.22	1.12, 0.15	1.12, 0.11
$p = 2,\ w = 6$	0.99, 0.37	0.99, 0.26	1.00, 0.17	1.00, 0.12	1.00, 0.08
$p = 2,\ w = 12$	1.00, 0.68	1.00, 0.49	1.00, 0.31	1.00, 0.22	1.00, 0.16
$p = 2,\ w = 24$	x	1.00, 0.68	1.00, 0.44	1.00, 0.31	1.00, 0.22

Table 5.2: Numerically calculated normalized mean and standard variation $\widetilde{\mu}^N_p$, $\widetilde{\sigma}^N_p$. Window parameter $f = 0.5$.

	$N = 40$	$N = 80$	$N = 200$	$N = 400$	$N = 800$
$p = 0,\ c_{\mathcal{N},0.6} = 0.842$	0.582	0.583	0.597	0.603	0.593
$p = 0,\ c_{\mathcal{N},0.8} = 1.281$	0.775	0.785	0.791	0.803	0.789
$p = 0,\ c_{\mathcal{N},0.95} = 1.960$	0.926	0.936	0.943	0.947	0.948
$p = 0,\ c_{\mathcal{N},0.99} = 2.576$	0.973	0.984	0.988	0.988	0.990
$p = 1,\ c_{\mathcal{N},0.6} = 0.842$	0.547	0.569	0.587	0.598	0.596
$p = 1,\ c_{\mathcal{N},0.8} = 1.281$	0.712	0.755	0.781	0.789	0.794
$p = 1,\ c_{\mathcal{N},0.95} = 1.960$	0.846	0.901	0.931	0.942	0.949
$p = 1,\ c_{\mathcal{N},0.99} = 2.576$	0.896	0.951	0.979	0.985	0.989
$p = 2,\ c_{\mathcal{N},0.6} = 0.842$	0.576	0.589	0.595	0.600	0.595
$p = 2,\ c_{\mathcal{N},0.8} = 1.281$	0.770	0.785	0.794	0.796	0.801
$p = 2,\ c_{\mathcal{N},0.95} = 1.960$	0.927	0.939	0.944	0.946	0.953
$p = 2,\ c_{\mathcal{N},0.99} = 2.576$	0.975	0.983	0.989	0.988	0.991

Table 5.3: Numerically calculated confidence levels for parameters $g_0(p)$ using estimates $\widehat{\lambda}^N_{r_p r_p}$ with window parameters $w = 6,\ f = 0.5$.

	$N = 40$	$N = 80$	$N = 200$	$N = 400$	$N = 800$
$p = 0,\ c_{\mathcal{N},0.6} = 0.842$	0.549	0.565	0.591	0.602	0.590
$p = 0,\ c_{\mathcal{N},0.8} = 1.281$	0.724	0.759	0.781	0.794	0.787
$p = 0,\ c_{\mathcal{N},0.95} = 1.960$	0.860	0.907	0.932	0.941	0.946
$p = 0,\ c_{\mathcal{N},0.99} = 2.576$	0.915	0.962	0.981	0.986	0.989
$p = 1,\ c_{\mathcal{N},0.6} = 0.842$	0.521	0.554	0.577	0.590	0.595
$p = 1,\ c_{\mathcal{N},0.8} = 1.281$	0.672	0.724	0.772	0.785	0.793
$p = 1,\ c_{\mathcal{N},0.95} = 1.960$	0.797	0.868	0.922	0.936	0.944
$p = 1,\ c_{\mathcal{N},0.99} = 2.576$	0.849	0.920	0.971	0.982	0.987
$p = 2,\ c_{\mathcal{N},0.6} = 0.842$	0.544	0.572	0.589	0.596	0.596
$p = 2,\ c_{\mathcal{N},0.8} = 1.281$	0.716	0.755	0.784	0.791	0.799
$p = 2,\ c_{\mathcal{N},0.95} = 1.960$	0.866	0.912	0.935	0.941	0.949
$p = 2,\ c_{\mathcal{N},0.99} = 2.576$	0.924	0.963	0.983	0.985	0.988

Table 5.4: Numerically calculated confidence levels for parameters $g_0(p)$ using estimates $\widehat{\lambda}^N_{r_p r_p}$ with window parameters $w = 12,\ f = 0.5$.

Part II

Identification of a Nominal Model

Chapter 6

A Frequency Response Curve Fit Procedure

6.1 Introduction

In this chapter a frequency response curve fit algorithm is developed. In the next chapter this curve fit procedure is applied to the problem of system approximation in H_∞. However, the procedure is considered of independent interest as well. Therefore this chapter is written such that it can be read independently of the other chapters in this thesis.

The problem is considered of fitting a discrete or continuous time stable rational transfer function of a specified order to a set of complex frequency response data. In literature reasonable attention has been paid to curve fit problems, however, without restricting the model set to the set of stable models. Levy (1959) introduced a curve fit problem with a sum of squares (ℓ_2) criterion function and provided a solution for it. Sanathanan and Koerner (1963) proposed an iterative adjustment of the weighting in order to improve on the estimate. The asymptotic behaviour of this and other iterative linear schemes has been investigated by Whitfield (1987). Payne (1970) already noticed that if high order models are used in the curve fit procedure, the resulting estimate tends to be unstable.

Alternative criteria (other than ℓ_2) are used in Spanos (1991) and Sidman *et al.* (1991). In the first paper an iteration of weighted least squares problems is used to solve an unweighted least maximum amplitude (ℓ_∞) optimization problem. In the latter paper the ℓ_2-criterion is adjusted such that it is well suited for logarithmically spaced frequency response data.

An entirely different approach is taken in the area of identification in H_∞, see the discussion at the end of Subsection 2.4.5. There the objective is also to fit a transfer function to a set of frequency response data, but generally no model order restrictions are imposed. More is said about this approach in Remark 7.2.5 in Chapter 7.

In this chapter a weighted maximum amplitude (ℓ_∞) criterion is considered. An advantage of an ℓ_∞-criterion compared to an ℓ_2-criterion is that in the ℓ_∞-case each frequency is basically equally important. The frequency distribution (linearly or log-

arithmically, etc.) does not influence the resulting estimate as much as it does in the ℓ_2-case. Moreover, the weighting allows for an easy and effective shaping of the model error. A final motivation for a maximum amplitude criterion is that in many applications, for example robust control design, it is desirable that the maximum (weighted) distance between system and model is minimal, in order to have maximum robustness. A disadvantage of an ℓ_∞-criterion is that it is more sensitive to outliers (due to noise) than an ℓ_2-criterion.

The ℓ_∞-optimization problem is solved by transforming it into a smooth nonlinear constrained optimization problem. This approach is different from the one followed by Spanos (1991). The transfer function is guaranteed to be stable by parametrizing the denominator as a product of first and second order polynomials and adding linear constraints on the denominator parameters to the optimization problem. In fact in the discrete time case a bound can be specified for the amplitude of the transfer function poles. In the continuous time case a bound can be specified for the real part of the poles. The nonlinear programming problem can be tackled using standard software, but requires the specification of a good initial estimate. A procedure is presented to arrive at such an initial estimate. It involves a Sanathanan-Koerner iteration in ℓ_∞-setting, where in each step one linear programming problem has to be solved.

Extensions of the curve fit procedure to the MIMO case are presented. Frequency domain identification of MIMO systems has also been considered by e.g. Lin and Wu (1982) and Bayard (1994). In the first paper it has been proposed first to identify a set of SISO transfer functions from each input to each output separately, and then to compose the complete MIMO system by means of a minimal realization method. In the latter paper a scalar denominator polynomial is used for the MIMO system. In this chapter a different approach is taken by a parametrization with left and right diagonal denominator matrices.

The outline of the chapter is as follows. In Section 6.2 the curve fit problem is formulated and a solution is derived in the form of a nonlinear programming problem. In Section 6.3 a procedure is derived to provide an initial estimate for the nonlinear programming problem. Section 6.4 discusses some practical implementational aspects of the linear and nonlinear programming problems involved. Section 6.5 extends the entire procedure to the MIMO case. Section 6.6 contains a simulation example which shows the applicability of the curve fit procedure. Finally, in Section 6.7 the results are discussed.

6.2 The Curve Fit Procedure

Consider complex-valued frequency response data $G(\omega_j)$, $j = 1, \ldots, l$, and a positive real-valued frequency dependent weight $\widetilde{W}(\omega_j)$, $j = 1, \ldots, l$. The objective is to find a stable rational transfer function $\widehat{G}(\xi)$ in a certain model set $\mathcal{M}$ that optimally describes the frequency response data in a weighted ℓ_∞-sense,

$$\widehat{G}(\xi) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{j=1,\ldots,l} \left| \Big(G(\omega_j) - \widehat{G}'(\xi(\omega_j)) \Big) \widetilde{W}(\omega_j) \right|. \tag{6.1}$$

Following the notation in Bayard (1994) the complex variable ξ is used to present the continuous and discrete time case in a unified way. So, ξ can be thought of as the Laplace variable s or the z-transform variable z. And $\xi(\omega)$ represents the evaluation of the complex variable ξ as a function of frequency, thus $\xi(\omega) = i\omega$, $\xi(\omega) = e^{i\omega}$ respectively.

$\mathcal{M}$ is chosen to be the model set consisting of all rational transfer functions with the numerator and denominator of specified degrees and the roots of the denominator in some a priori specified (stability) region $\mathcal{R}$ in the complex plane.

$$\mathcal{M}: \ \widehat{G}'(\xi) = \frac{n(\xi)}{d(\xi)}, \ n(\xi) = \sum_{k=0}^{n} n_k \xi^k, \ d(\xi) = \xi^d + \sum_{k=0}^{d-1} d_k \xi^k, \ \xi_{\mathrm{d},k} \in \mathcal{R}, \ k = 1, \ldots, d, \tag{6.2}$$

where $\xi_{\mathrm{d},k}$, $k = 1, \ldots, d$, are the roots of the polynomial $d(\xi)$, i.e. all solutions to the equation $d(\xi) = 0$. In the discrete time case the region $\mathcal{R}$ is defined by $\mathcal{R} = \mathcal{R}_\rho$ with

$$\mathcal{R}_\rho = \{\xi_{\mathrm{d},k} \text{ s.t. } |\xi_{\mathrm{d},k}| \leq \rho, \ k = 1, \ldots, d\},$$

which corresponds to an upper bound ρ on the amplitude of the poles of the transfer function. In the continuous time case this region is defined by $\mathcal{R} = \mathcal{R}_r$ with

$$\mathcal{R}_r = \{\xi_{\mathrm{d},k} \text{ s.t. } \mathrm{Re}\,(\xi_{\mathrm{d},k}) \leq r, \ k = 1, \ldots, d\},$$

which corresponds to an upper bound r on the real part of the poles of the transfer function. Here ρ and r are user-defined constants. The choice $\rho = 1$ corresponds to stability in the discrete time case. The choice $r = 0$ corresponds to stability in the continuous time case.

In order to cope with the region $\mathcal{R}$ for the roots of the denominator polynomial in a computationally attractive way, an alternative parametrization is introduced for the denominator by representing the polynomial as a product of first and second order polynomials.

$$d(\xi) = \begin{cases} \displaystyle\prod_{k=1}^{d/2} \left(\xi^2 + a_k \xi + b_k\right), & d \text{ even} \\ (\xi + c) \displaystyle\prod_{k=1}^{(d-1)/2} \left(\xi^2 + a_k \xi + b_k\right), & d \text{ odd} \end{cases} \tag{6.3}$$

where the parameters a_k, b_k and c are real-valued. Then the following lemma gives necessary and sufficient conditions for the location of the roots of $d(\xi)$ by means of simple linear constraints on the parameters.

Lemma 6.2.1 *Let $\xi_{\mathrm{d},k}$, $k = 1, \ldots, d$, be the roots of $d(\xi)$, i.e. all solutions to $d(\xi) = 0$. Then for any $\rho \geq 0$, $r \in \mathbb{R}$,*

$$\textbf{(i)} \qquad |\xi_{\mathrm{d},k}| \leq \rho, \ \forall\, k \iff \begin{cases} b_k \leq \rho^2, & \forall\, k \\ a_k \rho \leq \rho^2 + b_k, & \forall\, k \\ -a_k \rho \leq \rho^2 + b_k, & \forall\, k \\ -\rho \leq c \leq \rho, & \text{if } d \text{ is odd} \end{cases}$$

$$\textbf{(ii)}\ \ \mathrm{Re}\,(\xi_{\mathrm{d},k}) \leq r,\ \forall k \Longleftrightarrow \begin{cases} b_k + a_k r + r^2 \geq 0, & \forall k \\ a_k \geq -2r, & \forall k \\ c \geq -r, & \text{if } d \text{ is odd} \end{cases}$$

Proof: See Appendix 6.A. □

Note that the first part of the lemma has a straightforward application in restricting the roots of the denominator to the region $\mathcal{R}_\rho$ by means of simple linear parameter constraints. The second part of the lemma can, obviously, be applied in restricting the roots of the denominator to the region $\mathcal{R}_r$ by means of simple linear parameter constraints. The conditions given in Lemma 6.2.1 are necessary and sufficient conditions, which implies that any polynomial $d(\xi)$ with roots in the specified regions in the complex plane can be parametrized in such a way that the parameter constraints of Lemma 6.2.1 are satisfied.

After this close look at the model parametrization the actual curve fit problem is considered. The objective (6.1) can be written as

$$\min_{\widehat{G}' \in \mathcal{M}} \max_{j=1,\ldots,l} \left| \left(G(\omega_j) - \widehat{G}'(\xi(\omega_j)) \right) \widetilde{W}(\omega_j) \right|,$$

which is equivalent to

$$\min_{h_\infty, \widehat{G}' \in \mathcal{M}} h_\infty \text{ s.t.}$$
$$\left| \left(G(\omega_j) - \widehat{G}'(\xi(\omega_j)) \right) \widetilde{W}(\omega_j) \right| \leq h_\infty,\ j = 1, \ldots, l,$$

which is equivalent to

$$\min_{h_\infty, n(\xi), d(\xi)} h_\infty^2 \text{ s.t.}$$
$$\left| \left(G(\omega_j) - \tfrac{n(\xi(\omega_j))}{d(\xi(\omega_j))} \right) \widetilde{W}(\omega_j) \right|^2 \leq h_\infty^2,\ j = 1, \ldots, l,\ \xi_{\mathrm{d},k} \in \mathcal{R},$$

which is in turn equivalent to

$$\min_{h_\infty, n(\xi), d(\xi)} h_\infty^2 \text{ s.t.}$$
$$\widetilde{W}^2(\omega_j) \left| G(\omega_j) d(\xi(\omega_j)) - n(\xi(\omega_j)) \right|^2 \leq h_\infty^2 |d(\xi(\omega_j))|^2,\ j = 1, \ldots, l,\ \xi_{\mathrm{d},k} \in \mathcal{R}, \quad (6.4)$$

which is a nonlinear (nonconvex) constrained optimization problem. Consequently, the solution of the curve fit problem (6.1) is given by the solution of the nonlinear constrained optimization problem (6.4). In this nonlinear programming problem the numerator $n(\xi)$ is parametrized as given in (6.2) and the denominator $d(\xi)$ is parametrized as given in (6.3). The roots of the denominator polynomial are restricted to the region $\mathcal{R}$ by means of additional linear constraints, see Lemma 6.2.1. The nonlinear optimization problem (6.4) can be tackled by applying standard software for nonlinear constrained optimization. In Section 6.4 some aspects of the practical implementation are discussed. At this moment it is only noted that (6.4) is a smooth nonlinear problem, as the partial derivatives of the constraints and the objective function to the parameters are continuous functions.

6.3 The Initial Estimate

The nonlinear constrained optimization problem (6.4), which provides the solution for the curve fit problem (6.1), can generally only be solved satisfactorily if a good initial estimate is available. Such an initial estimate should satisfy two properties. First, it should be a feasible solution for the nonlinear optimization problem, which in fact means that its denominator roots should be in the region $\mathcal{R}$. Second, it should be close to the optimal solution in the sense that a low, although not necessarily minimal, criterion value is achieved. To arrive at such an initial esimate a three-step procedure is proposed here. In the first step an iteration is performed in order to obtain a curve fit model which is expected to yield a low criterion value. However, the poles of the model are not restricted to the region $\mathcal{R}$. In the second step then the poles are mirrored into the region $\mathcal{R}$, which ensures feasibility of the model. But this may happen at the price of an increase in criterion value. Therefore in the third and final step a new optimization is performed with respect to the numerator, the denominator being fixed. This will then yield an initial model which satisfies the constraints in the nonlinear optimization problem (6.4) and also has a low criterion value h_∞^2.

The first step consists of an iterative procedure using linear programming techniques. The following model structure is considered,

$$\mathcal{M} : \widehat{G}'(\xi) = \frac{n(\xi)}{d(\xi)}, \; n(\xi) = \sum_{k=0}^{n} n_k \xi^k, \; d(\xi) = \xi^d + \sum_{k=0}^{d-1} d_k \xi^k.$$

So stability of the model is not required, or more specifically the location of the denominator roots is not restricted to the region $\mathcal{R}$. For some positive, real valued weight $\widetilde{W}^*(\omega_j)$ the following optimization problem is considered,

$$\min_{n(\xi),d(\xi)} \max_{j=1,\ldots,l} \left| \left(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))\right) \widetilde{W}^*(\omega_j) \right|, \tag{6.5}$$

or equivalently,

$$\min_{h_\infty,n(\xi),d(\xi)} h_\infty \text{ s.t.}$$
$$\left| \left(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))\right) \widetilde{W}^*(\omega_j) \right| \leq h_\infty, \; j = 1,\ldots,l. \tag{6.6}$$

Using the technique introduced in Lemma 3.4.1 in Chapter 3, the optimization problem (6.6) can be approximated by the optimization problem

$$\min_{h_\infty,n(\xi),d(\xi)} h_\infty \text{ s.t.}$$
$$\text{Re}\left(e^{2\pi \frac{k'}{m'} i} \left(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))\right) \widetilde{W}^*(\omega_j) \right) \leq h_\infty, \; j = 1,\ldots,l, \; k' = 1,\ldots,m',$$

which is equivalent to the linear programming problem

$$\min_{h_\infty,n_k,d_k} h_\infty \text{ s.t.}$$

$$\widetilde{W}^*(\omega_j)\,\mathrm{Re}\left(e^{2\pi\frac{k'}{m'}i}G(\omega_j)\xi^d(\omega_j)\right)+\sum_{k=0}^{d-1}d_k\widetilde{W}^*(\omega_j)\,\mathrm{Re}\left(e^{2\pi\frac{k'}{m'}i}G(\omega_j)\xi^k(\omega_j)\right)+ \quad (6.7)$$

$$-\sum_{k=0}^{n}n_k\widetilde{W}^*(\omega_j)\,\mathrm{Re}\left(e^{2\pi\frac{k'}{m'}i}\xi^k(\omega_j)\right)-h_\infty\le 0,\ j=1,\ldots,l,\ k'=1,\ldots,m'.$$

The following theorem quantifies the accuracy of the approximation.

Theorem 6.3.1 *Denote the optimal solution of (6.5) by* h'_∞, $n'(\xi)$, $d'(\xi)$ *and the optimal solution of (6.7) by* h''_∞, $n''(\xi)$, $d''(\xi)$. *Then*

(i)
$$h''_\infty \le \max_{j=1,\ldots,l}\left|(G(\omega_j)d'(\xi(\omega_j))-n'(\xi(\omega_j)))\,\widetilde{W}^*(\omega_j)\right| \le$$
$$\le \max_{j=1,\ldots,l}\left|(G(\omega_j)d''(\xi(\omega_j))-n''(\xi(\omega_j)))\,\widetilde{W}^*(\omega_j)\right| \le \frac{h''_\infty}{\cos\left(\frac{\pi}{m'}\right)},$$

(ii)
$$\lim_{m'\to\infty}\frac{n''(\xi)}{d''(\xi)}=\frac{n'(\xi)}{d'(\xi)}.$$

Proof: See Appendix 6.A. □

The optimization problem (6.7) can be solved exactly using standard linear programming software available. See Luenberger (1984) or Karloff (1991) for an extensive treatment of the linear programming problem and algorithms to solve it. According to Theorem 6.3.1 the solution to the optimization problem (6.7) is an arbitrarily good approximation to problem (6.5) for m' sufficiently large.

Now to obtain a curve fit model with a low criterion value h^2_∞ for the nonlinear problem (6.4) an iteration is proposed analogously to the Sanathanan-Koerner iteration for ℓ_2-curve fitting (Sanathanan and Koerner, 1963). In each iteration the weight $\widetilde{W}^*(\omega_j)$ is chosen such that

$$\widetilde{W}^*(\omega_j)=\widetilde{W}(\omega_j)\,|d_{\mathrm{prev}}(\xi(\omega_j))|^{-1},\ j=1,\ldots,l,$$

where $d_{\mathrm{prev}}(\xi)$ denotes the denominator that resulted in the previous iteration. The idea is that the optimization problem (6.5) more and more resembles the nonlinear problem (6.1). It is emphasized that, on the one hand, convergence of the iterative procedure is not guaranteed and, on the other hand, if it converges the ultimate outcome will generally not be equal to the optimal result of the non-linear problem (6.1). However, Sanathanan and Koerner reported good results obtained with this iteration in an ℓ_2-setting and extensive simulations show that the procedure also works adequately in the ℓ_∞-setting adopted here, i.e. generally there is rapid convergence and a low value of the nonlinear objective function is often achieved.

As mentioned before the model resulting from the iteration described above need not have its poles in the region $\mathcal{R}$, as no constraints of the kind of those in Lemma 6.2.1 are imposed on the denominator. Therefore now a second step is performed which consists of mirroring the poles into the region $\mathcal{R}$. In the discrete time case the proposed

procedure is to replace a denominator root $\xi_{d,k} = a$, with $|a| > \rho$, by a root $\xi_{d,k} = 1/a$ if $1/a \leq \rho$, or by a root $\xi_{d,k} = \rho a/|a|$ if $1/a > \rho$. Analogously in the continuous time case the proposed procedure is to replace a denominator root $\xi_{d,k} = a$, with $\mathrm{Re}\,(a) > r$, by a root $\xi_{d,k} = -\mathrm{Re}\,(a) + \mathrm{Im}\,(a)$ if $-\mathrm{Re}\,(a) \leq r$, or by a root $\xi_{d,k} = r + \mathrm{Im}\,(a)$ if $-\mathrm{Re}\,(a) > r$. An important feature of this procedure is, that no new poles will be created which are closer to the unit circle (or imaginary axis) than the poles before mirroring. If such poles would be created, easily ill-conditioning could occur when the model is evaluated over the unit circle (imaginary axis) in the nonlinear curve fit procedure.

In the second step described above, the poles are put into the region $\mathcal{R}$, but the achieved low criterion value h_∞^2 may be lost. Therefore a third step is performed which consists of a final linear optimization. This optimization is basically identical to the optimization (6.5) of the first step, but now with the denominator $d(\xi)$ fixed to the result of the second step. Only the numerator $n(\xi)$ is adjusted by minimizing the criterion (6.5) with

$$d(\xi(\omega_j)) = d_{\text{fixed}}(\xi(\omega_j)),\ \widetilde{W}^*(\omega_j) = \widetilde{W}(\omega_j)\,|d_{\text{fixed}}(\xi(\omega_j))|^{-1},\ j = 1, \ldots, l.$$

Note that in case a fixed denominator $d(\xi)$ is used, or equivalently the denominator degree d is chosen to be zero, then the optimization problem (6.5) is identical to the optimization problem (6.1). This in turn means that in that case the solution of the linear programming problem (6.7) is arbitrarily close to the solution of the nonlinear programming problem (6.4). The consequence of this third and final step will generally be that a model is obtained which yields a low criterion value h_∞^2 in (6.4) whereas it is a feasible solution to the constraints in the nonlinear programming problem. Therefore it can be expected that the resulting model is a suitable initial estimate for the nonlinear optimization problem (6.4). Extensive simulations confirm this expectation.

6.4 Implementation Aspects

The curve fit problem (6.1) appears solvable by using linear and nonlinear programming routines. In practice things like numerical accuracy play an important role. Therefore some modifications to the programming problems (6.4) and (6.7) may be necessary to get well-conditioned optimization problems.

From equation (6.4) it is apparent that in the continuous time case, $\xi(\omega_j) = i\omega_j$, the expressions on the left- and right-hand side can be very large if $\omega_j > 1$. This may cause numerical problems. The problem is dealt with by introducing a scaling factor. A proper scaling is achieved by multiplying all constraints corresponding to $\omega_j > 1$ in the continuous time case with the factor $f^2(\omega_j)$ where $f(\omega_j) = \omega_j^{-\max(n,d)}$. Thus, one such constraint is given by

$$f^2(\omega_j)\widetilde{W}^2(\omega_j)\,|G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))|^2 \leq h_\infty^2 f^2(\omega_j)|d(\xi(\omega_j))|^2.$$

Analogously the constraints in the linear programming problem (6.7) need to be scaled. In the continuous time case for $\omega_j > 1$ all terms appearing in these constraints should

be multiplied with the factor $f(\omega_j)$ as given above. In the discrete time case and in the continuous time case for frequencies smaller than or equal to 1 no scaling is required.

A slight improvement of the nonlinear programming problem (6.4) is obtained by replacing h_∞^2 by a new parameter $h_{\infty,2}$. A consequence of this is that the objective function becomes linear, which is in general preferable to a nonlinear objective function.

For the simulations in Section 6.6 the sequential quadratic programming method implemented in the FORTRAN NAG-library has been used, see Gill *et al.* (1981) for details about this method. This method requires the first partial derivatives of the nonlinear constraints. To compute these the fact can be used that for any complex-valued function $y(x_1,\ldots,x_n)$ of the real variables $x_1,\ldots,x_n$ the first (partial) derivative of the squared magnitude is given by

$$\begin{aligned}\frac{\partial|y(x_1,\ldots,x_n)|^2}{\partial x_k} &= \frac{\partial y'y}{\partial x_k} = y'\frac{\partial y}{\partial x_k} + \frac{\partial y'}{\partial x_k}y = y'\frac{\partial y}{\partial x_k} + \left(y'\frac{\partial y}{\partial x_k}\right)' = \\ &= 2\,\mathrm{Re}\left(y'(x_1,\ldots,x_n)\frac{\partial y(x_1,\ldots,x_n)}{\partial x_k}\right),\end{aligned}$$

where y' denotes the complex conjugate of y. Then the first partial derivatives of the constraints in (6.4) are straightforwardly computed as

$$\frac{\partial\widetilde{W}^2(\omega_j)\,|G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))|^2 - h_{\infty,2}|d(\xi(\omega_j))|^2}{\partial h_{\infty,2}} = -|d(\xi(\omega_j))|^2,$$

$$\begin{aligned}&\frac{\partial\widetilde{W}^2(\omega_j)\,|G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))|^2 - h_{\infty,2}|d(\xi(\omega_j))|^2}{\partial n_k} = \\ &= -2\widetilde{W}^2(\omega_j)\,\mathrm{Re}\left((G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'\frac{\partial n(\xi(\omega_j))}{\partial n_k}\right) = \\ &= -2\widetilde{W}^2(\omega_j)\,\mathrm{Re}\left((G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'\xi^k(\omega_j)\right),\end{aligned}$$

$$\begin{aligned}&\frac{\partial\widetilde{W}^2(\omega_j)\,|G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))|^2 - h_{\infty,2}|d(\xi(\omega_j))|^2}{\partial a_k} = \\ &= 2\,\mathrm{Re}\left(\left(\widetilde{W}^2(\omega_j)(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'G(\omega_j) - h_{\infty,2}d'(\xi(\omega_j))\right)\frac{\partial d(\xi(\omega_j))}{\partial a_k}\right) \\ &= 2\,\mathrm{Re}\left(\left(\widetilde{W}^2(\omega_j)(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'G(\omega_j) - h_{\infty,2}d'(\xi(\omega_j))\right)\cdot\right. \\ &\quad\left.\cdot\frac{d(\xi(\omega_j))\xi(\omega_j)}{\xi^2(\omega_j) + a_k\xi(\omega_j) + b_k}\right),\end{aligned}$$

$$\frac{\partial\widetilde{W}^2(\omega_j)\,|G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))|^2 - h_{\infty,2}|d(\xi(\omega_j))|^2}{\partial b_k} =$$

$$= 2\,\mathrm{Re}\left(\left(\widetilde{W}^2(\omega_j)(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'G(\omega_j) - h_{\infty,2}d'(\xi(\omega_j)) \right) \frac{\partial d(\xi(\omega_j))}{\partial b_k} \right)$$

$$= 2\,\mathrm{Re}\left(\left(\widetilde{W}^2(\omega_j)(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'G(\omega_j) - h_{\infty,2}d'(\xi(\omega_j)) \right) \cdot \right.$$

$$\left. \cdot \frac{d(\xi(\omega_j))}{\xi^2(\omega_j) + a_k\xi(\omega_j) + b_k} \right),$$

$$\frac{\partial \widetilde{W}^2(\omega_j)\,|G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j))|^2 - h_{\infty,2}|d(\xi(\omega_j))|^2}{\partial c} =$$

$$= 2\,\mathrm{Re}\left(\left(\widetilde{W}^2(\omega_j)(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'G(\omega_j) - h_{\infty,2}d'(\xi(\omega_j)) \right) \frac{\partial d(\xi(\omega_j))}{\partial c} \right)$$

$$= 2\,\mathrm{Re}\left(\left(\widetilde{W}^2(\omega_j)(G(\omega_j)d(\xi(\omega_j)) - n(\xi(\omega_j)))'G(\omega_j) - h_{\infty,2}d'(\xi(\omega_j)) \right) \frac{d(\xi(\omega_j))}{\xi(\omega_j) + c} \right).$$

If a scaling $f(\omega_j)$ is used the expressions should of course be changed accordingly.

Finally, it is noticed that the parametrization (6.3) is not unique as there is freedom in the ordering of the second order terms. It depends on the specific nonlinear programming routine that is used if this nonuniqueness in the solution space will cause trouble. The routine applied for the simulations in Section 6.6 has no difficulties with it. If a routine is used that has difficulties with the nonuniqueness, uniqueness may be enforced by specifying additional constraints which correspond to ordering the second order terms uniquely. These constraints may, for example, be the following linear constraints,

$$a_1 > a_2 > \cdots > a_p, \quad p = d/2 \text{ or } p = (d-1)/2,$$

yielding a unique parametrization, which, however, excludes the possibility of identical a-parameters in two second order polynomials.

6.5 The Multivariable Case

In this section extensions of the curve fit procedure to the multivariable case are presented.

6.5.1 The Curve Fit Procedure

Consider complex-valued frequency response data $G(\omega_j)$, $j = 1, \ldots, l$, where $G(\omega_j)$ is an $p' \times q'$ matrix, and the real-valued $p' \times q'$ weighting matrix $\widetilde{W}(\omega_j)$, $j = 1, \ldots, l$. The objective is to find a stable rational transfer function $\widehat{G}(\xi)$ in a certain model set $\mathcal{M}$ that optimally describes the frequency response data in a weighted ℓ_∞-sense,

$$\widehat{G}(\xi) = \arg \min_{\widehat{G} \in \mathcal{M}} \max_{i',j',j} \left| \left(G_{i'j'}(\omega_j) - \widehat{G}^l_{i'j'}(\xi(\omega_j)) \right) \widetilde{W}_{i'j'}(\omega_j) \right|. \tag{6.8}$$

For the model set $\mathcal{M}$ a parametrization is chosen which is highly flexible but still easy to handle,

$$\mathcal{M}: \ \widehat{G}'(\xi) = D_{\mathrm{L}}^{-1}(\xi)N(\xi)D_{\mathrm{R}}^{-1}(\xi),$$

$$N(\xi) = \begin{bmatrix} N_{11}(\xi) & \cdots & N_{1q'}(\xi) \\ \vdots & \ddots & \vdots \\ N_{p'1}(\xi) & \cdots & N_{p'q'}(\xi) \end{bmatrix}, \ N_{i'j'}(\xi) = \sum_{k=0}^{n_{i'j'}} n_{i'j'}(k)\xi^k,$$

$$D_{\mathrm{L}}(\xi) = \mathrm{diag}\,\{D_{\mathrm{L}1}(\xi), \ldots, D_{\mathrm{L}p'}(\xi)\}, \ D_{\mathrm{L}i'}(\xi) = \xi^{d_{\mathrm{L}i'}} + \sum_{k=0}^{d_{\mathrm{L}i'}-1} d_{\mathrm{L}i'}(k)\xi^k, \ \xi_{D_{\mathrm{L}i'},k} \in \mathcal{R},$$

$$D_{\mathrm{R}}(\xi) = \mathrm{diag}\,\{D_{\mathrm{R}1}(\xi), \ldots, D_{\mathrm{R}q'}(\xi)\}, \ D_{\mathrm{R}j'}(\xi) = \xi^{d_{\mathrm{R}j'}} + \sum_{k=0}^{d_{\mathrm{R}j'}-1} d_{\mathrm{R}j'}(k)\xi^k, \ \xi_{D_{\mathrm{R}j'},k} \in \mathcal{R},$$

where $\xi_{D_{\mathrm{L}i'},k}$ and $\xi_{D_{\mathrm{R}j'},k}$ denote the roots of the polynomials $D_{\mathrm{L}i'}(\xi)$ and $D_{\mathrm{R}j'}(\xi)$ respectively. Consequently, the transfer function from input j' to output i' is given by

$$\widehat{G}'_{i'j'}(\xi) = \frac{N_{i'j'}(\xi)}{D_{\mathrm{L}i'}(\xi)D_{\mathrm{R}j'}(\xi)}.$$

The generic (McMillan) order of the transfer function $\widehat{G}'(\xi)$ is $\sum_{i'=1}^{p'} d_{\mathrm{L}i'} + \sum_{j'=1}^{q'} d_{\mathrm{R}j'}$. The flexibility of this parametrization lies in the freedom of choosing values for the structure indices $n_{i'j'}$, $d_{\mathrm{L}i'}$, $d_{\mathrm{R}j'}$. This is in contrast with e.g. Bayard (1994) where a parametrization has been chosen with a scalar denominator polynomial, i.e. $D_{\mathrm{L}}(\xi) = \mathrm{diag}\,(d(\xi), \ldots, d(\xi))$, $D_{\mathrm{R}}(\xi) = I$. The consequence of that choice is that the generic McMillan order of the MIMO system is $\min(p', q')\deg(d(\xi))$. With such a parametrization in general nominal models are estimated which have a (much) higher order than the underlying system. This becomes clear from a simple example.

Example 6.5.1 *Consider a multivariable system with 2 inputs and 2 outputs, where each $G'_{i'j'}(z)$ is a first order transfer function, different from the other transfer functions (no common poles). Thus, the order of the MIMO system is 4. If a parametrization with a scalar denominator polynomial is used to identify the system, a denominator polynomial of degree 4 is required (in case of non-approximate identification). However, because of round-off errors and noise the resulting identified nominal model will in general have order $2 \cdot 4 = 8$! If the parametrization proposed above with diagonal left and right denominator matrices is used, it is, for example, possible to choose $d_{\mathrm{L}1} = d_{\mathrm{L}2} = d_{\mathrm{R}1} = d_{\mathrm{R}2} = 1$ or $d_{\mathrm{L}1} = d_{\mathrm{L}2} = 2$, $d_{\mathrm{R}1} = d_{\mathrm{R}2} = 0$ or $d_{\mathrm{L}1} = d_{\mathrm{L}2} = 0$, $d_{\mathrm{R}1} = d_{\mathrm{R}2} = 2$, each of these cases leading to 4th order nominal models being estimated.*

Of course there are examples in which the parametrization proposed here also requires a higher order nominal model to be estimated, but due to the higher flexibility in the parametrization this redundancy is less than in the case of a scalar denominator polynomial.

To determine appropriate values for $d_{\mathrm{L}i'}$ and $d_{\mathrm{R}j'}$ in practice it is suggested first to fit SISO transfer functions to each set of frequency response data $G_{i'j'}(\omega_j)$, $j = 1, \ldots, l$, using the techniques described in the previous sections. Now poles that are very close to each other (according to the user's judgement) can be assumed to be identical. Values for $d_{\mathrm{L}i'}$ and $d_{\mathrm{R}j'}$ can next be determined by observing if the poles repeat in the same row (output) or column (input) respectively. Next, the MIMO curve fit problem can be solved using the techniques described below.

By choosing the alternative parametrization for the denominator,

$$D_{\mathrm{L}i'}(\xi) = \begin{cases} \prod_{k=1}^{d_{\mathrm{L}i'}/2} \left(\xi^2 + a_{\mathrm{L}i'}(k)\xi + b_{\mathrm{L}i'}(k)\right), & d_{\mathrm{L}i'} \text{ even} \\ (\xi + c_{\mathrm{L}i'}) \prod_{k=1}^{(d_{\mathrm{L}i'}-1)/2} \left(\xi^2 + a_{\mathrm{L}i'}(k)\xi + b_{\mathrm{L}i'}(k)\right), & d_{\mathrm{L}i'} \text{ odd} \end{cases}$$

$$D_{\mathrm{R}j'}(\xi) = \begin{cases} \prod_{k=1}^{d_{\mathrm{R}j'}/2} \left(\xi^2 + a_{\mathrm{R}j'}(k)\xi + b_{\mathrm{R}j'}(k)\right), & d_{\mathrm{R}j'} \text{ even} \\ (\xi + c_{\mathrm{R}j'}) \prod_{k=1}^{(d_{\mathrm{R}j'}-1)/2} \left(\xi^2 + a_{\mathrm{R}j'}(k)\xi + b_{\mathrm{R}j'}(k)\right), & d_{\mathrm{R}j'} \text{ odd} \end{cases}$$

the roots of the denominator polynomials can be restricted to the (stability) region $\mathcal{R}$ by means of linear constraints on the parameters $a_{\mathrm{L}i'}(k)$, $a_{\mathrm{R}j'}(k)$, $b_{\mathrm{L}i'}(k)$, $b_{\mathrm{R}j'}(k)$, $c_{\mathrm{L}i'}$ and $c_{\mathrm{R}j'}$ as given in Lemma 6.2.1, analogously to the SISO case.

Analogously to the derivation of (6.4) the objective (6.8) can be written as

$$\min_{\widehat{G}' \in \mathcal{M}} \max_{i',j',j} \left| \left(G_{i'j'}(\omega_j) - \widehat{G}'_{i'j'}(\xi(\omega_j))\right) \widetilde{W}_{i'j'}(\omega_j) \right|,$$

which is equivelent to

$$\begin{gathered} \min_{h_\infty, N, D_{\mathrm{L}}, D_{\mathrm{R}}} h_\infty^2 \text{ s.t.} \\ \widetilde{W}_{i'j'}^2(\omega_j) \left|G_{i'j'}(\omega_j) D_{\mathrm{L}i'}(\xi(\omega_j)) D_{\mathrm{R}j'}(\xi(\omega_j)) - N_{i'j'}(\xi(\omega_j))\right|^2 \leq \\ \leq h_\infty^2 \left|D_{\mathrm{L}i'}(\xi(\omega_j)) D_{\mathrm{R}j'}(\xi(\omega_j))\right|^2, \; i' = 1, \ldots, p', \; j' = 1, \ldots, q', \; j = 1, \ldots, l, \quad (6.9) \\ \xi_{D_{\mathrm{L}i'},k} \in \mathcal{R}, \; i' = 1, \ldots, p', \; \xi_{D_{\mathrm{R}j'},k} \in \mathcal{R}, \; j' = 1, \ldots, q', \end{gathered}$$

which again is a smooth nonlinear (nonconvex) constrained optimization problem.

6.5.2 The Initial Estimate

To determine the initial estimate required for the nonlinear programming problem, the three-step procedure of Section 6.3 is proposed again. In the first step a different parametrization is used for the denominator,

$$D_{\mathrm{L}i'}(\xi) = \xi^{d_{\mathrm{L}i'}} + \sum_{k=0}^{d_{\mathrm{L}i'}-1} d_{\mathrm{L}i'}(k)\xi^k, \; D_{\mathrm{R}j'}(\xi) = \xi^{d_{\mathrm{R}j'}} + \sum_{k=0}^{d_{\mathrm{R}j'}-1} d_{\mathrm{R}j'}(k)\xi^k,$$

so without restricting the roots of the denominator to the region $\mathcal{R}$. Consider two subproblems, the first one given by

$$\min_{N(\xi),D_{\mathrm{L}}(\xi)} \max_{i',j',j} \left| (G_{i'j'}(\omega_j) D_{\mathrm{L}i'}(\xi(\omega_j)) F_{j'}(\omega_j) - N_{i'j'}(\xi(\omega_j))) \widetilde{W}^*_{i'j'}(\omega_j) \right|, \tag{6.10}$$

for some weight $\widetilde{W}^*(\omega_j)$ and constant $F_{j'}(\omega_j)$, $j' = 1, \ldots, q'$, to which will be come back later. The second subproblem is

$$\min_{N(\xi),D_{\mathrm{R}}(\xi)} \max_{i',j',j} \left| (G_{i'j'}(\omega_j) D_{\mathrm{R}j'}(\xi(\omega_j)) F_{i'}(\omega_j) - N_{i'j'}(\xi(\omega_j))) \widetilde{W}^*_{i'j'}(\omega_j) \right|, \tag{6.11}$$

for some weight $\widetilde{W}^*(\omega_j)$ and constant $F_{i'}(\omega_j)$, $i' = 1, \ldots, p'$, to which will be come back later as well. Consider the first subproblem. It is equivalent to the constrained optimization problem

$$\min_{h_\infty, N, D_{\mathrm{L}}} h_\infty \text{ s.t.}$$
$$\left| (G_{i'j'}(\omega_j) D_{\mathrm{L}i'}(\xi(\omega_j)) F_{j'}(\omega_j) - N_{i'j'}(\xi(\omega_j))) \widetilde{W}^*_{i'j'}(\omega_j) \right| \leq h_\infty,$$
$$i' = 1, \ldots, p',\ j' = 1, \ldots, q',\ j = 1, \ldots, l,$$

which can be split up into a set of p' independent optimization problems

$$\min_{h_\infty, N_{i'j'}, D_{\mathrm{L}i'}} h_\infty \text{ s.t.}$$
$$\left| (G_{i'j'}(\omega_j) D_{\mathrm{L}i'}(\xi(\omega_j)) F_{j'}(\omega_j) - N_{i'j'}(\xi(\omega_j))) \widetilde{W}^*_{i'j'}(\omega_j) \right| \leq h_\infty, \tag{6.12}$$
$$j' = 1, \ldots, q',\ j = 1, \ldots, l,$$

for $i' = 1, \ldots, p'$. Similar to the SISO case this optimization problem can be solved within any accuracy desired using linear programming techniques. The details of the linear programming problem are not presented here as they are quite technical though straightforward to derive.

The second subproblem (6.11) can be tackled similarly to the first one (6.10). In fact both subproblems are equivalent and can be recast into each other by interchanging the roles of i' and j'. One routine to solve one can also be applied to the other.

A comment is needed on how to choose the constants $F_{i'}(\omega_j)$, $F_{j'}(\omega_j)$ and the weight $\widetilde{W}^*(\omega_j)$. It is proposed to set these iteratively to

$$F_{i'}(\omega_j) = D^{\mathrm{prev}}_{\mathrm{L}i'}(\xi(\omega_j)),\ F_{j'}(\omega_j) = D^{\mathrm{prev}}_{\mathrm{R}j'}(\xi(\omega_j)),$$

$$\widetilde{W}^*_{i'j'}(\omega_j) = \widetilde{W}_{i'j'}(\omega_j) \left| D^{\mathrm{prev}}_{\mathrm{L}i'}(\xi(\omega_j)) D^{\mathrm{prev}}_{\mathrm{R}j'}(\xi(\omega_j)) \right|^{-1},$$

where the superscript prev denotes the outcome of a previous iteration. It is then proposed to solve the optimization problems (6.10) and (6.11) in turn. Again, the idea is that these optimization problems (6.10) and (6.11) more and more resemble the nonlinear problem (6.9).

The second step of the three-step procedure to arrive at an initial estimate consists of mirroring all roots of the denominators $D_{\mathrm{L}i'}$, $i' = 1, \ldots, p'$, and $D_{\mathrm{R}j'}$, $j' = 1, \ldots, q'$, into the region $\mathcal{R}$. This is entirely similar to the SISO case, for details one is referred to Section 6.3. The model resulting from this step will be a feasible solution to the nonlinear problem (6.9).

The third step consists of a final linear optimization of the numerator matrix $N(\xi)$, the denominator matrices $D_{\mathrm{L}}(\xi)$ and $D_{\mathrm{R}}(\xi)$ being fixed to the results of the second step. So, subproblem (6.10) is solved with

$$D_{\mathrm{L}i'}(\xi(\omega_j)) = D_{\mathrm{L}i'}^{\mathrm{fixed}}(\xi(\omega_j)),\ F_{j'}(\omega_j) = D_{\mathrm{R}j'}^{\mathrm{fixed}}(\xi(\omega_j)),$$

$$\widetilde{W}_{i'j'}^{*}(\omega_j) = \widetilde{W}_{i'j'}(\omega_j)\left|D_{\mathrm{L}i'}^{\mathrm{fixed}}(\xi(\omega_j))D_{\mathrm{R}j'}^{\mathrm{fixed}}(\xi(\omega_j))\right|^{-1}.$$

Note that in this case of fixed denominator matrices the optimization problems (6.10), (6.11) and (6.8) become identical. The model resulting from this third step is expected to be a good initial estimate for the nonlinear optimization problem (6.9).

6.5.3 Implementation Aspects

Analogously to the SISO case a proper scaling of (6.12) is obtained by multiplying a constraint corresponding to i', j' and $\omega_j > 1$ in the continuous time case with $f_{i'j'}(\omega_j)$ given by

$$f_{i'j'}(\omega_j) = \omega_j^{-\max(n_{i'j'}, d_{\mathrm{L}i'} + d_{\mathrm{R}j'})}.$$

A proper scaling of (6.9) is achieved by multiplying the constraint corresponding to i', j' and $\omega_j > 1$ in the continuous time case with $f_{i'j'}^2(\omega_j)$.

In addition to this scaling it is advisable to scale the frequency response data $G_{i'j'}(\omega_j)$ by means of multiplication with a constant scaling $s_{i'j'}$ such that each transfer function is approximately of the same order of magnitude. This will improve the conditioning of the linear and nonlinear programming problems. The optimization problem itself remains identical if the weighting $\widetilde{W}_{i'j'}(\omega_j)$ is adjusted accordingly by dividing it by the constant scaling $s_{i'j'}$. Afterwards the resulting numerators should be corrected by dividing each $N_{i'j'}(\xi)$ by $s_{i'j'}$.

Also in the multivariable case it is possible to replace the parameter h_∞^2 in (6.9) by a new parameter $h_{\infty,2}$.

The computation of the partial derivatives of the constraints to the parameters is completely analogous to the SISO case and is not performed here.

6.6 Example

In this section the curve fit procedure is illustrated by means of an example. In particular the procedure is applied to convert a MIMO system from discrete to continuous time.

Consider the 5th order discrete-time MIMO system given by

$$G_{\text{disc}}(z) = \begin{bmatrix} z-0.9 & 0 \\ 0 & z^2-0.5z+0.8 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0.1z^2-0.2 & 0.1z^2-0.2z+0.2 \\ -0.5z^2+0.2z+1 & z^2+0.4z-0.3 \end{bmatrix} \begin{bmatrix} z-0.8 & 0 \\ 0 & z-0.7 \end{bmatrix}^{-1}.$$

The objective is to transform this system to continuous time with a sampling time of 1 second, such that the approximation error is minimized in a weighted ℓ_∞-sense. This can be done with the curve fit procedure developed in this chapter. The system's frequency response $G_{\text{disc}}(e^{i\omega_j})$ is determined for 50 frequencies logarithmically distributed between $10^{-1.5}$ and π. The weighting is chosen equal to

$$\widetilde{W}(\omega_j) = \frac{e^{i\omega_j}-0.8}{e^{i\omega_j}-0.5} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

The model structure is as specified in Section 6.5, with $d_{\text{L1}} = 1$, $d_{\text{L2}} = 2$, $d_{\text{R1}} = 1$, $d_{\text{R2}} = 1$, $n_{11} = 2$, $n_{12} = 2$, $n_{21} = 3$ and $n_{22} = 3$, which coincides with the structure of the discrete-time system. The denominator roots are restricted to the left-half plane, i.e. $r = 0$.

Next, the continuous-time multivariable curve fit problem (6.8) has been solved with the procedure described in Section 6.5. The initial estimate has been determined with the three-step procedure. In the first step initially $D_{\text{L}i'}^{\text{prev}}(s)$ and $D_{\text{R}j'}^{\text{prev}}(s)$ have been set to 1, and 10 iterations have been performed, with m' equal to 4. In the third step m' has been set to 8. The resulting initial approximation is equal to

$$G_{\text{init}}(s) = \begin{bmatrix} s+0.01452 & 0 \\ 0 & s^2+0.17112s+1.59212 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0.4709s^2-0.2033s-0.00482 & 0.2336s^2+0.4370s+0.0584 \\ 0.2996s^3+0.3605s^2-1.1587s+0.3448 & -0.5538s^3-0.5675s^2-0.1208s+1.1214 \end{bmatrix} \cdot \begin{bmatrix} s+0.04661 & 0 \\ 0 & s+0.57598 \end{bmatrix}^{-1}.$$

Figure 6.1 depicts an amplitude Bode diagram of the discrete-time system, the initial continuous-time approximation and the approximation error.

The nonlinear constrained optimization problem has been solved, which yields the optimal continuous-time system

$$G_{\text{cont}}(s) = \begin{bmatrix} s+0.10906 & 0 \\ 0 & s^2+0.19970s+1.64145 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0.1619s^2+0.2093s-0.0745 & 0.1884s^2-0.0204s+0.1951 \\ -0.2520s^3+0.8403s^2-2.1797s+0.6640 & -0.1827s^3-0.3852s^2+0.4011s+1.6644 \end{bmatrix} \cdot \begin{bmatrix} s+0.14328 & 0 \\ 0 & s+0.51650 \end{bmatrix}^{-1}.$$

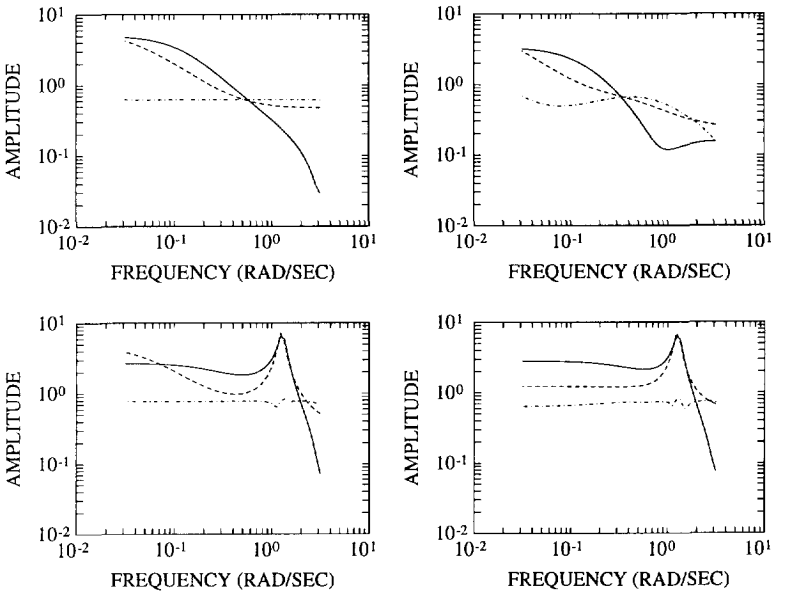

Fig. 6.1: Amplitude Bode diagram of the discrete-time system $G_{\text{disc}}(z)$ (solid), the initial continuous-time approximation $G_{\text{init}}(s)$ (dashed) and the approximation error $\frac{z-0.8}{z-0.5}\left(G_{\text{disc}}(z) - G_{\text{init}}(s)\right)$ (dash-dotted).

Figure 6.2 depicts an amplitude Bode diagram of the discrete-time system, the optimal continuous-time approximation and the approximation error.

The approximation can be optimized by tuning the weighting $\widetilde{W}(\omega_j)$. If, for example, the continuous-time approximation of transfer function $G^{12}_{\text{disc}}(z)$ is considered unsatisfactory, the weighting $\widetilde{W}_{12}(z)$ can be enlarged. The curve fit procedure will then yield a system with a better fit for this entry, obviously at the price of a deteriorated fit for the other entries. If this is considered unacceptable as well, one may consider to increase the order of the continuous-time system.

6.7 Discussion

In this chapter a procedure has been developed to fit a discrete or continuous-time rational transfer function to a set of frequency response data, minimizing a weighted ℓ_∞-criterion. An important feature of the curve fit procedure is that the resulting nominal model has its poles in a user-defined region in the complex plane. It is, for example, straightforward to restrict the model set to the set of stable models. The consequence of this then is that the resulting curve fit model will always be stable.

Obviously, the ℓ_∞-norm which is minimized in the curve fit procedure, is closely

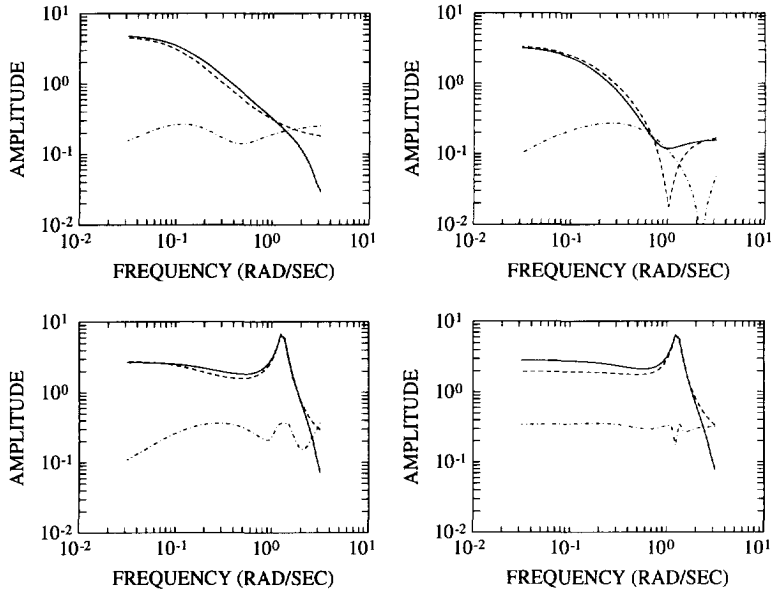

Fig. 6.2: Amplitude Bode diagram of the discrete-time system $G_{\text{disc}}(z)$ (solid), the optimal continuous-time approximation $G_{\text{cont}}(s)$ (dashed) and the approximation error $\frac{z-0.8}{z-0.5}\,(G_{\text{disc}}(z) - G_{\text{cont}}(s))$ (dash-dotted).

related to the H_∞-norm. In the SISO case both are approximately equal to each other if the system's frequency response is smooth and sufficiently many frequencies are used in the evaluation of the ℓ_∞-norm. In the next chapter bounds are derived for the influence of the approximation of the interval $[0, \pi]$ by a frequency grid. There it is also shown that, if the number of frequencies used in the curve fit procedure tends to infinity, the ℓ_∞-error criterion in (6.1) converges to an H_∞-criterion. In practice the number of frequencies can be determined by trial and error. It is recommended to take sufficiently many frequencies compared to the curve fit model order, and afterwards to check if the frequency response of the curve fit model is satisfactory for all frequencies.

Note that also in the MIMO case the norm used in the curve fit problem (6.8) is closely related to the H_∞-norm, as for any $p' \times q'$-matrix A the following bound can easily be established,

$$\max_{i',j'} |A_{i'j'}| \leq \|A\|_2 = \sigma_{\max}(A) \leq \sqrt{p'q'} \max_{i',j'} |A_{i'j'}|.$$

The curve fit procedure has a number of possible applications:

- The procedure can be used to construct a parametric model from frequency response measurements taken from some system. These measurements may either be direct, by means of sinusoidal input excitation, or indirect, for example by

using the frequency response uncertainty regions derived with the identification procedure of Chapter 4 or 5. In this thesis this is in fact the main application of the curve fit procedure. More details are provided in the next chapter, and examples can be found in the chapters to follow.

- As shown in the example in the previous section, the curve fit procedure can be applied to perform a transformation from discrete to continuous time or vice versa. If sufficiently many frequencies are used, this transformation will then in fact be H_∞-optimal.
- The procedure can be used to perform H_∞-optimal model reduction in discrete or continuous time. An example of this is given in Chapter 7.

6.A Proofs

Proof of Lemma 6.2.1

For the proof of part (i) first notice that $\xi + c = 0$ has a solution $|\xi| \leq \rho$ if and only if $|c| \leq \rho$. Next, consider a second order term $\xi^2 + a\xi + b$. This polynomial has all its roots in a disc with radius ρ if and only if the polynomial $(z\rho)^2 + a(z\rho) + b$ has all its roots in the unit disc. From the Schur-Cohn test, see Jacobs (1974, Ch. 3) it follows that this is the case if and only if $b^2 \leq \rho^4$ and $(\rho a)^2 \leq (\rho^2 + b)^2$ or equivalently $b \leq \rho^2$, $-b \leq \rho^2$, $\rho a \leq \rho^2 + b$ and $-\rho a \leq \rho^2 + b$, which yields the desired result (i) by noting that the inequality $-b \leq \rho^2$ is superfluous.

For the proof of part (ii) first notice that $\xi + c = 0$ has a solution $\mathrm{Re}(\xi) \leq r$ if and only if $\mathrm{Re}(c) = c \geq -r$. Next, consider a second order term $\xi^2 + a\xi + b$. This polynomial has roots with real parts less than or equal to r if and only if the polynomial $(s+r)^2 + a(s+r) + b$ has roots with nonnegative real parts. From the Routh-Hurwitz test, see Jacobs (1974, Ch. 3) it follows that this is the case if and only if $2r + a \geq 0$ and $r^2 + ar + b \geq 0$.

Proof of Theorem 6.3.1

The second inequality in part (i) follows from the definition of $n'(\xi)$, $d'(\xi)$. Next, it is noticed that the optimal solution to problem (6.7) has the property

$$h''_\infty = \max_{j=1,\ldots,l} f_{m'}\left((G(\omega_j)d''(\xi(\omega_j)) - n''(\xi(\omega_j)))\,\widetilde{W}^*(\omega_j)\right),$$

which yields the third inequality by applying Lemma 3.4.1. Finally, optimality of of $n''(\xi)$, $d''(\xi)$ implies that for some j',

$$f_{m'}\left((G(\omega_{j'})d'(\xi(\omega_{j'})) - n'(\xi(\omega_{j'})))\,\widetilde{W}^*(\omega_{j'})\right) \geq h''_\infty,$$

which yields the first inequality of part (i) by again applying Lemma 3.4.1. For $m' \to \infty$ the right-hand side in (i) converges to the left-hand side, which proves part (ii).

Chapter 7

System Approximation in H_∞ and ℓ_1

7.1 Introduction

In this chapter the problems of system approximation in H_∞ and ℓ_1, as formulated in Chapter 2, are considered. These problems concern the construction of a nominal model $\widehat{G}(z)$ with bounded, and small, H_∞- resp. ℓ_1-norm of the model error, defined as

$$\max_{G \in \mathcal{S}} \widetilde{W}(z) \left(G(z) - \widehat{G}(z) \right),$$

for some LTI, stable and stably invertible, finite dimensional weight $\widetilde{W}(z)$, and with $\mathcal{S}$ the system uncertainty set as identified in the uncertainty bounding identification procedure. If really the worst-case error is minimized, the system approximation is H_∞- resp. ℓ_1-optimal. In that case the nominal model is determined as

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'(z) \right) \right\|,$$

where the norm is either the H_∞- or the ℓ_1-norm.

This chapter is organized as follows. In the next section the problem of system approximation in H_∞ is considered. In section 7.3 the problem of system approximation in ℓ_1 is considered. Next, in Section 7.4 MIMO extensions are presented. In Section 7.5 some examples are presented. Finally, in Section 7.6 the results are discussed.

7.2 System Approximation in H_∞

7.2.1 Introduction

The H_∞-optimal system approximation problem is defined as

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'(z) \right) \right\|_\infty.$$

The model set $\mathcal{M}$ is chosen to be as follows,

$$\mathcal{M}: \ \widehat{G}'(z) = \frac{n(z)}{d(z)}, \ n(z) = \sum_{k=0}^{n} n_k z^k, \ d(z) = z^d + \sum_{k=0}^{d-1} d_k z^k, \tag{7.1}$$

with the numerator parameters n_k bounded by

$$|n_k| \leq \bar{n}, \ k = 0, \ldots, n, \tag{7.2}$$

and the denominator roots $z_{\mathrm{d},k}$, which are the (possibly complex-valued) solutions to the equation $d(z) = 0$, bounded by

$$|z_{\mathrm{d},k}| \leq \rho, \ k = 1, \ldots, d, \tag{7.3}$$

for some $\rho < 1$. The bound on the roots of the denominator implies stability of the nominal model, which is an essential prerequisite for a bounded H_∞-norm of the model error as defined above. The bound on the numerator parameters is added in order to enable a convergence analysis, as performed later. It will be shown that this bound can be removed under mild asymptotic conditions.

The set $\mathcal{S}$ is the system uncertainty set resulting from the uncertainty bounding identification procedure, with either the deterministic approach (Chapter 4) or the probabilistic approach (Chapter 5). This set is given by

$$\mathcal{S} = \left\{ G(z) \in \mathcal{A} \ \middle| \ G(e^{i\omega_j}) \in \mathcal{P}(\omega_j), \ \forall \omega_j \in \Omega, \ \left| \frac{\mathrm{d}\, G(e^{i\omega})}{\mathrm{d}\, \omega} \right| \leq \beta_1, \ \forall \omega \in [0, \pi] \right\}, \tag{7.4}$$

with $\mathcal{P}(\omega_j)$, $j = 1, \ldots, l$, and β_1 given.

It is mentioned that the H_∞-optimal system approximation problem as formulated above is not solved in this thesis, because of the fact that it is a very complicated problem. However, a closely related problem is formulated and solved. The following relation holds,

$$\begin{aligned} \max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'(z) \right) \right\|_\infty &= \max_{G \in \mathcal{S}} \max_{\omega \in [0,\pi]} \left| \widetilde{W}(e^{i\omega}) \left(G(e^{i\omega}) - \widehat{G}'(e^{i\omega}) \right) \right| = \\ &= \max_{G \in \mathcal{S}} \left| \widetilde{W}(e^{i\omega^*}) \left(G(e^{i\omega^*}) - \widehat{G}'(e^{i\omega^*}) \right) \right|, \end{aligned}$$

where ω^* denotes the (unknown) frequency for which the maximum is achieved. Using the triangle inequality the following bound can be given,

$$\begin{aligned} &\max_{G \in \mathcal{S}} \left| \widetilde{W}(e^{i\omega^*}) \left(G(e^{i\omega^*}) - \widehat{G}'(e^{i\omega^*}) \right) \right| \leq \\ &\leq \max_{G \in \mathcal{S}} \left| \widetilde{W}(e^{i\omega^*}) \left(G(e^{i\omega^*}) - \widehat{G}'_{\mathrm{C}}(\omega^*) \right) \right| + \left| \widetilde{W}(e^{i\omega^*}) \left(\widehat{G}'_{\mathrm{C}}(\omega^*) - \widehat{G}'(e^{i\omega^*}) \right) \right| \leq \\ &\leq \left| \widetilde{W}(e^{i\omega^*}) \right| \max_{G \in \mathcal{S}} \left| G(e^{i\omega^*}) - \widehat{G}'_{\mathrm{C}}(\omega^*) \right| + \max_{\omega \in [0,\pi]} \left| \widetilde{W}(e^{i\omega}) \left(\widehat{G}'_{\mathrm{C}}(\omega) - \widehat{G}'(e^{i\omega}) \right) \right|, \end{aligned} \tag{7.5}$$

for any complex-valued function $\widehat{G}'_{\mathrm{C}}(\omega)$, defined on the interval $[0, \pi]$. Now the system approximation is performed such that the two terms at the right-hand side are both

minimized. The consequence then is that the expression left is also made small, though not necessarily minimized.

The two terms are minimized by adopting a two-step procedure. In the first step the non-parametric, complex-valued so-called *central estimate* $\widehat{G}_C(\omega)$ is determined, which is defined as

$$\widehat{G}_C(\omega) = \arg \min_{\widehat{G}'_C(\omega)\in\mathbb{C}} \max_{G\in\mathcal{S}} \left| G(e^{i\omega}) - \widehat{G}'_C(\omega) \right|, \ \forall \omega \in [0,\pi]. \tag{7.6}$$

As this minimization is performed for each frequency in the interval $[0,\pi]$, it is also performed for the frequency ω^* and, consequently, this minimization corresponds to minimization of the first term in (7.5).

In the second step a nominal model $\widehat{G}(z)$ is determined as the H_∞*-optimal approximation* of this central estimate,

$$\widehat{G}(z) = \arg \min_{\widehat{G}'\in\mathcal{M}} \max_{\omega\in[0,\pi]} \left| \widetilde{W}(e^{i\omega}) \left(\widehat{G}_C(\omega) - \widehat{G}'(e^{i\omega}) \right) \right|, \tag{7.7}$$

which corresponds to minimization of the second term in (7.5).

Both minimization problems involve an evaluation over all frequencies in the set $[0,\pi]$. A computationally feasible solution to the minimization problems is provided by replacing the set $[0,\pi]$ by the set Ω, which consists of a finite number of frequencies. Next, it will be shown that the minimization problems (7.6) and (7.7) are indeed solved if Ω becomes dense in the interval $[0,\pi]$. Thus, the central estimate is computed for a finite number of frequencies,

$$\widehat{G}_C(\omega_j) = \arg \min_{\widehat{G}'_C(\omega_j)\in\mathbb{C}} \max_{G\in\mathcal{S}} \left| G(e^{i\omega_j}) - \widehat{G}'_C(\omega_j) \right|, \ \forall \omega_j \in \Omega. \tag{7.8}$$

In the second step the nominal model is determined as the solution to the following minimization problem,

$$\widehat{G}_\Omega(z) = \arg \min_{\widehat{G}'\in\mathcal{M}} \max_{\omega_j\in\Omega} \left| \widetilde{W}(e^{i\omega_j}) \left(\widehat{G}_C(\omega_j) - \widehat{G}'(e^{i\omega_j}) \right) \right|, \tag{7.9}$$

where also a finite number of frequencies is taken into account. In general the solution $\widehat{G}_\Omega(z)$ of the optimization problem (7.9) is not equal to the solution $\widehat{G}(z)$ of the optimization problem (7.7). However, in the sequel it will be shown that the difference between both models is bounded and that the approximation becomes arbitrarily good if the set Ω becomes dense in the interval $[0,\pi]$.

7.2.2 The Central Estimate

In this subsection the problem is considered of calculating the central estimate $\widehat{G}_C(\omega_j)$ as defined by (7.8). In case the set $\mathcal{S}$, given by (7.4), has been derived with the procedure of Chapter 4 or Chapter 5, the following relation holds,

$$\widehat{G}_C(\omega_j) = \arg \min_{\widehat{G}'_C(\omega_j)\in\mathbb{C}} \max_{G(\omega_j)\in\mathcal{P}(\omega_j)} \left| G(\omega_j) - \widehat{G}'_C(\omega_j) \right|, \ \forall \omega_j \in \Omega.$$

This means that the central estimate can be calculated for a certain frequency ω_j by calculating the center of the uncertainty region $\mathcal{P}(\omega_j)$ for that frequency.

In case the procedure for probabilistic uncertainty bounding identification (Chapter 5) has been used to construct $\mathcal{S}$, the regions $\mathcal{P}(\omega_j)$ are either ellipsoids or rectangulars, for which the centers are straightforwardly computed. However, the outcome of the procedure for deterministic uncertainty bounding identification, described in Chapter 4, can be more complicated. In that case the regions $\mathcal{P}(\omega_j)$ can be arbitrary convex polytopes in the complex plane. It is explained here how to calculate the centers of these polytopes.

Consider a convex polytope $\mathcal{P}$ with vertices v_k, $k = 1, \ldots, m$ (note that the dependency on ω_j is not relevant here and has been dropped for ease of notation). The problem is to calculate the center $\widehat{G}_C$ given by

$$\widehat{G}_C = \arg \min_{\widehat{G}'_C \in \mathbb{C}} \max_{G \in \mathcal{P}} \left| G - \widehat{G}'_C \right| = \arg \min_{\widehat{G}'_C \in \mathbb{C}} \max_{k=1,\ldots,m} \left| v_k - \widehat{G}'_C \right|. \tag{7.10}$$

Introducing the notation

$$\widehat{G}'_C = a + bi, \ a, b \in \mathbb{R},$$

the problem is to find optimal values for a and b. Two procedures to do this will be presented in the sequel.

The first procedure is arrived at by noting that the optimization problem (7.10) is equivalent to the following convex nonlinear constrained optimization problem,

$$\min_{\mu,a,b} \mu \text{ s.t.}$$
$$|v_k - a - bi| \leq \mu, \ k = 1, \ldots, m,$$

or equivelently

$$\min_{\mu,a,b} \mu^2 \text{ s.t.}$$
$$(\mathrm{Re}\,(v_k) - a)^2 + (\mathrm{Im}\,(v_k) - b)^2 \leq \mu^2, \ k = 1, \ldots, m.$$

The latter problem is a convex nonlinear programming problem with a differentiable objective function and differentiable constraints. It can be solved with standard nonlinear programming software. Because of convexity always the global optimum will be found.

The second procedure requires linear programming techniques instead of nonlinear programming. The reason to present this procedure is that software for solving linear programming problems is more widely available than software for nonlinear programming problems. In the procedure the mechanism of Lemma 3.4.1 in Chapter 3 is used. The procedure is presented in the next proposition.

Proposition 7.2.1 *Consider for some $m' \geq 3$ the linear programming problem*

$$\min_{\mu,a,b} \mu \ \textit{s.t.}$$
$$\mathrm{Re}\left(e^{i2\pi k'/m'} v_k\right) - \cos(2\pi k'/m')a + \sin(2\pi k'/m')b \leq \mu,$$
$$k' = 1, \ldots, m', \ k = 1, \ldots, m.$$

Let the solution of this LP problem be given by μ_{opt}, a_{opt}, b_{opt} *and let* $\widehat{G}_{\text{C}}$ *be given by (7.10). Then*

$$\textbf{(i)} \quad \mu_{\text{opt}} \leq \max_k \left| v_k - \widehat{G}_{\text{C}} \right| \leq \max_k |v_k - a_{\text{opt}} - b_{\text{opt}} i| \leq \frac{\mu_{\text{opt}}}{\cos(\pi/m')},$$

$$\textbf{(ii)} \quad \lim_{m' \to \infty} a_{\text{opt}} + b_{\text{opt}} i = \widehat{G}_{\text{C}}.$$

Proof: See Appendix 7.A. □

The linear programming approach thus provides an approximate solution for the problem of calculating the center $\widehat{G}_{\text{C}}$ of the polytope $\mathcal{P}$. The approximation can be made arbitrarily accurate by choosing m' large enough. The price paid for this is an increasing number of linear constraints, and hence, an increasing computational complexity.

Altogether, with either of the two methods presented the central estimate $\widehat{G}_{\text{C}}(\omega_j)$ can be calculated for all frequencies ω_j in the set Ω.

Finally, it is mentioned that $\widehat{G}_{\text{C}}(\omega)$ as defined in (7.6) is a continuous and smooth function in the entire frequency interval $[0, \pi]$, with the derivative bounded by

$$\left| \frac{\mathrm{d}\, \widehat{G}_{\text{C}}(\omega)}{\mathrm{d}\, \omega} \right| \leq \beta_1, \forall \omega \in [0, \pi], \tag{7.11}$$

which is due to the fact that any $G \in \mathcal{S}$ satisfies this bound. And with Lemma 4.4.7 in Chapter 4 also a bound β_2 can be established for the amplitude of $\widehat{G}_{\text{C}}(\omega)$,

$$\left| \widehat{G}_{\text{C}}(\omega) \right| \leq \beta_2, \ \forall \omega \in [0, \pi]. \tag{7.12}$$

This bound again holds because any system $G \in \mathcal{S}$ satisfies the bound. Both bounds are of use in the next subsection when establishing convergence properties of the system approximation procedure.

7.2.3 H_∞-Optimal Approximation

The second step in the two-step procedure is the construction of the nominal model $\widehat{G}(z)$ within the model set $\mathcal{M}$, that optimally fits the central estimate $\widehat{G}_{\text{C}}(\omega_j)$ in H_∞-sense, as expressed in (7.9). In fact this problem has been solved by the discrete-time version of the curve fit algorithm presented in the previous chapter. In comparison to the procedure as developed there it is only necessary to also include the linear constraints on the numerator parameters (7.2).

This implies that algorithms are available to carry out both steps of the two-step system approximation procedure in case the frequency set Ω is considered. Unfortunately in principle a good quality of the model for frequencies in the set Ω does not necessarily imply a good quality for frequencies outside this grid. Theoretically the nominal model might even be arbitrarily bad between two subsequent frequencies in Ω. However, it is shown next that this is not the case. It will be shown that the error due to the frequency sampling is bounded and that this error vanishes if more and more frequencies are taken into account, i.e. if the set Ω becomes dense in the interval $[0, \pi]$. For this purpose first the following auxiliary lemma is established.

Lemma 7.2.2 *Any element $\widehat{G}'(z)$ of the model set $\mathcal{M}$, as specified by (7.1), (7.2) and (7.3), satisfies the following bounds,*

$$\textbf{(i)} \qquad \left\|\widehat{G}'(z)\right\|_\infty \leq \beta_3(n,d,\rho,\bar{n}) := \frac{(n+1)\bar{n}}{(1-\rho)^d},$$

$$\textbf{(ii)} \qquad \left\|\frac{\mathrm{d}\,\widehat{G}'(z)}{\mathrm{d}\,z}\right\|_\infty \leq \beta_4(n,d,\rho,\bar{n}) := \frac{\bar{n}n(n+1)}{2(1-\rho)^d} + \frac{\bar{n}(n+1)d(1+\rho)^{d-1}}{(1-\rho)^{2d}}.$$

Proof: See Appendix 7.A. □

Next, define the scalars β_5 and β_6 as,

$$\beta_5 := \left\|\widetilde{W}(z)\right\|_\infty, \quad \beta_6 := \left\|\frac{\mathrm{d}\,\widetilde{W}(z)}{\mathrm{d}\,z}\right\|_\infty, \tag{7.13}$$

which are both finite due to the fact that $\widetilde{W}(z)$ is stable and finite dimensional.

Then the following theorem establishes boundedness and convergence of the identified nominal model $\widehat{G}_\Omega(z)$.

Theorem 7.2.3 *Define λ_j, $j = 1,\ldots,l$, as*

$$\begin{aligned}
\lambda_1 &:= \max\{2\omega_1, \omega_2 - \omega_1\},\\
\lambda_j &:= \max\{\omega_j - \omega_{j-1}, \omega_{j+1} - \omega_j\},\ j = 2,\ldots,l-1,\\
\lambda_l &:= \max\{\omega_l - \omega_{l-1}, 2(\pi - \omega_l)\}.
\end{aligned}$$

Consider the bound β_1 in (7.11), β_2 in (7.12), $\beta_3(n,d,\rho,\bar{n})$ as defined in part (i) of Lemma 7.2.2, $\beta_4(n,d,\rho,\bar{n})$ as defined in part (ii) of Lemma 7.2.2, and β_5 and β_6 as defined in (7.13). Define,

$$\beta(n,d,\rho,\bar{n}) := \beta_1\beta_5 + \beta_2\beta_6 + \beta_3(n,d,\rho,\bar{n})\beta_6 + \beta_4(n,d,\rho,\bar{n})\beta_5.$$

Then the following bound can be established for any model $\widehat{G}'(z) \in \mathcal{M}$,

$$\textbf{(i)} \qquad \max_{\omega\in[0,\pi]} \left|\widetilde{W}(e^{i\omega})\left(\widehat{G}_{\mathrm{C}}(\omega) - \widehat{G}'(e^{i\omega})\right)\right| \leq$$

$$\leq \max_{j=1,\ldots,l} \left[\left|\widetilde{W}(e^{i\omega_j})\left(\widehat{G}_{\mathrm{C}}(\omega_j) - \widehat{G}'(e^{i\omega_j})\right)\right| + \frac{1}{2}\lambda_j\beta(n,d,\rho,\bar{n})\right].$$

Moreover, let $\widehat{G}_\Omega(z)$ be defined by (7.9), and $\widehat{G}(z)$ by (7.7). Then, up to uniqueness of the solutions,

$$\textbf{(ii)} \qquad \lim_{\Omega\to[0,\pi]} \widehat{G}_\Omega(z) = \widehat{G}(z).$$

This convergence still takes place if

$$\bar{n} \to \infty,\ n \to \infty,\ d \to \infty,\ \rho \to 1,$$

provided $\max_j \lambda_j$ *tends to 0 sufficiently fast,*

$$\left(\max_j \lambda_j\right) \beta(n, d, \rho, \bar{n}) \to 0.$$

In particular, if n, d *and* ρ *are fixed,* $\bar{n}$ *may tend to infinity provided*

$$\left(\max_j \lambda_j\right) \bar{n} \to 0.$$

Proof: See Appendix 7.A. □

The theorem states that the model $\widehat{G}_\Omega(z)$ in (7.9) converges to the optimal model $\widehat{G}(z)$ in (7.7) if more and more frequencies ω_j are used in the curve fit procedure. Note that the choice of frequency grid Ω is a user's choice, independent of the experimental data available or whatsoever. With both the procedure for deterministic uncertainty bounding identification, developed in Chapter 4, and the procedure for probabilistic uncertainty bounding identification, developed in Chapter 5, the frequencies ω_j can be chosen arbitrarily. Subsequently, these frequencies can be used in the curve fit algorithm developed in Chapter 6. The price paid for using more frequencies is an increasing computational effort.

Note that in theorem 7.2.3 the speed of convergence is related to β. However, this is a worst-case value. In practice convergence can be much faster. In particular it is often not necessary to bound the numerator parameters with the bound $\bar{n}$, especially if n and d are relatively small.

Remark 7.2.4 *Due to the use of the triangly inequality in (7.5) the system approximation is suboptimal rather than optimal. However, consider the situation that the system uncertainty set* $\mathcal{S}$ *converges to one frequency response, which is then equal to* $\widehat{G}_C(\omega)$. *In that case the first term at the right-hand side in (7.5) vanishes. Then the system approximation is optimal, and the model* $\widehat{G}_\Omega(z)$ *is the solution to the* H_∞*-optimal system approximation problem. In particular this situation occurs if the number of data* N *used in the uncertainty bounding identification procedure of Chapter 5, tends to infinity. If the data are informative enough, the system uncertainty set* $\mathcal{S}$ *will typically converge to the true system* $G_0(z)$.

Remark 7.2.5 *The result can be compared to results obtained in the area of identification in* H_∞. *There the aim is to identify a nominal model on the basis of frequency response uncertainty regions, such that the* H_∞*-norm of the model error is bounded, and such that the model error vanishes under the following asymptotic conditions:*

- *The frequency grid* Ω *becomes dense in* $[0, \pi]$.
- *The model order (*n *and* d*) tends to infinity.*
- *The noise level tends to zero, or, in other words, the system uncertainty set* $\mathcal{S}$ *converges to the central estimate* $\widehat{G}_C(\omega)$.

In fact the result of Theorem 7.2.3 is much stronger than this. This result states that the model error is not only bounded in H_∞-norm, but is also minimized in case the first asymptotic condition of the three listed above is satisfied. In particular the optimality of the algorithm is independent of the model order n, d. Obviously, the optimality property implies the desired convergence property. The H_∞-norm of the model error vanishes if the three asymptotic conditions listed above are satisfied.

A drawback of the procedure is perhaps that it requires nonlinear, nonconvex optimizations to compute the optimal model $\widehat{G}_\Omega(z)$. However, this drawback is eliminated in case the denominator $d(z)$ is chosen fixed. Then the model parametrization is linear and the optimization problem is convex, which means that always the global optimum is found, see Chapter 6.

Remark 7.2.6 *Theorem 7.2.3 has been stated with respect to any continuous function $\widehat{G}_C(\omega)$ defined on the interval $[0, \pi]$. This implies that the results are straightforwardly applicable to the situation that the curve fit algorithm presented in the previous chapter is applied to e.g. conversion from continuous to discrete time, or discrete-time model reduction. In those cases $\widehat{G}_C(\omega)$ is equal to the frequency response of the given model that needs to be reduced or converted to discrete time. The theorem then states that the procedure is H_∞-optimal in case the frequency grid Ω is chosen sufficiently fine. By means of the computable derivative bound $\beta(n, d, \rho, \bar{n})$ in Theorem 7.2.3 it can be determined beforehand how many frequencies should be used in the set Ω in order to have an approximation of H_∞-optimality within a certain specified accuracy.*

Remark 7.2.7 *Results similar to Theorem 7.2.3 are obtainable in case a continuous-time model $\widehat{G}(s)$ is determined with the curve fit procedure described in Chapter 6. The situation becomes slightly more complicated compared to the discrete-time case, as the frequency axis goes to infinity. Boundedness for the high-frequency range can be established by putting proper constraints on the numerator and denominator parameters. In Baratchart and Leblond (1993) typical difficulties for continuous-time frequency domain identification are highlighted and possible solution strategies are indicated.*

7.3 System Approximation in ℓ_1

7.3.1 Introduction

The ℓ_1-optimal system approximation problem is defined as

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'(z) \right) \right\|_{\ell_1}. \tag{7.14}$$

The model set $\mathcal{M}$ considered here is different from the one used in the context of system approximation in H_∞. It is chosen to be as follows,

$$\mathcal{M} : \ \widehat{G}'(z) = \frac{n(z)}{d(z)}, \ n(z) = \sum_{k=0}^{n} n_k z^k, \tag{7.15}$$

with the denominator $d(z)$ some *fixed* and *given* polynomial of degree d, with $d \geq n$, which has all its roots inside the open unit disc. Thus, the model is linearly parametrized, in terms of the parameters n_k. This linear model parametrization is used in order to obtain tractable algorithms. As will become clear later, no tractable solution to the problem of system approximation in ℓ_1 can be provided in case the denominator $d(z)$ is parametrized as well. Finally, it is mentioned that the given parametrization (7.15) allows the user to effectively control the model order. This is the reason why the linear model parametrization in terms of basis functions, as used in Chapter 4 and 5, is not used here.

The set $\mathcal{S}$ is the system uncertainty set resulting from the uncertainty bounding identification procedure, with either the deterministic approach (Chapter 4) or the probabilistic approach (Chapter 5). This set is given by

$$\mathcal{S} = \left\{ G(z) \in \mathcal{A} \;\middle|\; G(z) = W^{-1}(z) \sum_{k=0}^{\infty} p(k) z^{-k},\ p(k) \in [p_l(k), p_u(k)],\ k = 0, \ldots, \infty \right\}, \tag{7.16}$$

with $\{p_l(k)\}_{k=0,\ldots,\infty}$, $\{p_u(k)\}_{k=0,\ldots,\infty}$ given and $W(z)$ some given stable and stably invertible finite dimensional and biproper weighting. It is assumed that the following equality is satisfied,

$$p_l(k) = -p_u(k),\ k = \bar{k}, \ldots, \infty,$$

for some (large) integer $\bar{k}$. If one of the uncertainty bounding identification procedures presented in this thesis has been used to construct the set $\mathcal{S}$, such a $\bar{k}$ always exists.

The problem as formulated above is not solved in this thesis. Analogously to the solution strategy in the H_∞-case, the problem is approximated by a closely related problem, which is solved. However, in addition to this it will be shown that this approximation provides an exact solution for the ℓ_1-optimal system approximation problem in the special case that the two weights $\widetilde{W}(z)$ and $W(z)$ are identical.

Using the triangle inequality and the multiplicative property of the ℓ_1-norm, the following bound can be given,

$$\begin{aligned}
&\max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'(z) \right) \right\|_{\ell_1} \leq \\
&\leq \max_{G \in \mathcal{S}} \left\| \widetilde{W}(z) \left(G(z) - \widehat{G}'_C(z) \right) \right\|_{\ell_1} + \left\| \widetilde{W}(z) \left(\widehat{G}'_C(z) - \widehat{G}'(z) \right) \right\|_{\ell_1} \leq \\
&\leq \max_{G \in \mathcal{S}} \left\| \widetilde{W} W^{-1} \right\|_{\ell_1} \left\| W(z) \left(G(z) - \widehat{G}'_C(z) \right) \right\|_{\ell_1} + \left\| \widetilde{W}(z) \left(\widehat{G}'_C(z) - \widehat{G}'(z) \right) \right\|_{\ell_1}.
\end{aligned} \tag{7.17}$$

Now the system approximation is performed such that the two terms at the right-hand side are both minimized. The consequence then is that the expression left is also made small, though not necessarily minimized.

The two terms at the right-hand side in (7.17) are minimized by adopting a two-step procedure. In the first step the so-called *central estimate* is calculated, which is defined as

$$\widehat{G}_C(z) = \arg \min_{\widehat{G}'_C \in \mathcal{A}} \max_{G \in \mathcal{S}} \left\| W(z) \left(G(z) - \widehat{G}'_C(z) \right) \right\|_{\ell_1}, \tag{7.18}$$

which corresponds to minimizing the first term in (7.17).

In the second step a nominal model $\widehat{G}(z)$ is determined as the *ℓ_1-optimal approximation* of this central estimate,

$$\widehat{G}(z) = \arg\min_{\widehat{G}' \in \mathcal{M}} \left\| \widetilde{W}(z) \left(\widehat{G}_{\mathrm{C}}(z) - \widehat{G}'(z) \right) \right\|_{\ell_1}, \tag{7.19}$$

which corresponds to minimization of the second term in (7.17).

The two-step procedure has been motivated by the triangle inequality, which introduces conservatism. This means that the ℓ_1-optimal system approximation problem formulated in (7.14) is not solved exactly. However, it appears that if $W(z) = \widetilde{W}(z)$, the inequalities in (7.17) become equalities, in which case the ℓ_1-optimal system approximation problem (7.14) is solved exactly. This is formulated in the next theorem.

Theorem 7.3.1 *Consider the set $\mathcal{S}$ given in (7.16), any model $\widehat{G}'(z) \in \mathcal{A}$ and the central estimate $\widehat{G}_{\mathrm{C}}(z)$ as defined in (7.18). Then,*

$$\begin{aligned} &\max_{G \in \mathcal{S}} \left\| W(z) \left(G(z) - \widehat{G}'(z) \right) \right\|_{\ell_1} = \\ &= \max_{G \in \mathcal{S}} \left\| W(z) \left(G(z) - \widehat{G}_{\mathrm{C}}(z) \right) \right\|_{\ell_1} + \left\| W(z) \left(\widehat{G}_{\mathrm{C}}(z) - \widehat{G}'(z) \right) \right\|_{\ell_1}. \end{aligned}$$

This implies that the solution of problem (7.14) is identical to the solution of problem (7.19) if $\widetilde{W}(z) = W(z)$.

Proof: Given in the proof of Theorem 7.3.2. □

The first step in the two-step procedure is the construction of the central estimate $\widehat{G}_{\mathrm{C}}(z)$ as defined in (7.18). The following theorem gives an expression for this central estimate.

Theorem 7.3.2 *Consider the set $\mathcal{S}$ given in (7.16). Define*

$$P_{\mathrm{C}}(z) := \sum_{k=0}^{\infty} p_{\mathrm{c}}(k) z^{-k},$$

with

$$p_{\mathrm{c}}(k) := \frac{1}{2} \left(p_{\mathrm{l}}(k) + p_{\mathrm{u}}(k) \right), \quad k = 0, \ldots, \infty.$$

Then the central estimate $\widehat{G}_{\mathrm{C}}(z)$ defined in (7.18) is equal to

$$\widehat{G}_{\mathrm{C}}(z) = W^{-1}(z) P_{\mathrm{C}}(z).$$

Proof: See Appendix 7.A. □

The theorem shows that the central estimate can straightforwardly be computed. As $p_{\mathrm{c}}(k) = 0$, $\forall\, k \geq \bar{k}$, it is a stable, finite dimensional transfer function, although its order can be very high.

7.3.2 ℓ_1-Optimal Model Reduction for Linearly Parametrized Reduced Order Models

The second step in the two-step procedure is the construction of the nominal model $\widehat{G}(z)$ within the model set $\mathcal{M}$ that optimally fits the central estimate $\widehat{G}_{\mathrm{C}}(z)$ in ℓ_1-sense, as expressed in (7.19). In fact this problem is an ℓ_1-*optimal model reduction* problem. An algorithm will next be presented which solves this problem in case of the linearly parametrized model set $\mathcal{M}$ given in (7.15).

Any model in the model set $\mathcal{M}$ is equal to a quotient of two polynomials in z. By dividing numerator and denominator by z^d this is changed to a quotient of two polynomials in z^{-1} (note that the degree of $n(z)$ is smaller than or equal to the degree of $d(z)$),

$$\mathcal{M}: \ \widehat{G}'(z) = \frac{\widetilde{n}(z^{-1})}{\widetilde{d}(z^{-1})},$$

with

$$\widetilde{n}(z^{-1}) = n(z)z^{-d} = \sum_{k=0}^{n} n_k z^{k-d} = \sum_{k=d-n}^{d} n_{d-k} z^{-k},$$

and $\widetilde{d}(z^{-1}) = d(z)z^{-d}$ a given polynomial in z^{-1}.

Next, as $\widehat{G}_{\mathrm{C}}(z)$ and $\widetilde{W}(z)$ are finite dimensional transfer functions, they can always be written as a quotient of two polynomials in z^{-1},

$$\begin{aligned} \widehat{G}_{\mathrm{C}}(z) &= \frac{n_{\mathrm{G}}(z^{-1})}{d_{\mathrm{G}}(z^{-1})}, \\ \widetilde{W}(z) &= \frac{n_{\mathrm{W}}(z^{-1})}{d_{\mathrm{W}}(z^{-1})}, \end{aligned}$$

with $n_{\mathrm{G}}(z^{-1})$, $d_{\mathrm{G}}(z^{-1})$, $n_{\mathrm{W}}(z^{-1})$ and $d_{\mathrm{W}}(z^{-1})$ given polynomials in z^{-1} (of finite degree).

Moreover, for future reference the following polynomials are introduced,

$$\begin{aligned} a(z^{-1}) &:= d_{\mathrm{W}}(z^{-1}) d_{\mathrm{G}}(z^{-1}) \widetilde{d}(z^{-1}), \\ b(z^{-1}) &:= n_{\mathrm{W}}(z^{-1}) d_{\mathrm{G}}(z^{-1}), \\ c(z^{-1}) &:= n_{\mathrm{W}}(z^{-1}) n_{\mathrm{G}}(z^{-1}) \widetilde{d}(z^{-1}), \end{aligned}$$

which can all three be calculated. For convenience the following notation is introduced,

$$\begin{aligned} a(z^{-1}) &= \sum_{k=0}^{\infty} a(k) z^{-k}, \ a(k) = 0, \ \forall\, k > \deg\left(a(z^{-1})\right), \\ b(z^{-1}) &= \sum_{k=0}^{\infty} b(k) z^{-k}, \ b(k) = 0, \ \forall\, k > \deg\left(b(z^{-1})\right), \\ c(z^{-1}) &= \sum_{k=0}^{\infty} c(k) z^{-k}, \ c(k) = 0, \ \forall\, k > \deg\left(c(z^{-1})\right). \end{aligned}$$

The ℓ_1-optimal model reduction problem

$$\min_{\widehat{G}' \in \mathcal{M}} \left\| \widetilde{W}(z) \left(\widehat{G}_{\mathrm{C}}(z) - \widehat{G}'(z) \right) \right\|_{\ell_1}$$

is equivalent to

$$\min_{\widetilde{n}(z^{-1}), \Delta(z)} \|\Delta(z)\|_{\ell_1} \text{ s.t.}$$

$$\Delta(z) = \frac{n_{\mathrm{W}}(z^{-1})}{d_{\mathrm{W}}(z^{-1})} \left(\frac{n_{\mathrm{G}}(z^{-1})}{d_{\mathrm{G}}(z^{-1})} - \frac{\widetilde{n}(z^{-1})}{\widetilde{d}(z^{-1})} \right),$$

which is equivalent to

$$\min_{\widetilde{n}(z^{-1}), \Delta(z)} \sum_{k=0}^{\infty} |\delta(k)| \text{ s.t.}$$

$$\Delta(z) = \sum_{k=0}^{\infty} \delta(k) z^{-k},$$

$$\Delta(z) d_{\mathrm{W}}(z^{-1}) d_{\mathrm{G}}(z^{-1}) \widetilde{d}(z^{-1}) + \widetilde{n}(z^{-1}) n_{\mathrm{W}}(z^{-1}) d_{\mathrm{G}}(z^{-1}) + - n_{\mathrm{W}}(z^{-1}) n_{\mathrm{G}}(z^{-1}) \widetilde{d}(z^{-1}) = 0,$$

which in turn is equivalent to

$$\min_{\{n_k\}_{k=0,\ldots,\infty}, \{\delta(k)\}_{k=0,\ldots,\infty}} \sum_{k=0}^{\infty} |\delta(k)| \text{ s.t.}$$

$$a(z^{-1}) \sum_{k=0}^{\infty} \delta(k) z^{-k} + b(z^{-1}) \sum_{k=0}^{n} n_k z^{k-d} - c(z^{-1}) = 0,$$

which is in fact an infinite dimensional nonlinear constrained optimization problem. By a variable transformation (Lu and Wang, 1988) this can be transformed into an infinite dimensional linear programming problem. Introduce for that purpose $\delta_{\mathrm{p}}(k)$ and $\delta_{\mathrm{n}}(k)$ such that

$$\delta(k) = \delta_{\mathrm{p}}(k) - \delta_{\mathrm{n}}(k), \ \min \{\delta_{\mathrm{p}}(k), \delta_{\mathrm{n}}(k)\} = 0, \ k = 0, \ldots, \infty,$$

and, hence, $|\delta(k)| = \delta_{\mathrm{p}}(k) + \delta_{\mathrm{n}}(k)$. Then the nonlinear optimization problem is equivalent to

$$\min_{\{n_k\}, \{\delta_{\mathrm{p}}(k)\}, \{\delta_{\mathrm{n}}(k)\}} \sum_{k=0}^{\infty} (\delta_{\mathrm{p}}(k) + \delta_{\mathrm{n}}(k)) \text{ s.t.}$$

$$a(z^{-1}) \sum_{k=0}^{\infty} (\delta_{\mathrm{p}}(k) - \delta_{\mathrm{n}}(k)) z^{-k} + b(z^{-1}) \sum_{k=d-n}^{d} n_{d-k} z^{-k} - c(z^{-1}) = 0,$$

$$\delta_{\mathrm{p}}(k) \geq 0, \ \delta_{\mathrm{n}}(k) \geq 0, \ k = 0, \ldots, \infty.$$

This equivalence is due to the fact that in the optimum indeed either $\delta_p(k)$ or $\delta_n(k)$ is 0, see Lu and Wang (1988). The constraint in the given optimization problem states that a polynomial of infinite degree should be equal to 0. The polynomial has coefficients which tend to zero if k tends to infinity. The polynomial is equal to 0 if and only if all its coefficients are equal to 0. This means that the given optimization problem is equivalent to the following one,

$$\min_{\{n_k\},\{\delta_p(k)\},\{\delta_n(k)\}} \sum_{k=0}^{\infty} (\delta_p(k) + \delta_n(k)) \text{ s.t.}$$

$$\sum_{j=0}^{k} a(k-j)(\delta_p(j) - \delta_n(j)) + \sum_{j=d-n}^{\min(d,k)} b(j-k)n_{d-j} - c(k) = 0, \; k = 0, \ldots, \infty, \quad (7.20)$$

$$\delta_p(k) \geq 0, \; \delta_n(k) \geq 0, \; k = 0, \ldots, \infty,$$

which is an infinite dimensional linear programming problem, with infinite unknowns and infinite constraints. Note that the notation is such that the summation $\sum_{j=d-n}^{\min(d,k)}$ just cancels if $\min(d,k) < d-n$.

Resuming, the solution to the ℓ_1-optimal model reduction problem (7.19) is given by the solution of the infinite dimensional linear programming problem (7.20). As there are no algorithms which can solve an infinite dimensional LP problem, a solution to this problem will be provided by approximating it by a finite dimensional LP problem. The accuracy of the approximation will be specified and it will be shown that, when increasing the size of the finite dimensional LP problem, the approximate solution converges to the solution of (7.20). Analogously to the truncation applied in the context of ℓ_1-optimal control design (McDonald and Pearson, 1991) a truncation parameter l is introduced, such that

$$l \geq \max\{\deg(a(z^{-1}), \deg(b(z^{-1}), \deg(c(z^{-1})\}.$$

Consider the following finite dimensional linear programming problem, derived from (7.20) by truncation,

$$\min_{\{n_k\},\{\delta_p(k)\},\{\delta_n(k)\}} \sum_{k=0}^{l} (\delta_p(k) + \delta_n(k)) \text{ s.t.}$$

$$\sum_{j=0}^{k} a(k-j)(\delta_p(j) - \delta_n(j)) + \sum_{j=d-n}^{\min(d,k)} b(j-k)n_{d-j} - c(k) = 0, \; k = 0, \ldots, l, \quad (7.21)$$

$$\delta_p(k) \geq 0, \; \delta_n(k) \geq 0, \; k = 0, \ldots, l,$$

which is an LP problem with $n + 2l + 3$ unknowns subject to $3(l+1)$ linear equality and inequality constraints. The calculable solution to this problem (7.21) provides an approximate solution for problem (7.20). The accuracy of the approximation is specified in the following theorem.

Theorem 7.3.3 *Consider the infinite dimensional LP problem (7.20) and the finite dimensional LP problem (7.21). Denote the (unknown) solution of (7.20) by*

$$\left\{\{n_k^{\text{opt}}\}_{k=0,\ldots,n},\ \{\delta_{\text{p}}^{\text{opt}}(k)\}_{k=0,\ldots,\infty},\ \{\delta_{\text{n}}^{\text{opt}}(k)\}_{k=0,\ldots,\infty}\right\},$$

and define the corresponding transfer functions as

$$\begin{aligned}\widehat{G}_{\text{opt}}(z) &= \frac{\sum_{k=0}^{n} n_k^{\text{opt}} z^{k-d}}{\widetilde{d}(z^{-1})},\\ \Delta_{\text{opt}}(z) &= \sum_{k=0}^{\infty}\left(\delta_{\text{p}}^{\text{opt}}(k) - \delta_{\text{n}}^{\text{opt}}(k)\right) z^{-k}.\end{aligned}$$

Denote the (calculable) solution of (7.21) by

$$\left\{\{n_k^{\text{T,opt}}\}_{k=0,\ldots,n},\ \{\delta_{\text{p}}^{\text{T,opt}}(k)\}_{k=0,\ldots,l},\ \{\delta_{\text{n}}^{\text{T,opt}}(k)\}_{k=0,\ldots,l}\right\},$$

and define the corresponding transfer functions as

$$\begin{aligned}\widehat{G}_{\text{T,opt}}(z) &= \frac{\sum_{k=0}^{n} n_k^{\text{T,opt}} z^{k-d}}{\widetilde{d}(z^{-1})},\\ \Delta_{\text{T,opt}}(z) &= \sum_{k=0}^{l}\left(\delta_{\text{p}}^{\text{T,opt}}(k) - \delta_{\text{n}}^{\text{T,opt}}(k)\right) z^{-k}.\end{aligned}$$

Then,

$$\textbf{(i)}\quad \|\Delta_{\text{T,opt}}(z)\|_{\ell_1} \le \|\Delta_{\text{opt}}(z)\|_{\ell_1} \le \left\|\widetilde{W}(z)\left(\widehat{G}_{\text{C}}(z) - \widehat{G}_{\text{T,opt}}(z)\right)\right\|_{\ell_1},$$

$$\textbf{(ii)}\quad \lim_{l\to\infty}\left(\left\|\widetilde{W}(z)\left(\widehat{G}_{\text{C}}(z) - \widehat{G}_{\text{T,opt}}(z)\right)\right\|_{\ell_1} - \|\Delta_{\text{T,opt}}(z)\|_{\ell_1}\right) = 0,$$

and hence, up to uniqueness of the solutions, the solution of LP problem (7.21) converges to the solution of LP problem (7.20),

$$\lim_{l\to\infty}\left\|\widetilde{W}(z)\left(\widehat{G}_{\text{opt}}(z) - \widehat{G}_{\text{T,opt}}(z)\right)\right\|_{\ell_1} = 0.$$

Proof: See Appendix 7.A. □

The theorem states that the solution of the finite dimensional LP problem (7.21) can be made arbitrarily close to the solution of the model reduction problem (7.19), which is equivalent to the solution of the infinite dimensional LP problem (7.20). The accuracy of the approximation can be calculated with the upper and lower bound established in part (i). Part (ii) states that this upper and lower bound converge to the optimal solution of (7.20) if the size of the LP problem (7.21) is increased. This means that iteratively an acceptable value of l can be determined, yielding a solution within any prespecified accuracy desired. Obviously, the price paid for high accuracy is a high computational complexity.

Remark 7.3.4 *A linear parametrization has been used for the model set $\mathcal{M}$ in (7.15). This has been done in order to end up with a tractable numerical problem, as the linear programming problem (7.21) is easily solvable with standard software. Parametrizing the denominator $d(z)$ as well would lead to a highly untractable nonconvex infinite dimensional nonlinear programming problem, which generally can not be solved. In the approach adopted here the denominator is specified a priori, and then the numerator parameters are tuned optimally in ℓ_1-sense. In practice the denominator can be specified by first applying any other model reduction tool to the model $\widehat{G}_C(z)$, for example Hankel-norm (Glover, 1984), balanced model reduction (Al-Saggaf and Franklin, 1988), or frequency weighted balanced model reduction (Wortelboer and Bosgra, 1992; Wortelboer, 1994).*

7.4 The Multivariable Case

In this section extensions of the problem of system approximation in H_∞ and ℓ_1 to the multivariable case are presented. Only the main differences with the SISO case are discussed and most details are omitted.

In case of a MIMO system $G_0(z)$ with q' inputs and p' outputs, the uncertainty bounding identification procedures presented in the Chapters 4 and 5 yield a system uncertainty set $\mathcal{S}$ in the following format

$$\mathcal{S} = \{G(z) \mid G_{i'j'}(z) \in \mathcal{S}_{i'j'},\ i' = 1,\ldots,p',\ j' = 1,\ldots,q'\}.$$

Here $\mathcal{S}_{i'j'}$ is the system uncertainty set for one entry of the MIMO system, which has a format as specified by (2.3) for the H_∞-case, and (2.4) for the ℓ_1-case. The special structure of this multivariable uncertainty description is due to the fact that the system uncertainty is evaluated for each entry of the MIMO transfer function separately.

In the H_∞-case the objective is defined to be: construct a nominal model as the solution to the following minimization problem,

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{G \in \mathcal{S}} \max_{i',j'} \left\| \widetilde{W}_{i'j'}(z) \left(G_{i'j'}(z) - \widehat{G}'_{i'j'}(z) \right) \right\|_\infty .$$

The model set $\mathcal{M}$ is as specified in Section 6.5 and the system uncertainty set $\mathcal{S}$ is as specified above. Each transfer function $\widetilde{W}_{i'j'}(z)$ is a stable and stably invertible, LTI, finite dimensional weighting.

Similar to the SISO case this problem is approximately solved by a two-step procedure, which is motivated by the triangle inequality. In the first step the central estimate $\widehat{G}_C(\omega_j)$ is computed, which is defined as

$$\widehat{G}_C^{i'j'}(\omega_j) = \arg \min_{\widehat{G}'_C(\omega_j) \in \mathbb{C}} \max_{G_{i'j'} \in \mathcal{S}_{i'j'}} \left| G_{i'j'}(e^{i\omega_j}) - \widehat{G}'_C(\omega_j) \right|,\ \forall \omega_j \in \Omega.$$

This step can be solved with the algorithms developed for the SISO case. In the second step the H_∞-optimal parametric approximation of this central estimate is computed

as the solution to the optimization problem,

$$\widehat{G}_\Omega(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{i'j'} \max_{\omega_j \in \Omega} \left| \widetilde{W}_{i'j'}(e^{i\omega_j}) \left(\widehat{G}_C^{i'j'}(\omega_j) - \widehat{G}'_{i'j'}(e^{i\omega_j}) \right) \right|.$$

This second step is solved by the MIMO curve fit algorithm developed in Section 6.5. Entirely similar to the SISO case optimality properties of the model $\widehat{G}_\Omega(z)$ can be derived for the asymptotic situation that Ω becomes dense in $[0, \pi]$. Details are omitted.

In the ℓ_1-case the objective is defined to be: construct a nominal model as the solution to the following minimization problem,

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{G \in \mathcal{S}} \max_{i',j'} \left\| \widetilde{W}_{i'j'}(z) \left(G_{i'j'}(z) - \widehat{G}'_{i'j'}(z) \right) \right\|_{\ell_1}.$$

The system uncertainty set $\mathcal{S}$ and weightings $\widetilde{W}_{i'j'}(z)$ are as specified above. The model set $\mathcal{M}$ is linearly parametrized,

$$\mathcal{M}: \ \widehat{G}'(z) = \begin{bmatrix} \widehat{G}'_{11}(z) & \cdots & \widehat{G}'_{1q'}(z) \\ \vdots & \ddots & \vdots \\ \widehat{G}'_{p'1}(z) & \cdots & \widehat{G}'_{p'q'}(z) \end{bmatrix},$$

where each $G'_{i'j'}(z)$ is an element of the set $\mathcal{M}_{i'j'}$ given by

$$\mathcal{M}_{i'j'}: \ \widehat{G}'_{i'j'}(z) = \frac{n_{i'j'}(z)}{d_{i'j'}(z)}, \ n_{i'j'}(z) = \sum_{k=0}^{n_{i'j'}} n_k^{i'j'} z^k,$$

with each $d_{i'j'}(z)$ some fixed and given polynomial of degree $d_{i'j'} \geq n_{i'j'}$, which has all its roots inside the open unit disc.

Again, the problem is approximately solved by adopting a two-step procedure. In the first step the central estimate is computed,

$$\widehat{G}_C^{i'j'}(z) = \arg \min_{\widehat{G}'_C \in \mathcal{A}} \max_{G_{i'j'} \in \mathcal{S}_{i'j'}} \left\| W_{i'j'}(z) \left(G_{i'j'}(z) - \widehat{G}'_C(z) \right) \right\|_{\ell_1},$$

a solution for which is provided in Theorem 7.3.2. In the second step the ℓ_1-optimal approximation of this central estimate is computed as the solution to the optimization problem,

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \max_{i',j'} \left\| \widetilde{W}_{i'j'}(z) \left(\widehat{G}_C^{i'j'}(z) - \widehat{G}'_{i'j'}(z) \right) \right\|_{\ell_1}.$$

Due to the linear model parametrization this MIMO problem can be split up into $p'q'$ independent SISO problems, as there is no interaction between the different entries of the MIMO transfer function,

$$\widehat{G}_{i'j'}(z) = \arg \min_{\widehat{G}'_{i'j'} \in \mathcal{M}_{i'j'}} \left\| \widetilde{W}_{i'j'}(z) \left(\widehat{G}_C^{i'j'}(z) - \widehat{G}'_{i'j'}(z) \right) \right\|_{\ell_1}, \ \forall i', j'.$$

These are just SISO ℓ_1-optimal model reduction problems, a solution for which is provided in Theorem 7.3.3.

7.5 Examples

7.5.1 Example 1: H_∞-Optimal Model Reduction

In this subsection examples of H_∞-optimal model reduction are given. This model reduction problem concerns the computation of a stable model $\widehat{G}(z)$ of some specified order, which optimally approximates a given system $G_0(z)$ in the following sense,

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \left\| \widetilde{W}(z) \left(G_0(z) - \widehat{G}'(z) \right) \right\|_\infty .$$

Untill now this problem has not been solved. Only a characterization of the solution, but even not the solution itself, has been given in Kavranoğlu and Bettayeb (1993) and Kavranoğlu (1993) for $\widetilde{W}(z) = 1$. According to Theorem 7.2.3 a solution for the general problem is provided by the curve fit procedure developed in Chapter 6.

Consider the 5th order discrete-time system

$$G_0(z) = \frac{0.1z^5 - 0.268z^4 + 0.2768z^3 - 0.1365z^2 + 0.0285z}{z^5 - 4.2706z^4 + 7.4176z^3 - 6.5781z^2 + 2.9894z - 0.5582},$$

which is stable and minimum-phase.

First, this system is reduced to 4th order, minimizing the H_∞-norm of the additive error, $\widetilde{W}(z) = 1$. The curve fit procedure of Chapter 6 has been applied to this problem, using a frequency grid Ω of 50 frequencies logarithmically distributed between 0.01 and π. The poles of the model have been restricted to the unit disc, but no constraints on the numerator parameters have been used. The curve fit algorithm yields the following optimal reduced order model,

$$\widehat{G}_1(z) = \frac{0.106239385z^4 - 0.212004596z^3 + 0.186880001z^2 - 0.109776190z + 0.035525358}{z^4 - 3.364803891z^3 + 4.314862622z^2 - 2.512867722z + 0.563669118}.$$

In Figure 7.1 an amplitude Bode diagram is given of the 5th order system, the reduced order model and the additive error. Obviously, the model reduction error is all-pass, which could have been expected. The H_∞-norm of the additive error is equal to 0.01983 and the ℓ_1-norm of the additive error is equal to 0.12136.

Next, the 5th order system is reduced to 3rd order, minimizing the H_∞-norm of the relative error, $\widetilde{W}(z) = G_0(z)^{-1}$. Again, this has been solved by applying the curve fit procedure of Chapter 6, using the same frequency grid as previously and restricting the poles to the unit disc. The following optimal model is obtained,

$$\widehat{G}_2(z) = \frac{0.098925559z^3 + 0.017583191z^2 + 0.023268372z + 0.023644254}{z^3 - 1.422805316z^2 + 0.368179543z + 0.072207994}.$$

In Figure 7.2 an amplitude Bode diagram is given of the system, the reduced order model and the relative error. The H_∞-norm of the relative error is equal to 0.16102. The ℓ_1-norm of this error is equal to 0.72147.

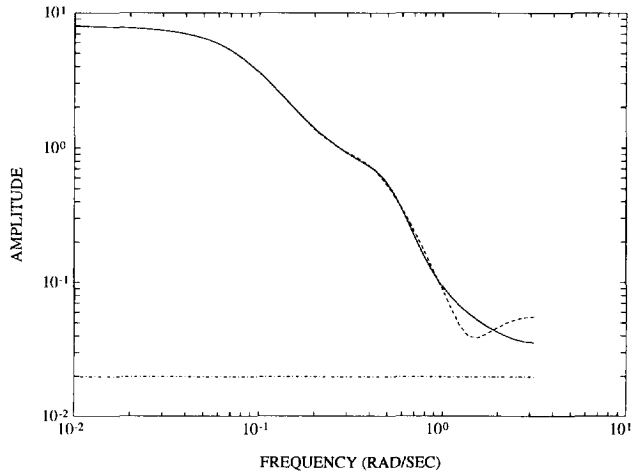

Fig. 7.1: Amplitude Bode diagram of the system $G_0(z)$ (solid), the reduced order model $\widehat{G}_1(z)$ (dashed) and the additive error $G_0(z) - \widehat{G}_1(z)$ (dash-dotted).

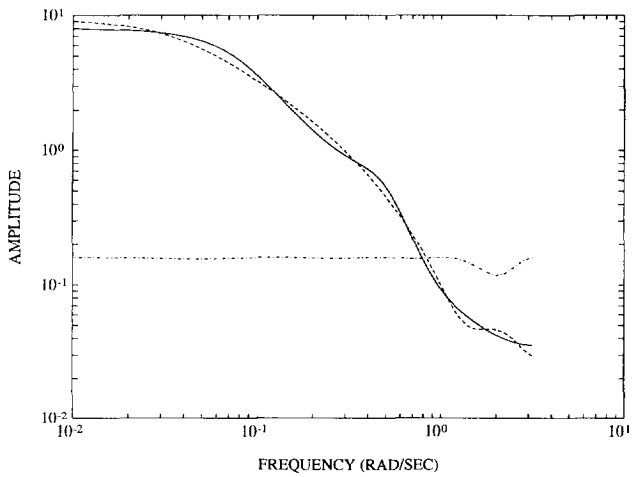

Fig. 7.2: Amplitude Bode diagram of the system $G_0(z)$ (solid), the reduced order model $\widehat{G}_2(z)$ (dashed) and the error $\left(G_0(z) - \widehat{G}_2(z)\right) G_0(z)^{-1}$ (dash-dotted).

7.5.2 Example 2: ℓ_1-Optimal Model Reduction

In this subsection examples of ℓ_1-optimal model reduction are given. This model reduction problem concerns the computation of a stable model $\widehat{G}(z)$ of some specified

order, which optimally approximates a given system $G_0(z)$ in the following sense,

$$\widehat{G}(z) = \arg \min_{\widehat{G}' \in \mathcal{M}} \left\| \widetilde{W}(z) \left(G_0(z) - \widehat{G}'(z) \right) \right\|_{\ell_1} .$$

A solution for this problem has been given in Subsection 7.3.2 in case of a linearly parametrized model set $\mathcal{M}$.

Consider the system $G_0(z)$ as specified in the previous example. The system is reduced to 4th order, minimizing the ℓ_1-norm of the additive error, $\widetilde{W}(z) = 1$. The denominator of the model has been fixed to the denominator of the H_∞-optimal reduced order model $\widehat{G}_1(z)$ as given above. The truncation parameter l in the linear programming problem (7.21) has been chosen 200. The following ℓ_1-optimal reduced order model has been computed,

$$\widehat{G}_3(z) = \frac{0.1z^4 - 0.194687443z^3 + 0.168702440z^2 - 0.100994990z + 0.033831954}{z^4 - 3.364803891z^3 + 4.314862622z^2 - 2.512867722z + 0.563669118}.$$

In Figure 7.3 an amplitude Bode diagram is given of the system, the reduced order model and the additive error. The H_∞-norm of the additive error is equal to 0.04684

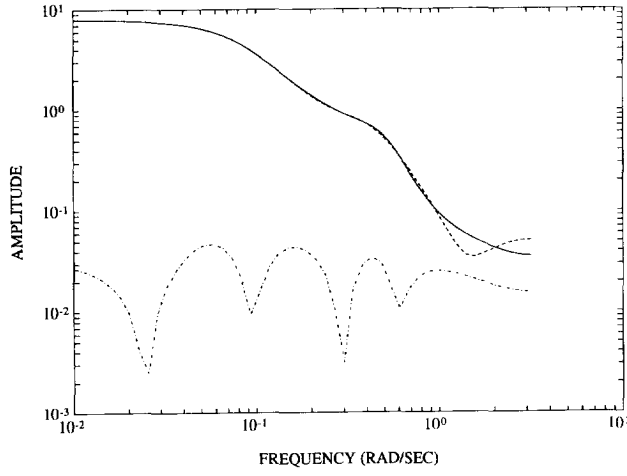

Fig. 7.3: Amplitude Bode diagram of the system $G_0(z)$ (solid), the reduced order model $\widehat{G}_3(z)$ (dashed) and the additive error $G_0(z) - \widehat{G}_3(z)$ (dash-dotted).

and the ℓ_1-norm of the additive error is equal to 0.11302.

Next, the system $G_0(z)$ is reduced to 3rd order, minimizing the ℓ_1-norm of the relative error, $\widetilde{W}(z) = G_0(z)^{-1}$. The denominator of the model has been fixed to the denominator of the H_∞-optimal reduced order model $\widehat{G}_2(z)$ as specified above. Again, the truncation parameter l has been chosen 200. The following ℓ_1-optimal reduced order model has been computed,

$$\widehat{G}_4(z) = \frac{0.1z^3 + 0.016779468z^2 + 0.024828176z + 0.017593997}{z^3 - 1.422805316z^2 + 0.368179543z + 0.072207994}.$$

In Figure 7.4 an amplitude Bode diagram is shown of the system, the reduced order model and the relative error. The H_∞-norm of the relative error is 0.20840 and its

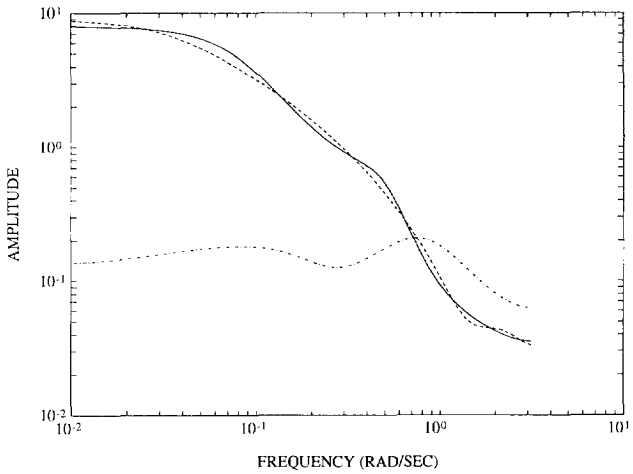

Fig. 7.4: Amplitude Bode diagram of the system $G_0(z)$ (solid), the reduced order model $\widehat{G}_4(z)$ (dashed) and the error $\left(G_0(z) - \widehat{G}_4(z)\right) G_0(z)^{-1}$ (dash-dotted).

ℓ_1-norm is 0.61489.

7.5.3 Example 3: H_∞-Identification of a Nominal Model

In this subsection the H_∞-part of Example 2 in Section 5.9 is continued. There frequency response uncertainty regions have been identified. Here these are used to construct a nominal model with the H_∞-identification procedure of Section 7.2.

First, the centers of the frequency response uncertainty regions depicted in Figure 5.3 are calculated. This yields the central estimate $\widehat{G}_C(\omega_j)$ as defined in (7.8). Next, with the curve fit procedure of Chapter 6 a 4th order model is constructed, minimizing the criterion (7.9) with $\widetilde{W}(z) = 1$. The resulting nominal model, denoted by $\widehat{G}_\infty(z)$, is equal to

$$\widehat{G}_\infty(z) = \frac{0.08062z^4 - 0.15993z^3 + 1.18131z^2 - 1.51746z + 0.79717}{z^4 - 2.67311z^3 + 3.29946z^2 - 2.01906z + 0.54769}.$$

In Figure 7.5 an amplitude Bode diagram is shown of the central estimate, the optimal nominal model and the fit error.

Similar to the procedure in Subsection 5.9.2 the frequency response uncertainty regions can be used to establish an upper bound on the H_∞-norm of the model error. On the basis of the rectangular confidence regions depicted in Figure 5.3, the additive model error bound $\delta_u(\omega_j)$ is calculated with respect to the model $\widehat{G}_\infty(z)$. The result

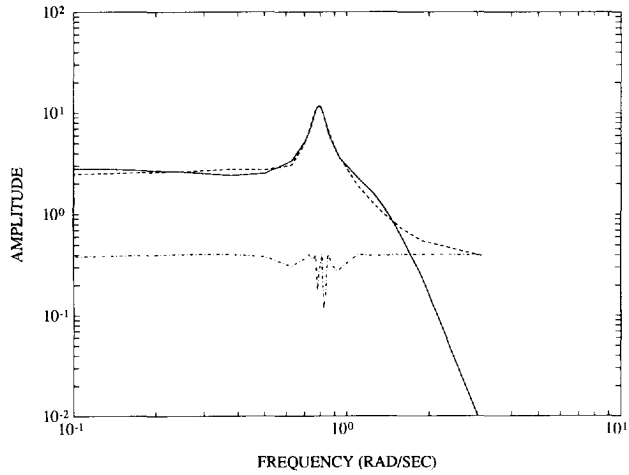

Fig. 7.5: Amplitude Bode diagram of the central estimate $\widehat{G}_C(\omega_j)$ (solid), the model $\widehat{G}_\infty(z)$ (dashed) and the fit error $\left|\widehat{G}_C(\omega_j) - \widehat{G}_\infty(e^{i\omega_j})\right|$ (dash-dotted).

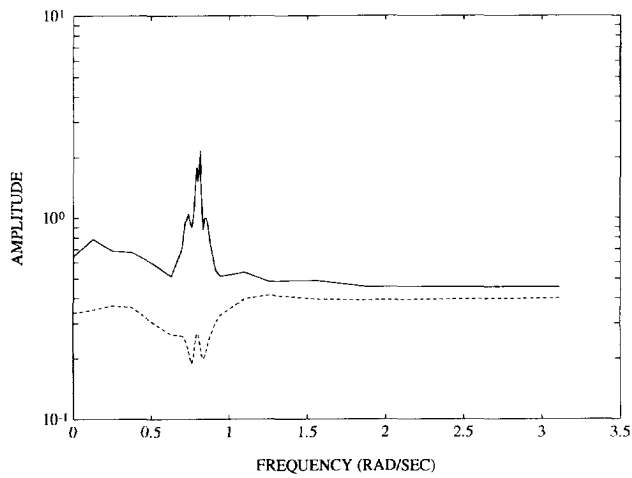

Fig. 7.6: Additive frequency response model error bound $\delta_u(\omega_j)$ (solid) and additive model error $\left|\widehat{G}_\infty(e^{i\omega_j}) - G_0(e^{i\omega_j})\right|$ (dashed).

is depicted in Figure 7.6. Neglecting the influence of the intersample frequencies, this yields an H_∞-model error bound equal to 2.16, whereas the H_∞-norm of the model

error is 0.414. Obviously, the model $\widehat{G}_\infty(z)$ is a lot better than the model $\widehat{G}(z)$ used in Subsection 5.9.2. An even better model can be obtained if the model order is increased.

7.5.4 Example 4: ℓ_1-Identification of a Nominal Model

In this subsection the ℓ_1-part of Example 2 in Section 5.9 is continued. There pulse response uncertainty regions have been identified. Here these are used to construct a nominal model with the ℓ_1-identification procedure described in Section 7.3.

First, the centers of the pulse response confidence intervals depicted in Figure 5.4 are calculated. This yields the central estimate $\widehat{G}_C(z)$ as defined in (7.18). Next, with the ℓ_1-model reduction procedure described in Subsection 7.3.2 a 4th order model is constructed, minimizing the criterion (7.19) with $\widetilde{W}(z) = 1$. The denominator is chosen equal to the denominator of the model $\widehat{G}_\infty$ as determined in the previous example. The resulting nominal model, denoted by $\widehat{G}_{\ell_1}(z)$, is equal to

$$\widehat{G}_{\ell_1}(z) = \frac{0.006673z^4 + 0.20250z^3 + 0.60150z^2 - 1.06513z + 0.65704}{z^4 - 2.67311z^3 + 3.29946z^2 - 2.01906z + 0.54769}.$$

In Figure 7.7 a diagram is shown with the pulse response of the central estimate, denoted by $p_c(k)$, the pulse response of the optimal nominal model, denoted by $\widehat{p}_{\ell_1}(k)$, and the pulse response error.

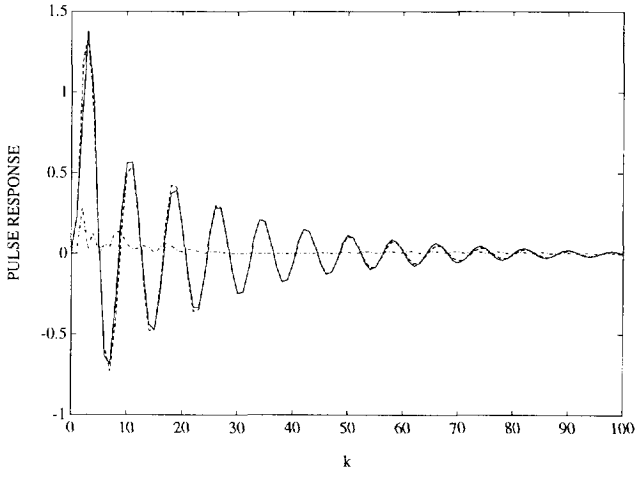

Fig. 7.7: Pulse responses of the central estimate $\widehat{G}_C(z)$ (solid) and the model $\widehat{G}_{\ell_1}(z)$ (dashed) and the pulse response error $|p_c(k) - \widehat{p}_{\ell_1}(k)|$ (dash-dotted).

Similar to the procedure in Subsection 5.9.2 the pulse response uncertainty regions can be used to establish an upper bound on the ℓ_1-norm of the model error. Figure 7.8 depicts the pulse response error bound. This yields an ℓ_1-model error bound equal to

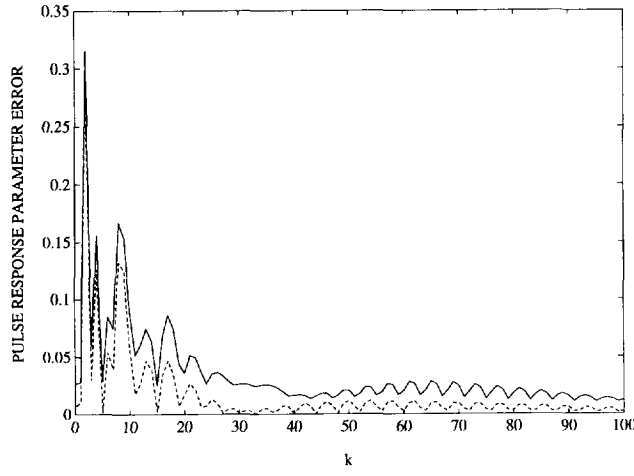

Fig. 7.8: Additive pulse response model error bound $\max[p_u(k) - \widehat{p}_{\ell_1}(k), \widehat{p}_{\ell_1}(k) - p_l(k)]$ (solid) and additive model error $|\widehat{p}_{\ell_1}(k) - p_0(k)|$ (dashed).

3.6, whereas the ℓ_1-norm of the model error is 1.73. Also, the model $\widehat{G}_{\ell_1}(z)$ appears to be a lot better than the model $\widehat{G}(z)$ used in Subsection 5.9.2.

7.6 Discussion

In this chapter the problem has been considered of constructing a nominal model of some user-defined order, such that a certain worst-case system-norm is minimized. For that purpose the problems of system approximation in H_∞- and ℓ_1-norm have been formulated and solved in good approximation. In the H_∞-case the solution is based on the curve fit algorithm developed in Chapter 6. In the ℓ_1-case the solution is based on an algorithm for model reduction in ℓ_1-norm, which has been developed in this chapter. MIMO extensions have been presented as well.

The procedures developed in this chapter are applicable to the problem of identifying a model that is good in H_∞- or ℓ_1-sense. Notice that in both the H_∞- and the ℓ_1-case the nominal model is not constructed directly on the basis of measured data. In both cases the starting point is the system uncertainty set as identified in an uncertainty bounding identification procedure, such as the ones developed in Chapter 4 and 5. These system uncertainty sets specify uncertainty regions for the system's frequency or pulse response. The problem of constructing a nominal model as considered in this chapter, is the problem of constructing a suitable parametric approximation to these uncertainty sets. Obviously, the resulting nominal model is indirectly determined by the measurement data, as the system uncertainty sets are identified on the basis of data.

Examples 1 and 2 in Section 7.5 show that the tools developed in this and the previous chapter have possible applications outside the area of identification. In particular these tools provide exact solutions for H_∞- and ℓ_1-optimal model reduction problems.

7.A Proofs

Proof of Proposition 7.2.1

The second inequality in part (i) follows from the definition of $\widehat{G}_{\mathrm{C}}$. Next, it is noticed that

$$\mathrm{Re}\left(e^{i2\pi k'/m'}v_k\right) - \cos(2\pi k'/m')a + \sin(2\pi k'/m')b = \mathrm{Re}\left(e^{i2\pi k'/m'}\left(v_k - a - bi\right)\right),$$

and, consequently, the solution of the LP problem has the property

$$\mu_{\mathrm{opt}} = \max_k f_{m'}\left(v_k - a_{\mathrm{opt}} - b_{\mathrm{opt}}i\right),$$

with the function $f_{m'}(\cdot)$ defined in Lemma 3.4.1. Application of this lemma yields the third inequality in part (i). Finally, optimality of a_{opt}, b_{opt} implies that for some k^*,

$$f_{m'}\left(v_{k^*} - \widehat{G}_{\mathrm{C}}\right) \geq \mu_{\mathrm{opt}},$$

which again with Lemma 3.4.1 yields the first inequality of part (i).

For $m' \to \infty$ the right-hand side in (i) converges to the left-hand side and hence, as $\widehat{G}_{\mathrm{C}}$ is a unique complex number, $a_{\mathrm{opt}} + b_{\mathrm{opt}}i$ converges to $\widehat{G}_{\mathrm{C}}$, which proves part (ii).

Proof of Lemma 7.2.2

First, it is noted that any model $\widehat{G}'(z) \in \mathcal{M}$ is stable, due to the bounds on the denominator roots. As the poles of $\widehat{G}'(z)$ are identical to those of $\frac{\mathrm{d}\,\widehat{G}'(z)}{\mathrm{d}\,z}$ (up to multiplicity), this derivative is also stable.

Next, the following bounds can be established for the numerator $n(z)$ and denominator $d(z)$ of any model in the model set $\mathcal{M}$,

$$\sup_{|z|=1}|n(z)| = \sup_{|z|=1}\left|\sum_{k=0}^{n} n_k z^k\right| \leq \sup_{|z|=1}\sum_{k=0}^{n}|n_k||z^k| = \sum_{k=0}^{n}|n_k| \leq (n+1)\bar{n},$$

$$\begin{aligned}\inf_{|z|=1}|d(z)| &= \inf_{|z|=1}\left|z^d + \sum_{k=0}^{d-1} d_k z^k\right| = \inf_{|z|=1}\left|\prod_{k=1}^{d}(z - z_{\mathrm{d},k})\right| = \inf_{|z|=1}\prod_{k=1}^{d}|z - z_{\mathrm{d},k}| \geq \\ &\geq \inf_{|z|=1}\prod_{k=1}^{d}(|z| - |z_{\mathrm{d},k}|) \geq \prod_{k=1}^{d}(1-\rho) = (1-\rho)^d,\end{aligned}$$

$$\sup_{|z|=1}\left|\frac{\mathrm{d}\,n(z)}{\mathrm{d}\,z}\right| = \sup_{|z|=1}\left|\sum_{k=0}^{n} k n_k z^{k-1}\right| \le \sum_{k=0}^{n} k|n_k| \le \bar{n}\sum_{k=0}^{n} k = \frac{\bar{n}n(n+1)}{2},$$

$$\sup_{|z|=1}\left|\frac{\mathrm{d}\,d(z)}{\mathrm{d}\,z}\right| = \sup_{|z|=1}\left|\frac{\mathrm{d}\prod_{k=1}^{d}(z-z_{\mathrm{d},k})}{\mathrm{d}\,z}\right| = \sup_{|z|=1}\left|\sum_{k'=1}^{d}\prod_{k\neq k'}(z-z_{\mathrm{d},k})\right| \le$$

$$\le \sup_{|z|=1}\sum_{k'=1}^{d}\prod_{k\neq k'}(|z|+|z_{\mathrm{d},k}|) \le d(1+\rho)^{d-1}.$$

These results lead to the following bounds,

$$\left\|\widehat{G}'(z)\right\|_\infty = \sup_{|z|=1}\left|\frac{n(z)}{d(z)}\right| \le \frac{\sup_{|z|=1}|n(z)|}{\inf_{|z|=1}|d(z)|} \le \frac{(n+1)\bar{n}}{(1-\rho)^d},$$

which proves the first part, and

$$\left\|\frac{\mathrm{d}\,\widehat{G}'(z)}{\mathrm{d}\,z}\right\|_\infty = \left\|\frac{\mathrm{d}\,\frac{n(z)}{d(z)}}{\mathrm{d}\,z}\right\|_\infty = \left\|d^{-2}(z)\left(d(z)\frac{\mathrm{d}\,n(z)}{\mathrm{d}\,z} - n(z)\frac{\mathrm{d}\,d(z)}{\mathrm{d}\,z}\right)\right\|_\infty =$$

$$= \sup_{|z|=1}\left|d^{-2}(z)\left(d(z)\frac{\mathrm{d}\,n(z)}{\mathrm{d}\,z} - n(z)\frac{\mathrm{d}\,d(z)}{\mathrm{d}\,z}\right)\right| \le$$

$$\le \sup_{|z|=1}\left|d^{-1}(z)\frac{\mathrm{d}\,n(z)}{\mathrm{d}\,z}\right| + \left|d^{-2}(z)n(z)\frac{\mathrm{d}\,d(z)}{\mathrm{d}\,z}\right| \le$$

$$\le \frac{\sup_{|z|=1}\left|\frac{\mathrm{d}\,n(z)}{\mathrm{d}\,z}\right|}{\inf_{|z|=1}|d(z)|} + \frac{\sup_{|z|=1}|n(z)|\sup_{|z|=1}\left|\frac{\mathrm{d}\,d(z)}{\mathrm{d}\,z}\right|}{\inf_{|z|=1}|d(z)|^2} \le$$

$$\le \frac{\bar{n}n(n+1)}{2(1-\rho)^d} + \frac{\bar{n}(n+1)d(1+\rho)^{d-1}}{(1-\rho)^{2d}},$$

which proves the second part.

Proof of Theorem 7.2.3

The fact is used that for any ℓ_2-stable $X(z)$,

$$\left|\frac{\mathrm{d}\,|X(e^{i\omega})|}{\mathrm{d}\,\omega}\right| \le \left|\frac{\mathrm{d}\,X(e^{i\omega})}{\mathrm{d}\,\omega}\right|,$$

and that

$$\sup_{\omega\in[0,\pi]}\left|\frac{\mathrm{d}\,X(e^{i\omega})}{\mathrm{d}\,\omega}\right| = \left\|\frac{\mathrm{d}\,X(z)}{\mathrm{d}\,z}\right\|_\infty.$$

Next, for any $\omega \in [0,\pi]$ and any $\widehat{G}'(z) \in \mathcal{M}$, the following bound can be established,

$$\begin{aligned}
&\left|\frac{\mathrm{d}\left|\widetilde{W}(e^{i\omega})\left(\widehat{G}_{\mathrm{C}}(\omega)-\widehat{G}'(e^{i\omega})\right)\right|}{\mathrm{d}\,\omega}\right| \leq \\
&\leq \left|\frac{\mathrm{d}\,\widetilde{W}(e^{i\omega})\widehat{G}_{\mathrm{C}}(\omega)}{\mathrm{d}\,\omega}\right| + \left|\frac{\mathrm{d}\,\widetilde{W}(e^{i\omega})\widehat{G}'(e^{i\omega})}{\mathrm{d}\,\omega}\right| \leq \\
&\leq \left|\frac{\mathrm{d}\,\widetilde{W}(e^{i\omega})}{\mathrm{d}\,\omega}\right|\left|\widehat{G}_{\mathrm{C}}(\omega)\right| + \left|\widetilde{W}(e^{i\omega})\right|\left|\frac{\mathrm{d}\,\widehat{G}_{\mathrm{C}}(\omega)}{\mathrm{d}\,\omega}\right| + \\
&\quad + \left|\frac{\mathrm{d}\,\widetilde{W}(e^{i\omega})}{\mathrm{d}\,\omega}\right|\left|\widehat{G}'(e^{i\omega})\right| + \left|\widetilde{W}(e^{i\omega})\right|\left|\frac{\mathrm{d}\,\widehat{G}'(e^{i\omega}))}{\mathrm{d}\,\omega}\right| \leq \\
&\leq \beta_6\beta_2 + \beta_5\beta_1 + \beta_6\beta_3(n,d,\rho,\bar{n}) + \beta_5\beta_4(n,d,\rho,\bar{n}) = \beta(n,d,\rho,\bar{n}).
\end{aligned}$$

Thus, $\forall\omega$, ω_j, linear interpolation gives the following bound,

$$\left|\widetilde{W}(e^{i\omega})\left(\widehat{G}_{\mathrm{C}}(\omega)-\widehat{G}'(e^{i\omega})\right)\right| \leq \left|\widetilde{W}(e^{i\omega_j})\left(\widehat{G}_{\mathrm{C}}(\omega_j)-\widehat{G}'(e^{i\omega_j})\right)\right| + |\omega_j-\omega|\beta(n,d,\rho,\bar{n}).$$

Performing this interpolation for every $\omega \in [\omega_j - \lambda_j/2, \omega_j + \lambda_j/2]$ and for every j gives the desired result (i).

As result (i) holds for any $\widehat{G}'(z) \in \mathcal{M}$, it holds for $\widehat{G}_\Omega(z)$ defined in (7.9). Using this and optimality of $\widehat{G}(z)$ and $\widehat{G}_\Omega(z)$, according to (7.7) and (7.9), the following result follows,

$$\begin{aligned}
&\max_{j=1,\ldots,l}\left|\widetilde{W}(e^{i\omega_j})\left(\widehat{G}_{\mathrm{C}}(\omega_j)-\widehat{G}_\Omega(e^{i\omega_j})\right)\right| \leq \max_{j=1,\ldots,l}\left|\widetilde{W}(e^{i\omega_j})\left(\widehat{G}_{\mathrm{C}}(\omega_j)-\widehat{G}(e^{i\omega_j})\right)\right| \leq \\
&\leq \max_{\omega\in[0,\pi]}\left|\widetilde{W}(e^{i\omega})\left(\widehat{G}_{\mathrm{C}}(\omega)-\widehat{G}(e^{i\omega})\right)\right| \leq \max_{\omega\in[0,\pi]}\left|\widetilde{W}(e^{i\omega})\left(\widehat{G}_{\mathrm{C}}(\omega)-\widehat{G}_\Omega(e^{i\omega})\right)\right| \leq \\
&\leq \max_{j=1,\ldots,l}\left[\left|\widetilde{W}(e^{i\omega_j})\left(\widehat{G}_{\mathrm{C}}(\omega_j)-\widehat{G}_\Omega(e^{i\omega_j})\right)\right| + \frac{1}{2}\lambda_j\beta(n,d,\rho,\bar{n})\right].
\end{aligned}$$

From the first, third and fifth terms in this expression it follows that minimization (7.7) is sandwiched between minimization (7.9) and minimization (7.9) plus some error term, the interpolation term $\lambda_j\beta(n,d,\rho,\bar{n})/2$. This error term vanishes if

$$\left(\max_j \lambda_j\right)\beta(n,d,\rho,\bar{n}) \to 0.$$

In that case both minimization problems (7.9) and (7.7) turn out to be equivalent, which proves part (ii).

Proof of Theorem 7.3.2

First, it is noted that $p_{\mathrm{c}}(k) = 0$ for k large enough (larger than $\bar{k}-1$), and therefore $P_{\mathrm{C}}(z)$ is finite dimensional and stable. Moreover, $W^{-1}(z)$ is finite dimensional, causal

and stable. Therefore $W^{-1}(z)P_C(z)$ is finite dimensional and stable and, consequently, an element of the set $\mathcal{A}$.

Consider the set

$$\mathcal{T} := \left\{ G(z) \in \mathcal{A} \;\middle|\; G(z) = \sum_{k=0}^{\infty} p(k) z^{-k},\ p(k) \in [p_l(k), p_u(k)],\ k = 0, \ldots, \infty \right\},$$

which is a symmetric orthotope which has $P_C(z)$ as center. This set $\mathcal{T}$ is related to the set $\mathcal{S}$ in the sense that any $G(z)$ is an element of $\mathcal{S}$ if and only if $W(z)G(z)$ is an element of $\mathcal{T}$. Using the orthotopic structure of the set $\mathcal{T}$ the following result is obtained for any model $\widehat{G}'_C(z) \in \mathcal{A}$,

$$\begin{aligned}
&\max_{G \in \mathcal{S}} \left\| W(z) \left(G(z) - \widehat{G}'_C(z) \right) \right\|_{\ell_1} = \\
&= \max_{G \in \mathcal{T}} \left\| G(z) - P_C(z) + P_C(z) - W(z)\widehat{G}'_C(z) \right\|_{\ell_1} = \\
&= \max_{G \in \mathcal{T}} \| G(z) - P_C(z) \|_{\ell_1} + \left\| P_C(z) - W(z)\widehat{G}'_C(z) \right\|_{\ell_1} = \\
&= \max_{G \in \mathcal{S}} \| W(z)G(z) - P_C(z) \|_{\ell_1} + \left\| P_C(z) - W(z)\widehat{G}'_C(z) \right\|_{\ell_1} \geq \\
&\geq \max_{G \in \mathcal{S}} \| W(z)G(z) - P_C(z) \|_{\ell_1} = \sum_{k=0}^{\infty} \frac{1}{2} \left(p_u(k) - p_l(k) \right),
\end{aligned}$$

and equality is achieved for $\widehat{G}'_C(z) = W^{-1}(z)P_C(z)$. This completes the proof.

Moreover, the derivation shows that the following equality holds,

$$\begin{aligned}
&\max_{G \in \mathcal{S}} \left\| W(z) \left(G(z) - \widehat{G}'(z) \right) \right\|_{\ell_1} = \\
&= \max_{G \in \mathcal{S}} \left\| W(z) \left(G(z) - \widehat{G}_C(z) \right) \right\|_{\ell_1} + \left\| W(z) \left(\widehat{G}_C(z) - \widehat{G}'(z) \right) \right\|_{\ell_1},
\end{aligned}$$

for any $\widehat{G}'(z) \in \mathcal{A}$. This implies that minimization of the left-hand side by a suitable choice of $\widehat{G}'(z)$ is equivalent to minimization of the right-hand side, which in turn is equivalent to minimization of the second term at the right-hand side by a suitable choice of $\widehat{G}'(z)$. This proves Theorem 7.3.1.

Proof of Theorem 7.3.3

First, it is noted that the set of constraints in (7.21) is a subset of the set of constraints in (7.20). Consequently, due to the fact that

$$\left\{ \{n_k^{\text{opt}}\}_{k=0,\ldots,n},\ \{\delta_p^{\text{opt}}(k)\}_{k=0,\ldots,\infty},\ \{\delta_n^{\text{opt}}(k)\}_{k=0,\ldots,\infty} \right\}$$

is a feasible solution of the LP problem (7.20),

$$\left\{ \{n_k^{\text{opt}}\}_{k=0,\ldots,n},\ \{\delta_p^{\text{opt}}(k)\}_{k=0,\ldots,l},\ \{\delta_n^{\text{opt}}(k)\}_{k=0,\ldots,l} \right\}$$

is a feasible solution of the LP problem (7.21). Together with the fact that

$$\left\{\{n_k^{\mathrm{T,opt}}\}_{k=0,\ldots,n},\ \{\delta_{\mathrm{p}}^{\mathrm{T,opt}}(k)\}_{k=0,\ldots,l},\ \{\delta_{\mathrm{n}}^{\mathrm{T,opt}}(k)\}_{k=0,\ldots,l}\right\}$$

is the optimal solution of (7.21) this gives

$$\begin{aligned}\|\Delta_{\mathrm{T,opt}}(z)\|_{\ell_1} &= \sum_{k=0}^{l}\left(\delta_{\mathrm{p}}^{\mathrm{T,opt}}(k)+\delta_{\mathrm{n}}^{\mathrm{T,opt}}(k)\right) \le \sum_{k=0}^{l}\left(\delta_{\mathrm{p}}^{\mathrm{opt}}(k)+\delta_{\mathrm{n}}^{\mathrm{opt}}(k)\right) \le \\ &\le \sum_{k=0}^{\infty}\left(\delta_{\mathrm{p}}^{\mathrm{opt}}(k)+\delta_{\mathrm{n}}^{\mathrm{opt}}(k)\right) = \|\Delta_{\mathrm{opt}}(z)\|_{\ell_1},\end{aligned}$$

which proves the first inequality in part (i). The second inequality in part (i) is a direct implication of the fact that

$$\left\{\{n_k^{\mathrm{opt}}\}_{k=0,\ldots,n},\ \{\delta_{\mathrm{p}}^{\mathrm{opt}}(k)\}_{k=0,\ldots,\infty},\ \{\delta_{\mathrm{n}}^{\mathrm{opt}}(k)\}_{k=0,\ldots,\infty}\right\}$$

is the optimal solution of (7.20), which completes the proof of part (i).

Next, denote

$$\Delta_{\mathrm{T}}(z)=\sum_{k=0}^{\infty}\delta_{\mathrm{T}}(k)z^{-k} := \widetilde{W}(z)\left(\widehat{G}_{\mathrm{C}}(z)-\widehat{G}_{\mathrm{T,opt}}(z)\right).$$

Note that the solution of LP problem (7.20) has the property that

$$\Delta_{\mathrm{opt}}(z)=\widetilde{W}(z)\left(\widehat{G}_{\mathrm{C}}(z)-\widehat{G}_{\mathrm{opt}}(z)\right),$$

but with respect to LP problem (7.21) in general,

$$\Delta_{\mathrm{T,opt}}(z)\neq\widetilde{W}(z)\left(\widehat{G}_{\mathrm{C}}(z)-\widehat{G}_{\mathrm{T,opt}}(z)\right)=\Delta_{\mathrm{T}}(z).$$

However, it is shown that the difference between the left- and the right-hand side term vanishes if $l\to\infty$. Consider LP problem (7.21). If $l=l_1$ there is a set of constraints for each $\delta_{\mathrm{p}}(k)$ and $\delta_{\mathrm{n}}(k)$, $\forall\, 0\le k\le l_1$. If l is increased to some integer l_2, a number of constraints is added. However, due to the fact that $a(z^{-1})$ is a polynomial of finite degree, no new constraints are added for each $\delta_{\mathrm{p}}(k)$ and $\delta_{\mathrm{n}}(k)$ with $k\le l_1-\deg(a(z^{-1}))$. Denoting $\bar{l}:=l-\deg(a(z^{-1}))$, this implies that the first $\bar{l}+1$ pulse response parameters of $\Delta_{\mathrm{T}}(z)$ are exactly identical to the first $\bar{l}+1$ pulse response parameters of $\Delta_{\mathrm{T,opt}}(z)$, i.e.

$$\delta_{\mathrm{T}}(k)=\delta_{\mathrm{p}}^{\mathrm{T,opt}}(k)-\delta_{\mathrm{n}}^{\mathrm{T,opt}}(k),\ k=0,\ldots,\bar{l}.$$

Split the pulse response of $\Delta_{\mathrm{T}}(z)$ into two parts, the start and the tail,

$$\Delta_{\mathrm{T}}(z)=\widetilde{\Delta}_{\mathrm{T}}(z)+\bar{\Delta}_{\mathrm{T}}(z),\ \widetilde{\Delta}_{\mathrm{T}}(z)=\sum_{k=0}^{\bar{l}}\delta_{\mathrm{T}}(k)z^{-k},\ \bar{\Delta}_{\mathrm{T}}(z)=\sum_{k=\bar{l}+1}^{\infty}\delta_{\mathrm{T}}(k)z^{-k}.$$

Similarly, split the pulse response of $\Delta_{\mathrm{T,opt}}(z)$ into two parts,

$$\Delta_{\mathrm{T,opt}}(z) = \widetilde{\Delta}_{\mathrm{T,opt}}(z) + \bar{\Delta}_{\mathrm{T,opt}}(z),\ \widetilde{\Delta}_{\mathrm{T,opt}}(z) = \sum_{k=0}^{\bar{l}} \left(\delta_{\mathrm{p}}^{\mathrm{T,opt}}(k) - \delta_{\mathrm{n}}^{\mathrm{T,opt}}(k)\right) z^{-k},$$

$$\bar{\Delta}_{\mathrm{T,opt}}(z) = \sum_{k=\bar{l}+1}^{l} \left(\delta_{\mathrm{p}}^{\mathrm{T,opt}}(k) - \delta_{\mathrm{n}}^{\mathrm{T,opt}}(k)\right) z^{-k}.$$

As shown above, the following equality holds,

$$\widetilde{\Delta}_{\mathrm{T}}(z) = \widetilde{\Delta}_{\mathrm{T,opt}}(z),$$

and, consequently, obviously $\left\|\widetilde{\Delta}_{\mathrm{T}}(z)\right\|_{\ell_1} = \left\|\widetilde{\Delta}_{\mathrm{T,opt}}(z)\right\|_{\ell_1}$. Next, it is noted that $\widetilde{W}(z)$, $\widehat{G}_{\mathrm{C}}(z)$ and $\widehat{G}_{\mathrm{T,opt}}(z)$ are finite dimensional and stable and, consequently, $\Delta_{\mathrm{T}}(z)$ is finite dimensional and stable. Therefore, the pulse response of $\Delta_{\mathrm{T}}(z)$ exponentially decays. This implies that the ℓ_1-norm of $\bar{\Delta}_{\mathrm{T}}(z)$ tends to 0 if l, and hence $\bar{l}$, tends to infinity. Using the definition of the ℓ_1-norm and the bounds established in part (i), the following result is obtained,

$$\begin{aligned}
0 \le \|\Delta_{\mathrm{T}}(z)\|_{\ell_1} - \|\Delta_{\mathrm{T,opt}}(z)\|_{\ell_1} &= \\
= \left\|\widetilde{\Delta}_{\mathrm{T}}(z)\right\|_{\ell_1} + \left\|\bar{\Delta}_{\mathrm{T}}(z)\right\|_{\ell_1} - \left\|\widetilde{\Delta}_{\mathrm{T,opt}}(z)\right\|_{\ell_1} - \left\|\bar{\Delta}_{\mathrm{T,opt}}(z)\right\|_{\ell_1} &= \\
= \left\|\bar{\Delta}_{\mathrm{T}}(z)\right\|_{\ell_1} - \left\|\bar{\Delta}_{\mathrm{T,opt}}(z)\right\|_{\ell_1} \le \left\|\bar{\Delta}_{\mathrm{T}}(z)\right\|_{\ell_1} &\to 0 \text{ if } l \to \infty.
\end{aligned}$$

Substitution of the definition of $\Delta_{\mathrm{T}}(z)$ gives,

$$\lim_{l\to\infty} \left(\left\| \widetilde{W}(z) \left(\widehat{G}_{\mathrm{C}}(z) - \widehat{G}_{\mathrm{T,opt}}(z) \right) \right\|_{\ell_1} - \|\Delta_{\mathrm{T,opt}}(z)\|_{\ell_1} \right) = 0.$$

With the bounds established in part (i) this implies that

$$\lim_{l\to\infty} \|\Delta_{\mathrm{T}}(z)\|_{\ell_1} = \|\Delta_{\mathrm{opt}}(z)\|_{\ell_1},$$

and hence, if the solution to the LP problem (7.20) is unique,

$$\lim_{l\to\infty} \|\Delta_{\mathrm{T}}(z) - \Delta_{\mathrm{opt}}(z)\|_{\ell_1} = 0.$$

Noting that

$$\begin{aligned}
\Delta_{\mathrm{T}}(z) - \Delta_{\mathrm{opt}}(z) &= \widetilde{W}(z) \left(\widehat{G}_{\mathrm{C}}(z) - \widehat{G}_{\mathrm{T,opt}}(z) \right) - \widetilde{W}(z) \left(\widehat{G}_{\mathrm{C}}(z) - \widehat{G}_{\mathrm{opt}}(z) \right) = \\
&= \widetilde{W}(z) \left(\widehat{G}_{\mathrm{opt}}(z) - \widehat{G}_{\mathrm{T,opt}}(z) \right),
\end{aligned}$$

finally yields

$$\lim_{l\to\infty} \left\| \widetilde{W}(z) \left(\widehat{G}_{\mathrm{opt}}(z) - \widehat{G}_{\mathrm{T,opt}}(z) \right) \right\|_{\ell_1} = 0,$$

which completes the proof of part (ii).

Part III

Practical Application

Chapter 8

A Robust-Control-Oriented Identification Procedure

8.1 Introduction

The main purpose of this thesis is the development of a practically applicable identification procedure which yields a nominal model and a model error bound suitable for use in high-performance robust control. For that purpose in Chapter 2 a new identification procedure has been introduced, see Figure 2.4. Several steps in this identification procedure need a more detailed specification. In particular this concerns the procedures for actually constructing a system uncertainty set and a nominal model.

In the chapters 3 till 7 new tools have been developed to identify a system uncertainty set, a nominal model and model error bounds. In this chapter it is explained how these tools can be incorporated in the identification procedure of Figure 2.4, such that a practically applicable control-oriented identification procedure is obtained. Application of this procedure then yields both an accurate, control-relevant nominal model and a non-conservative model error bound. These can straightforwardly be used in procedures for the design of a high-performing robust controller.

The outline of the chapter is as follows. In Section 8.2 the details of the identification procedure are outlined. Next, in Section 8.3 a solution strategy is presented in case the system to be identified is unstable. In Section 8.4 the identification procedure is illustrated by means of two examples. The chapter concludes with a discussion of the results in Section 8.5.

8.2 The Identification Procedure

Consider the identification procedure depicted in Figure 2.4 in Chapter 2. This scheme serves as basis for the practical, control-oriented identification procedure. Based on the theory developed in this thesis there are two possible approaches, the H_∞- and the ℓ_1-approach. In the first approach frequency response system uncertainty regions are identified. Based on these uncertainty regions a nominal model can be constructed by

minimizing an H_∞-error criterion. Moreover, the system uncertainty regions can be used to determine an upper bound on the H_∞-norm of the model error, the deviation between system and identified nominal model. In the second approach (weighted) pulse response system uncertainty regions are identified. These uncertainty regions can be used to construct a nominal model by minimizing an ℓ_1-error criterion. And these regions can also be used to construct an upper bound on the ℓ_1-norm of the model error.

These two approaches are not equally powerful. In fact the H_∞-approach is to be preferred in general. The reasons for this are three-fold:

- As follows from the discussion in Subsection 2.2.3 in Chapter 2, an ℓ_1-robustness criterion is conservative in case linear uncertainty blocks are considered. By this is meant that the conditions for robust stability and/or robust performance are sufficient but not necessary in case of a linear system. On the other hand, the H_∞- or μ-robustness tests are necessary and sufficient in case of linear perturbations. In fact easily the situation can occur that robust stability can not be proven with the ℓ_1-test of Theorem 2.2.2, whereas it can be proven with the μ-test of Theorem 2.2.1.

  In this thesis attention is mainly restricted to linear systems, or systems which are approximately linear in an operating point. The uncertainty bounding identification procedures of Chapter 4 and 5 have been developed for linear systems. For this class of systems the H_∞-approach is more natural than the ℓ_1-approach.

- As mentioned in Remark 4.5.5 in Chapter 4 only a conservative upper bound on the ℓ_1-norm of the model error can be identified with the procedure of Chapter 4. On the other hand, the upper bound on the H_∞-norm of the model error can be identified non-conservatively. Something similar can be said of the procedure developed in Chapter 5. This implies additional conservatism for ℓ_1-robustness tests compared to H_∞-robustness tests in case identified model error bounds are used.

- In the H_∞-approach it is easier to include and interpret frequency dependent weightings. As discussed in Section 2.4 in Chapter 2, it is important to choose the design variables in the identification procedure such that the identified model and model error bound are optimally tuned towards the application. In particular this concerns a proper choice of input signal spectrum, the weighting in the nominal identification criterion and the weighting in the criterion for the selection of an uncertainty structure. It appears that the influence of weightings is less transparent in the ℓ_1-case than in the H_∞-case. This is due to the fact that in the H_∞-case a frequency domain interpretation can be given, which is not possible in the ℓ_1-case.

  The restrictions of the ℓ_1-approach especially become clear by the fact that a bound on the ℓ_1-norm of the weighted model error can only be calculated for an *a priori* specified weighting. In the H_∞-approach this weighting can be determined *a posteriori*, when the distribution of the model error over frequency has become

visible. Then the H_∞-norm of the model error can be determined with respect to a weighting which is optimally tuned towards the shape of the model error. See the discussion about uncertainty structure selection in Subsection 2.4.6.

For these three reasons the H_∞-approach is recommended for use in the control-oriented identification procedure. However, note that sometimes the use of the ℓ_1-robustness test is unavoidable, even if the system is linear. This is the case if robust performance of a controller is evaluated, with respect to certain time-domain specifications in terms of induced ℓ_∞-norms. This then also requires the specification of a bound on the model error in terms of the ℓ_∞-induced norm. Currently there is no test available which can handle time-domain performance specifications in combination with a frequency domain system uncertainty description.

Within the H_∞-approach there are two alternatives with respect to the construction of system uncertainty regions. These regions can be identified with the procedure for deterministic identification, presented in Chapter 4, or the procedure for probabilistic identification, presented in Chapter 5. Again, these two alternatives are not equally powerful, but the probabilistic approach performs better than the deterministic approach. The reasons for this are two-fold:

- In general the identified probabilistic system uncertainty regions are much smaller than the deterministic ones. This is not only true for the simulation examples considered in Chapter 4 and 5, but to the author's experience this is typical behaviour. This conservatism of a deterministic approach is well-known in case time-domain amplitude noise constraints are used as prior information. For that reason the cross-covariance noise constraints have been introduced in Chapter 3. But even the use of these constraints in Chapter 4 yields system uncertainty regions which are larger than the ones obtainable with the stochastic approach of Chapter 5.

  The conservatism of the deterministic approach is due to the fact that the cross-covariance noise bounds have to be chosen large in order to guarantee their correctness. In the stochastic approach there is no need for this, as the choice of a certain size for the system uncertainty region uniquely corresponds to a certain calculable probability and vice versa. This in turn has everything to do with the fact that the stochastic approach optimally exploits the property that the noise is independent of the input signal.

- The procedure for probabilistic identification is less sensitive to the prior information. Basically a small error in one of the priors implies a small error in the resulting system uncertainty regions. With respect to the procedure for deterministic identification such a statement can not be made. Theoretically it is possible that an erratic noise bound leads to system uncertainty regions which are complete nonsense.

For these reasons the procedure for probabilistic uncertainty bounding identification is recommended for use in the entire identification procedure. Note, however, that sometimes the use of the procedure for deterministic uncertainty bounding identification can not be avoided. This is the case if the output is disturbed by non-stationary

noise, or noise which is heavily correlated with the input signal without an independent external reference signal being available. The probabilistic approach is not applicable then, as the stochastic assumptions about the noise are not valid. However, a bounded amplitude noise assumption may still be valid, which means that the deterministic approach can be applied then.

The combination of the identification scheme of Figure 2.4 with the choices made in the previous discussion, leads to the following identification procedure:

Procedure 8.2.1

1. *Design an appropriate identification experiment and collect data.*
2. *Identify confidence regions for the system's frequency response with the procedure for probabilistic uncertainty bounding identification of Chapter 5.*
3. *Select an appropriate weighting in the identification criterion and identify a nominal model of some specified order with the frequency response curve fit procedure of Chapter 6.*
4. *Select an appropriate weighting in the uncertainty structure and calculate the upper bound on the H_∞-norm of the model error using Theorem 4.4.8.*
5. *If desired a model validation step may be carried out, or the nominal model with the model error bound may directly be used to design a robust controller.*

The most essential step of this identification procedure is step 2, the construction of frequency response uncertainty regions. In the first place these regions are used in step 3 to construct a nominal model. In the second place these regions are used in step 4 to calculate a model error bound. In fact step 3 is less essential. One may, for example, prefer another identification method to identify a control-relevant nominal model. Then step 4 can still be carried out, yielding an H_∞-model error bound for that particular nominal model. But step 4 often need not be carried out in full detail either. It is, for example, possible to use a non-parametric frequency dependent upper bound on the model error to prove robust stability or performance. This upper bound can be established on the basis of the frequency response system confidence regions without specifying a parametric weighting function.

The identification procedure 8.2.1 can be further refined. As becomes clear from the description in Chapter 5, the procedure for probabilistic uncertainty bounding identification makes use of a system expansion in terms of orthonormal basis functions. A proper choice of basis functions can mean a considerable reduction in the size of the uncertainty regions. The basis functions are properly chosen, if the model that generates the basis functions is a good description of the system, in the sense that their poles are close to each other. Consequently, the problem reduces to selecting a good nominal model, such that small system confidence regions can be identified in step 2 of the identification procedure. Obviously, a nominal model becomes available in step 3 of the identification procedure. This suggests an iterative approach to the identification problem. Before continuing to step 4 of the identification procedure a few iterations between step 2 and 3 are carried out. This iteration can be as follows:

Procedure 8.2.2

1. *Choose some set of basis functions* $\{B_k(z)\}_{k=0,\ldots,\infty}$, *for example the standard FIR basis* $B_k(z) = z^{-k}$.

2. *Identify frequency response system confidence regions with the procedure summarized in Section 5.7, using the specified set of basis functions.*

3. *Construct a nominal model with the curve fit procedure of Chapter 6, on the basis of the centers of the system confidence regions identified in the previous step.*

4. *Use this nominal model to construct a new set of basis functions* $\{B_k(z)\}_{k=0,\ldots,\infty}$ *with the procedure described in Heuberger et al. (1994).*

5. *Repeat step 2, 3 and 4 of this iteration scheme till satisfactory results are obtained, i.e. a good nominal model and small frequency response uncertainty regions.*

Convergence of this iteration is not guaranteed. But to the author's experience it leads in general to both an accurate nominal model and small uncertainty regions. An example of such an iteration is given in Section 8.4.

The identification results can even be further improved by introducing another iteration. This is the well-known iteration of identification and control design, as investigated by many researchers, see the discussion in Subsection 2.4.5 in Chapter 2. The purpose of such an iteration is to incorporate knowledge of the controller in the identification procedure, such that the identification can put additional emphasis on the frequency range important for the specific control application. In the present context this iteration could look as follows.

Procedure 8.2.3

1. *Design an open-loop experiment and take measurements of the system to be controlled.*

2. *Identify frequency response system confidence regions, a nominal model and a model error bound with the identification procedure 8.2.1, possibly applying the iterative procedure 8.2.2.*

3. *Design a controller based on the nominal model and the model error bound, such that maximum robust performance is achieved.*

4. *Design a new experiment with the system operating in closed loop with the designed controller and collect data.*

5. *Repeat step 2, 3 and 4 of this iteration scheme till a satisfactory controller is obtained.*

If the iterative procedure is carried out properly, it should be possible to gradually increase the control performance specifications during the iteration. The maximum

achievable performance in each iteration step automatically becomes clear from the given model error bound. In any case the controller should be designed to be robustly stable with respect to this model error bound. Thus, an important feature of this iterative scheme is that it can not lead to a controller which, designed for the model, destabilizes the system. In Section 8.4 an example is given of an iteration of identification and control design.

Remark 8.2.4 *Each time step 2 of iteration scheme 8.2.3 is carried out, a new set of system confidence regions becomes available. If the system is (approximately) linear, or if it operates in one operating point such that it can be considered as such, the new and old information can be combined by taking the* intersection *of these system confidence regions. Then the natural situation occurs that after each identification experiment the system becomes more accurately known. And then it is always possible to increase the control performance requirements in each iteration, as less controller robustness is required. Note that this approach is not possible, if the system is nonlinear and operates in a certain operating point which changes each time a new controller is implemented.*

Finally, note that the iteration in Procedure 8.2.2 is based on one data set, whereas Procedure 8.2.3 requires a new identification experiment in each iteration.

8.3 Unstable Systems

Until now it has always been assumed that the system to be identified and controlled is ℓ_∞-stable. If this is not the case, direct identification of the unstable system with the uncertainty bounding identification procedure of Chapter 5 (or 4) is not possible, as this procedure has been developed for stable systems. The identification problem can be solved in an indirect way. It is assumed that a stabilizing controller C is available, see Figure 8.1. Note that the presence of such a stabilizing controller makes it possible to collect data of the unstable system G_0.

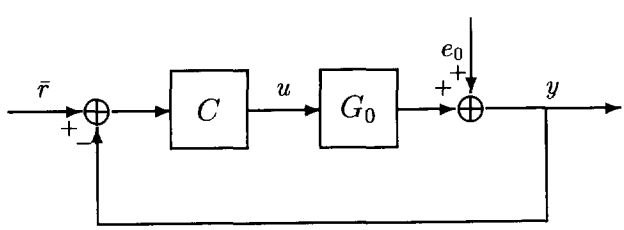

Fig. 8.1: System G_0 controlled by stabilizing controller C.

It is assumed that the external reference signal $\bar{r}$ is independent of the noise signal e_0. The system G_0 and controller C can be multivariable, but for simplicity it is assumed that the dimension (number of entries) of the signal $\bar{r}$ equals the dimension of

the input signal u. Note that in general the transfer function G_0 can not be uniquely identified if the dimension of $\bar{r}$ is less than the dimension of u.

The following relations hold,

$$\begin{aligned} u &= C(I + G_0C)^{-1}\bar{r} - C(I + G_0C)^{-1}e_0, \\ y &= G_0C(I + G_0C)^{-1}\bar{r} + (I + G_0C)^{-1}e_0. \end{aligned}$$

Denote the transfer function from $\bar{r}$ to u by $T_{u\bar{r}}$ and the one from $\bar{r}$ to y by $T_{y\bar{r}}$,

$$T_{u\bar{r}} = C(I + G_0C)^{-1}, T_{y\bar{r}} = G_0C(I + G_0C)^{-1}.$$

As the controller C is stabilizing, the transfer functions $T_{y\bar{r}}$ and $T_{u\bar{r}}$ are stable. Under the assumption that $\bar{r}$ and u have equal dimensions, the transfer function G_0 is given by

$$G_0 = T_{y\bar{r}}T_{u\bar{r}}^{-1}.$$

Instead of trying to directly identify the unstable system G_0, one can identify the transfer functions $T_{y\bar{r}}$ and $T_{u\bar{r}}$. For this measurements are needed of the external reference signal $\bar{r}$, the input signal u and the output signal y. As the signals $\bar{r}$ and e_0 are assumed independent, this then leads to two open loop identification problems. With the identification procedure 8.2.1 models $\widehat{T}_{y\bar{r}}$ and $\widehat{T}_{u\bar{r}}$ can be identified, in combination with model error bounds for each of these transfer functions. This leads to a structured uncertainty description. Based on the model and the structured uncertainty, a controller can be designed for the system G_0 with, for example, a μ-control design procedure.

In Schrama (1992a, 1992b) a different approach is described, which is also applicable in the present situation. Instead of knowledge of the external reference signal $\bar{r}$, that approach requires knowledge of the controller C. It is shown that the problem of identifying an unstable system G_0 in closed loop, can be recast into the problem of identifying two stable coprime factors of G_0 in open loop. This is similar to the approach described above. Successful application of this approach in combination with a frequency response uncertainty bounding identification procedure has been reported in De Vries (1994). In Van den Hof *et al.* (1993b) a comparison is made between this closed loop identification approach and the approach described above.

8.4 Examples

8.4.1 Example 1: Iteration to Determine Basis Functions

In this subsection the iterative Procedure 8.2.2 is illustrated by means of an example.

Consider the system

$$G_0(z) = \frac{0.0273z^4 - 0.08128z^3 + 0.09329z^2 - 0.05112z + 0.01186}{z^5 - 4.295z^4 + 7.61z^3 - 6.954z^2 + 3.283z - 0.643}.$$

The input signal $u(t)$ is chosen to be a sum of 100 sinusoids of equal amplitude $\sqrt{2}/10$, with frequencies logarithmically distributed between 0.01 and π, and such that the

signal is periodic with a period time of 1000 samples. In order to reduce time-domain peaks in the input signal, random phases have been assigned to the sinusoids. In Figure 8.2 the absolute value of the discrete Fourier transform of the input signal is depicted, in particular $\left|\frac{1}{\sqrt{1000}}\sum_{t=1}^{1000} u(t)e^{-i\omega_j t}\right|$.

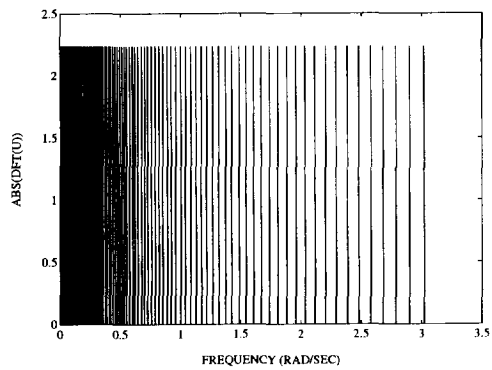

Fig. 8.2: Absolute value of the DFT of the input signal $u(t)$.

The measured output sigal $y(t)$ is disturbed by low-pass filtered white noise,

$$y(t) = G_0(q)u(t) + e_0(t),\ e_0(t) = \frac{0.1q}{q-0.5}w_0(t),$$

where $w_0(t)$ is gaussian white noise with variance 1.

The system has been initiated at $t = 1$, but the first 1000 samples are not used for identification. The samples from $t = 1001$ till $t = 3000$ are used for identification purposes.

As recommended in Remark 3.3.4 in Chapter 3, the signal

$$\hat{e}(t) := (y(t+1000) - y(t))/\sqrt{2},\ t = 1001, \ldots, 2000,$$

is used to estimate the auto-covariance function of the noise process, which is needed to estimate frequency response system confidence regions.

First, a standard FIR basis is used in the uncertainty bounding identification procedure of Chapter 5, i.e. $B_k(z) = z^{-k}$, $\forall k$. A prior pulse response bound $\bar{g}(k)$ has been chosen. This bound is depicted in Figure 8.3, together with the pulse response of the system $G_0(z)$.

For each of the 100 frequencies excited by the input signal, ellipsoidal frequency response confidence regions have been identified with $c_{\chi,\alpha} = 9.21$, which corresponds to a confidence level of 99%. In the procedure (see equation (5.32) in Chapter 5) use has been made of a Tukey-window, as specified by (5.34) in Appendix 5.B, with window-parameters $f = 0.5$ and $w = 60$. An optimal value of the truncation parameter n (see equation 5.6) appears to be 100. This value is optimal in the sense that the sum

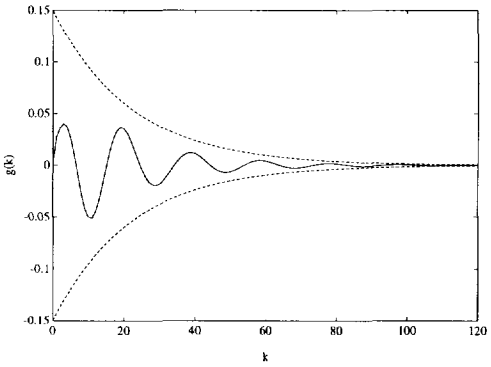

Fig. 8.3: Prior pulse response bound $\pm\bar{g}(k)$ (dashed) and pulse response of the system $G_0(z)$ (solid).

of bias and variance errors is minimal for this choice of n. The identification result is depicted in a Nyquist diagram in Figure 8.4.

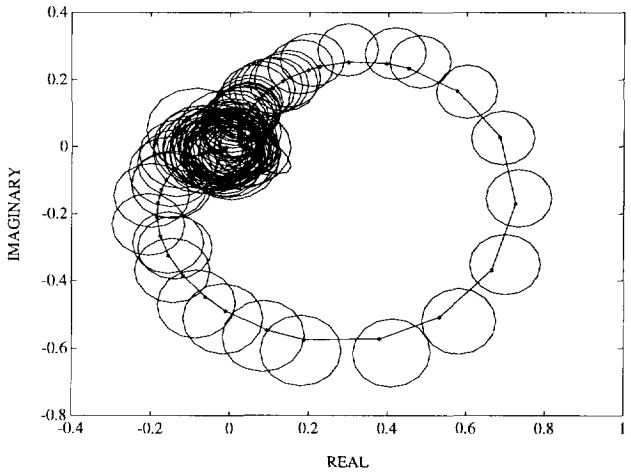

Fig. 8.4: Nyquist diagram identified 99%-confidence regions for the system's frequency response (ellipsoids) and system's frequency response $G_0(e^{i\omega_j})$ (solid,$*$). In the identification use has been made of a standard FIR basis.

With the curve fit procedure described in Chapter 6 a third order nominal model $\widehat{G}(z)$ has been identified, on the basis of the centers of the uncertainty regions depicted in Figure 8.4. The weighting $\widetilde{W}(\omega_j)$ has been chosen equal to 1. The resulting model

is equal to

$$\widehat{G}(z) = \frac{0.01574z^3 - 0.03212z^2 + 0.05729z - 0.04174}{z^3 - 2.33616z^2 + 1.87885z - 0.49042}.$$

In Figure 8.5 an amplitude Bode diagram is depicted of the system $G_0(z)$, the central estimate $\widehat{G}_C(\omega_j)$ and the curve fit model $\widehat{G}(z)$.

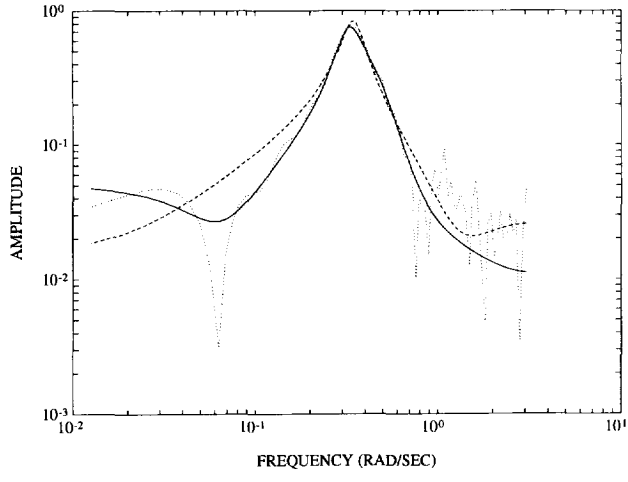

Fig. 8.5: Amplitude Bode diagram of the system $G_0(z)$ (solid), the central estimate $\widehat{G}_C(\omega_j)$ (dotted) and the curve fit model $\widehat{G}(z)$ (dashed).

Next, with the method described in Heuberger *et al.* (1994) orthonormal basis functions $B_k(z)$ have been constructed, on the basis of the model $\widehat{G}(z)$. A prior bound $\bar{g}(k)$ has been chosen for the generalized pulse response of the system $G_0(z)$. This bound is depicted in Figure 8.6, together with the orthonormal expansion coefficients $g_0(k)$ of the system $G_0(z)$.

With this new set of orthonormal basis functions the procedure for probabilistic uncertainty bounding identification of Chapter 5 has been applied. For each of the 100 frequencies excited by the input signal ellipsoidal frequency response confidence regions have been identified, corresponding to a confidence level of 99%. Now an optimal value of the truncation parameter n (see equation 5.6) appears to be 27. The identification result is depicted in a Nyquist diagram in Figure 8.7.

A measure for the size of the identified uncertainty regions is the so-called *radius of uncertainty*. For a given frequency this radius is defined here as the maximum distance of the central estimate to the boundary of the ellipsoidal confidence region for that frequency. Radii of uncertainty can be calculated for both the uncertainty regions depicted in Figure 8.4 and the ones shown in Figure 8.7. The result is shown in Figure 8.8.

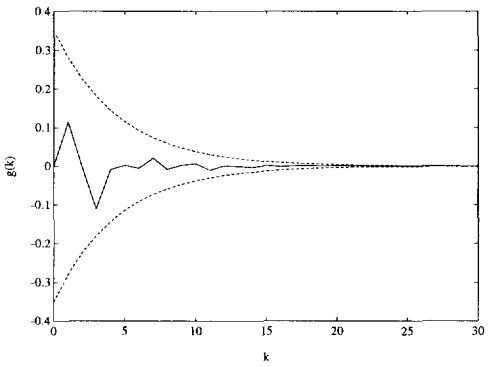

Fig. 8.6: Prior bound $\pm\bar{g}(k)$ on the system's generalized pulse response (dashed) and generalized pulse response $g_0(k)$ of the system $G_0(z)$ (solid).

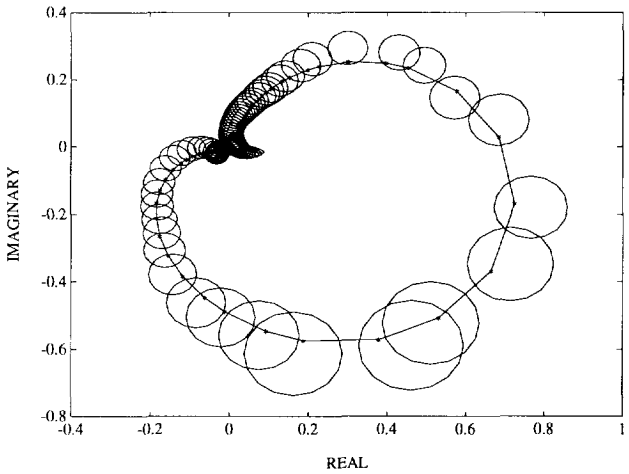

Fig. 8.7: Nyquist diagram identified 99%-confidence regions for the system's frequency response (ellipsoids) and system's frequency response $G_0(e^{i\omega_j})$ (solid,$*$). In the identification use has been made of generalized orthonormal basis functions generated by the model $\widehat{G}(z)$.

A comparison between Figure 8.4 and 8.7, or a look at Figure 8.8, shows that the system uncertainty regions in Figure 8.7 are significantly smaller in the low- and high-frequency ranges. In the medium frequency range the system uncertainty regions in Figure 8.4 are slightly smaller than the ones shown in Figure 8.7. It becomes apparent that a proper use of orthonormal basis functions can substantially improve the quality

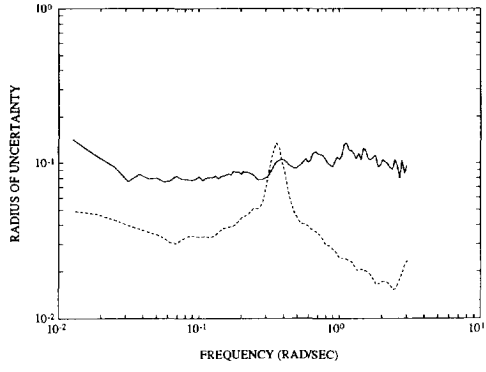

Fig. 8.8: Radii of uncertainty as a function of frequency for the uncertainty regions depicted in Figure 8.4 (solid) and 8.7 (dashed).

of the identification result. This is explained by the fact that, due to faster convergence, less parameters need to be identified, which implies a decrease in the variance error (compare the values of n used in both cases). Probably further improvements can be obtained by continuing the iterative Procedure 8.2.2, but this is not carried out here.

It is interesting to show one more feature of the uncertainty bounding identification procedure in combination with the use of orthonormal basis functions. Namely it is possible to identify frequency response confidence regions for frequencies which have *not* been excited by the input signal. This feature is due to the fact that a parametric identification method interpolates between frequencies. In order to show this feature, frequency response system uncertainty regions have been identified for frequencies *between* the 100 frequencies excited by the input signal, on the basis of both the standard FIR basis and the generalized orthonormal basis. The resulting confidence regions are not shown here. Similar to before radii of uncertainty can be calculated for the identified frequency response confidence regions. These radii are depicted in Figure 8.9.

The figure shows that the interpolation works fine in case use is made of the generalized orthonormal basis. The system uncertainty is not larger than the system uncertainty at the excitation frequencies. The use of a FIR basis leads to acceptable results in the low-frequency range. But in the high-frequency range the use of a FIR basis leads to extremely large frequency response uncertainty regions. If a FIR basis is used, many parameters need to be estimated ($n = 100$), which causes identifiability problems for the high frequencies which have not been excited by the input signal.

Remark 8.4.1 *In the uncertainty bounding identification procedure the inverse of a matrix* $\frac{1}{N}\sum_{t=t_s}^{N} V(t)U^T(t)$ *needs to be computed, see Step 11 in the overview in Section 5.7. This may be an ill-conditioned operation. Especially if the input signal is insufficiently informative, numerical problems can occur.*

In fact this situation occurs in the first part of the present example, where the stan-

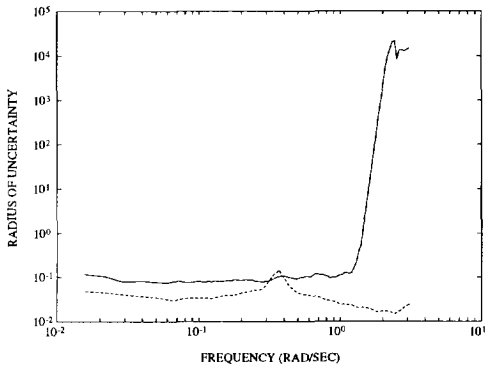

Fig. 8.9: Radii of uncertainty for frequencies between the excitation frequencies. Uncertainty bounding identification has been performed with the standard FIR basis (solid) and the model-based generalized orthonormal basis (dashed).

dard FIR basis has been used. The matrix $\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t)$ *has a condition number of about* 10^{-17}*, which makes it impossible to accurately compute the inverse. The explanation for this ill-conditioning is the fact that a large number of model parameters is estimated in comparison to the number of sinusoids in the input signal. The consequence is that the system uncertainty is large in some frequency regions, especially the high-frequency range. However, the system uncertainty is not that large for the frequencies which have been excited by the input signal.*

In order to circumvent numerical problems, the following ad hoc procedure has been applied in the example in this section. A singular value decomposition is determined of the matrix to be inverted,

$$\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t) = \widetilde{U}\Sigma\widetilde{V}^T.$$

Basically the inverse of the given matrix is equal to

$$\left[\frac{1}{\widetilde{N}} \sum_{t=t_s}^{N} V(t)U^T(t)\right]^{-1} = \widetilde{V}\Sigma^{-1}\widetilde{U}^T,$$

but its computation can be ill-conditioned due to near-singularities in the matrix Σ*. To overcome this problem, singular values in the matrix* Σ *which are almost identically zero, are set to a certain small value, for example* 10^{-12}*. The singular values larger than* 10^{-12} *remain unchanged. This yields a modified matrix* Σ_{M}*. Finally, the matrix*

$$\widetilde{V}\Sigma_{\mathrm{M}}^{-1}\widetilde{U}^T$$

is used to replace the inverse

$$\left[\frac{1}{\widetilde{N}}\sum_{t=t_s}^{N} V(t)U^T(t)\right]^{-1}.$$

In the example in this section this procedure has been applied and works satisfactorily. The results for frequencies with small frequency response system uncertainty are hardly affected by the operation. Nevertheless, it certainly deserves future research attention to develop a less heuristic procedure which solves the numerical problems. It seems possible to adjust the uncertainty bounding identification procedure in such a way that no inverse needs to be computed, which circumvents the numerical problems.

8.4.2 Example 2: Iteration of Identification and Robust Control Design

In this subsection the iterative Procedure 8.2.3 is illustrated by means of an example.

Consider the system

$$G_0(z) = \frac{0.05\prod_{k=1}^{6}(z - z_k)}{\prod_{k=1}^{7}(z - p_k)},$$

with the zeros of the system given by

$$z_1 = 0.63 + 0.8i,\ z_2 = 0.63 - 0.8i,\ z_3 = 0.85 + 0.54i,$$
$$z_4 = 0.85 - 0.54i,\ z_5 = 0.95 + 0.26i,\ z_6 = 0.95 - 0.26i,$$

and the poles of the system given by

$$p_1 = 0.75 + 0.57i,\ p_2 = 0.75 - 0.57i,\ p_3 = 0.84 + 0.37i,\ p_4 = 0.84 - 0.37i,$$
$$p_5 = 0.91,\ p_6 = 0.93 + 0.23i,\ p_7 = 0.93 - 0.23i.$$

The output of the system is disturbed by low-pass filtered white noise,

$$y(t) = G_0(q)u(t) + e_0(t),\ e_0(t) = \frac{0.1q}{q - 0.5}w_0(t),$$

where $w_0(t)$ is gaussian white noise with variance 1.

First, an open loop identification experiment has been performed. Starting from zero initial conditions a gaussian white input signal with unit variance has been applied to the system. Altogether 1000 measurements have been collected for identification purposes.

On the basis of the measurement data, frequency response uncertainty regions are identified with the procedure described in Chapter 5. For that purpose it is necessary to have available an estimate of the auto-covariance function of the noise process. The

procedure introduced in Subsection 5.6.1 is used. A standard output error identification method has been applied to the first 500 measured data samples and a strictly proper 5th order model has been identified. An amplitude Bode diagram of the resulting model is shown in Figure 8.10.

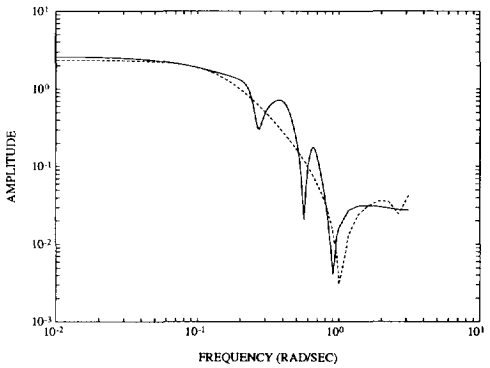

Fig. 8.10: Amplitude Bode diagram of the system $G_0(z)$ (solid) and the identified 5th order output error model (dashed).

The simulation or output error of this 5th order model has been determined for the second 500 data samples. The auto-covariance function of this output error is used as an estimate of the auto-covariance function of the noise process. As explained in Subsection 5.6.1 this leads to conservative results in general. In Figure 8.11 the true and estimated auto-covariance function of the noise process are depicted.

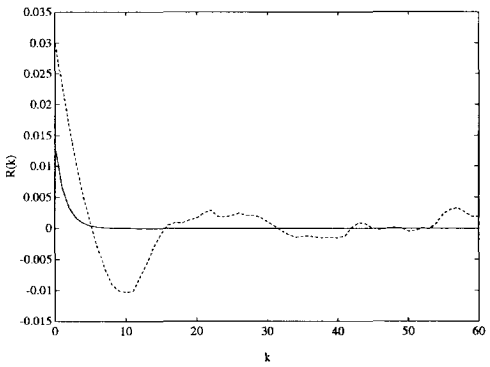

Fig. 8.11: Auto-covariance function of the noise process $\{e_0(t)\}$ (solid) and estimated auto-covariance function (dashed).

A standard FIR basis is used in the uncertainty bounding identification procedure.

The prior pulse response bound that has been used, is given by

$$\bar{g}(k) = 0.4(0.95)^k, \ k = 0, \ldots, \infty.$$

In fact this bound is slightly too small for large values of k. But, as argued before, the identification procedure of Chapter 5 is robust to small errors in the prior assumptions. The identification result is only little influenced by the specific choice of prior pulse response bound.

For 100 frequencies between 0.01 and π rectangular frequency response confidence regions $\mathcal{P}_1(\omega_j)$ have been identified with $c_{\mathcal{N},\alpha} = 3.5$, which corresponds to a confidence level of 99.9%. In the procedure use has been made of a Tukey-window, as specified by (5.34) in Appendix 5.B, with window-parameters $f = 0.5$ and $w = 60$. An optimal value of the truncation parameter n (see equation 5.6) appears to be 100. The identification result is depicted in a Nyquist diagram in Figure 8.12.

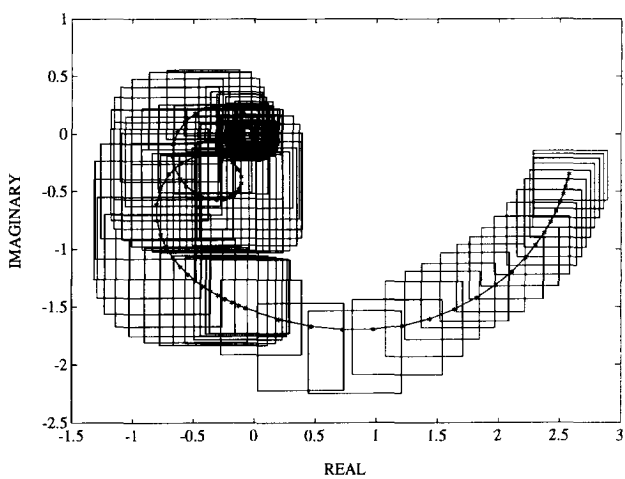

Fig. 8.12: Nyquist diagram system's frequency response $G_0(e^{i\omega_j})$ (solid,$*$) and 99.9%-confidence regions $\mathcal{P}_1(\omega_j)$ (rectangulars) identified on the basis of open loop measurement data.

With the curve fit procedure described in Chapter 6 a strictly proper fifth order nominal model $\widehat{G}_1(z)$ has been identified, on the basis of the centers of the uncertainty regions $\mathcal{P}_1(\omega_j)$. The weighting $\widetilde{W}(\omega_j)$ has been chosen equal to 1. The resulting model is equal to

$$\widehat{G}_1(z) = \frac{0.06022z^4 - 0.20686z^3 + 0.29813z^2 - 0.21191z + 0.06274}{z^5 - 4.40076z^4 + 7.90253z^3 - 7.23511z^2 + 3.37883z - 0.64461}.$$

In Figure 8.13 an amplitude Bode diagram is depicted of the system $G_0(z)$, the central estimate $\widehat{G}_C(\omega_j)$ and the curve fit model $\widehat{G}_1(z)$.

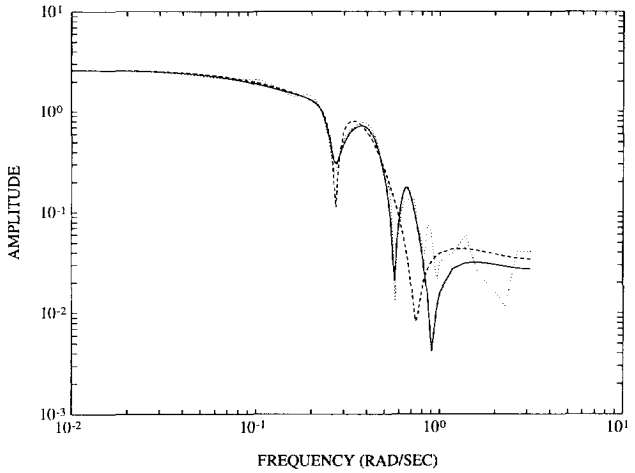

Fig. 8.13: Amplitude Bode diagram of the system $G_0(z)$ (solid), the central estimate $\widehat{G}_C(\omega_j)$ (dotted) and the curve fit model $\widehat{G}_1(z)$ (dashed).

On the basis of the model $\widehat{G}_1(z)$ a controller is designed with the H_∞-control design procedure described in Bongers and Bosgra (1990), Bongers (1994) and McFarlane and Glover (1989). The most important features of this (multivariable) H_∞-control design procedure are as follows.

Procedure 8.4.2 *Let $\widehat{G}(z)$ be a nominal model, possibly multivariable, and let $W_1(z)$ and $W_2(z)$ be weighting matrices. Define $\widetilde{G}(z)$ as*

$$\widetilde{G}(z) := W_1(z)\widehat{G}(z)W_2(z).$$

The control problem concerns the minimization of

$$\left\| T(\widetilde{G}(z), \widetilde{C}(z)) \right\|_\infty$$

over stabilizing $\widetilde{C}(z)$ of some specified order. Here the matrix T is defined as

$$T(\widetilde{G}, \widetilde{C}) := \begin{bmatrix} I \\ \widetilde{G} \end{bmatrix} (I + \widetilde{C}\widetilde{G})^{-1} \begin{bmatrix} I & \widetilde{C} \end{bmatrix}.$$

The controller

$$C(z) := W_2(z)\widetilde{C}(z)W_1(z)$$

stabilizes $\widehat{G}(z)$ and achieves a certain performance and robustness, dependent on the choice of weighting matrices $W_1(z)$ and $W_2(z)$. The controller $C(z)$ is designed for the

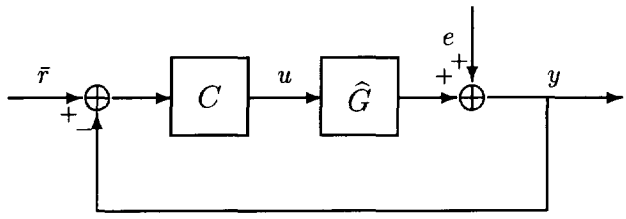

Fig. 8.14: Standard closed loop configuration.

standard configuration depicted in Figure 8.14. In this figure u is the input signal, y the (disturbed) output signal, $\bar{r}$ the external reference signal and e the output noise.

In the SISO example considered here the weighting functions in the control design are chosen as follows,

$$W_1(z) = 1, \ W_2(z) = \frac{\alpha z}{z-1}.$$

The presence of an integrator in $W_2(z)$ provides integrating action in the resulting controller. The parameter α can be used to make a trade-off between controller performance and robustness. A large value of α generally gives rise to a high bandwidth of the closed loop system, i.e. a high controller performance, but also little controller robustness. On the other hand, a small value of α leads to a very robust, but low-performing, controller.

The control design so far only makes use of the nominal model $\widehat{G}_1(z)$. But in the control design procedure the identified uncertainty regions $\mathcal{P}_1(\omega)$ can also be taken into account as explained in the sequel.

The so-called $M = 1.3$-circle in the complex plane is the circle with center $\frac{-(1.3)^2}{(1.3)^2-1}$ and radius $\frac{1.3}{(1.3)^2-1}$. Each (complex-valued) element $c = a + bi$ of this circle has the property that $|c/(1+c)| = 1.3$. Moreover, each point c inside the circle has the property that $|c/(1+c)| > 1.3$, and for each point c outside the circle $|c/(1+c)| < 1.3$.

The importance of the $M = 1.3$-circle becomes clear from the following observation. Let a system G be controlled by a controller C. Let the product CG be plotted in a Nyquist diagram, together with the $M = 1.3$-circle. Then visual inspection immediately shows if the amplitude of the complementary sensitivity function $\frac{CG}{1+CG}$ exceeds the value 1.3 or not. If the Nyquist curve of CG is entirely outside the $M = 1.3$-disc, then $\left|\frac{CG}{1+CG}\right| < 1.3$ for all frequencies. In fact the robust performance of the controller C is visually evaluated in this way.

Remark 8.4.3 *Note that the use of the $M = 1.3$-circle is more or less arbitrary. One might prefer another robust performance margin. In general the $M = m$-circle is the circle with center $\frac{-m^2}{m^2-1}$ and radius $\frac{m}{m^2-1}$. It specifies all points c in the complex plane for which*

$$\left|\frac{c}{1+c}\right| = m.$$

This circle can be used to bound the amplitude of the complementary sensitivity function of the closed loop system. In the special case $m = \infty$ this reduces to a robust stability test, as then the M-circle reduces to the point -1.

Another possibility is to consider the circle with center -1 and radius $1/n$. This circle specifies all points c in the complex plane for which

$$\left|\frac{1}{1+c}\right| = n.$$

This circle can be used to bound the amplitude of the sensitivity function of the closed loop system. In the special case $n = \infty$ this again reduces to a robust stability test.

An important property of the above described visual robust performance analysis method is its applicability in case the model G is replaced by a set of models, for example uncertainty regions $\mathcal{P}(\omega_j)$. Instead of one curve GC, a set of curves $\mathcal{P}(\omega_j)C(\omega_j)$ can be plotted in the Nyquist diagram, and possible intersection with the $M = 1.3$-circle can be noticed.

This leads to the following robust control design procedure, which makes use of both the nominal model $\widehat{G}_1(z)$ and the uncertainty regions $\mathcal{P}_1(\omega_j)$. For a given value of α a controller $C_1(z)$ can be designed on the basis of the nominal model $\widehat{G}_1(z)$. Now the parameter α is chosen as large as possible, which corresponds to maximum performance and minimum robustness, but under the side-constraint that the set of Nyquist curves $\mathcal{P}_1(\omega_j)C_1(e^{i\omega_j})$ does not intersect the $M = 1.3$-circle. As explained above the latter is checked by visual inspection. In this way the parameter α is iteratively adjusted to an optimal value, which appears to be $\alpha = 0.057$. The optimal controller is equal to

$$C_1(z) = \frac{0.09655 \prod_{k=1}^{6} (z - z_k)}{\prod_{k=1}^{6} (z - p_k)},$$

with the zeros being given by

$$z_1 = 0.823 - 0.314i,\ z_2 = 0.823 + 0.314i,\ z_3 = 0.927 - 0.251i,$$
$$z_4 = 0.927 + 0.251i,\ z_5 = 0.888,\ z_6 = 0,$$

and the poles by

$$p_1 = 0.81 + 0.349i,\ p_2 = 0.81 - 0.349i,\ p_3 = 0.942 + 0.259i,$$
$$p_4 = 0.942 - 0.259i,\ p_5 = 0.721,\ p_6 = 1.$$

In Figure 8.15 a Nyquist diagram is depicted of $\widehat{G}_1(z)C_1(z)$, $\mathcal{P}_1(\omega_j)C_1(e^{i\omega_j})$ and the $M = 1.3$-circle.

Note that in the design of the controller $C_1(z)$ an optimal trade-off has been made between robustness and performance of the controller. In a non-conservative way

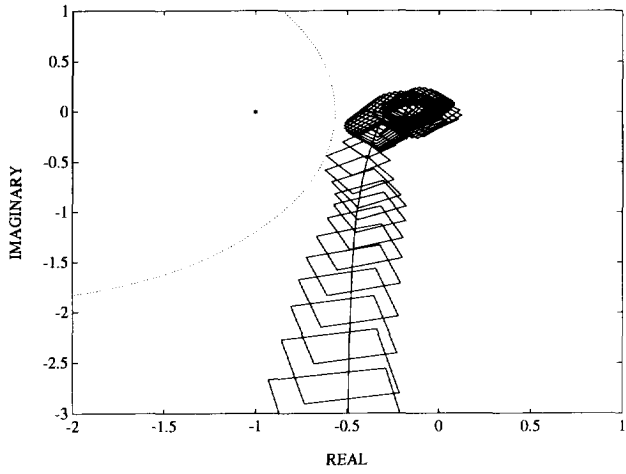

Fig. 8.15: Nyquist diagram of $\widehat{G}_1(z)C_1(z)$ (solid), $\mathcal{P}_1(\omega_j)C_1(e^{i\omega_j})$ (boxes), $M = 1.3$-circle (dotted) and the critical point -1 $(*)$.

explicit use has been made of the identified frequency response system uncertainty regions.

The results may be presented in yet another fashion. On the basis of the uncertainty regions $\mathcal{P}_1(\omega_j)C_1(e^{i\omega_j})$ it is straightforward to calculate upper and lower bounds on the amplitude of the sensitivity function $\frac{1}{1+C_1G_0}$. This is done by calculating for each frequency ω_j,

$$s_1(\omega_j) := \min_{G(\omega_j)\in\mathcal{P}_1(\omega_j)} \left|1 + C_1(e^{i\omega_j})G(\omega_j)\right|,$$

which can be solved with convex programming, and

$$s_u(\omega_j) := \max_{G(\omega_j)\in\mathcal{P}_1(\omega_j)} \left|1 + C_1(e^{i\omega_j})G(\omega_j)\right|,$$

which can be solved by evaluating the 4 vertices of $\mathcal{P}_1(\omega_j)$. Upper and lower bounds on the amplitude of the sensitivity function are then given by

$$\frac{1}{s_u(\omega_j)} \leq \left|\frac{1}{1 + C_1(e^{i\omega_j})G(\omega_j)}\right| \leq \frac{1}{s_1(\omega_j)}, \ \forall G(\omega_j) \in \mathcal{P}_1(\omega_j).$$

The calculated upper and lower bounds are shown in an amplitude Bode diagram in Figure 8.16. The sensitivity function of the system $G_0(z)$ is depicted as well.

Next, a new identification experiment is performed. The objective is to more accurately identify the system in the closed loop relevant frequency ranges. If the system is more accurately known for these frequencies, a new controller can be designed with increased robust performance. In order to emphasize the closed loop relevant

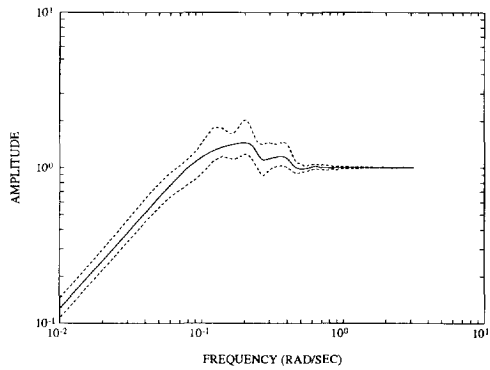

Fig. 8.16: Amplitude Bode diagram of the sensitivity function $\frac{1}{1+C_1(z)G_0(z)}$ (solid) and identified upper and lower bounds (dashed).

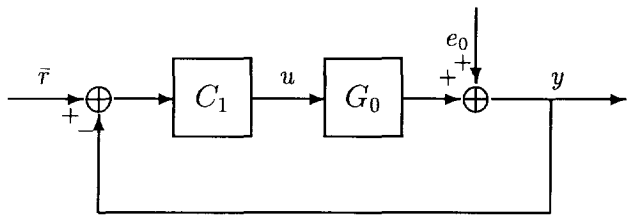

Fig. 8.17: Closed loop of system G_0 and controller C_1.

frequencies in the identification experiment, measurements are carried out in closed loop, see Figure 8.17.

During the identification experiment the external reference signal has been chosen as follows,

$$\bar{r}(t) = \frac{6}{1 + C_1(q)\widehat{G}_1(q)}\bar{w}(t),$$

with $\{\bar{w}(t)\}$ a white gaussian signal with variance 1. The filter gives additional emphasis on the feedback-relevant frequencies. The gain of the filter has been chosen such that the energy of the input signal $\{u(t)\}$ during the closed loop experiment equals the energy of the input signal during the open loop experiment carried out previously.

Again, 1000 data samples have been collected. The procedure of Chapter 5 is applied to identify frequency response uncertainty regions. The signal $\{r(t)\}$ (see step 3 in Section 5.7) is chosen equal to

$$r(t) = \frac{C_1(q)}{1 + C_1(q)\widehat{G}_1(q)}\bar{r}(t),$$

which corresponds to an estimate of the noise free part of the input signal $u(t)$. Again, a FIR basis is used. The prior pulse response bounds $\bar{g}(k)$ and the estimated autocovariance function of the noise process remain unchanged. The resulting uncertainty regions $\mathcal{P}_2(\omega_j)$ are depicted in a Nyquist diagram in Figure 8.18.

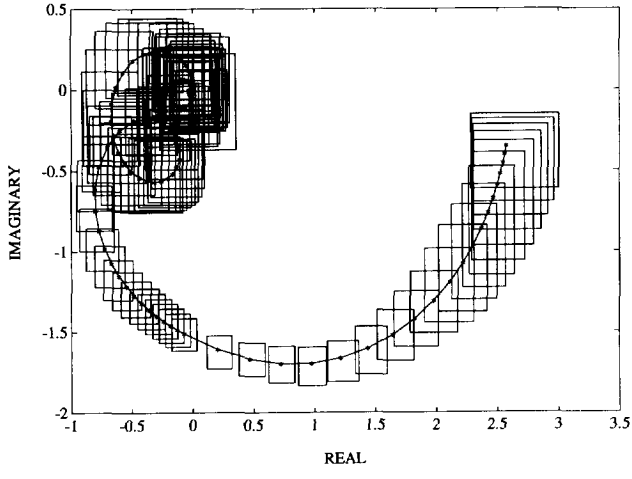

Fig. 8.18: Nyquist diagram system's frequency response $G_0(e^{i\omega_j})$ (solid,*) and 99.9%-confidence regions $\mathcal{P}_2(\omega_j)$ (rectangulars) identified on the basis of closed loop measurement data.

As explained in Example 1 in this section, a measure for the size of an uncertainty region is the radius of uncertainty. In Figure 8.19 the radii of uncertainty are shown for the uncertainty regions $\mathcal{P}_1(\omega_j)$ and $\mathcal{P}_2(\omega_j)$. The figure clearly shows that the uncertainty regions identified on the basis of the closed loop experiment are relatively small in the medium-frequency range, which is in this case the feedback-relevant frequency range. Obviously, an increase of input energy for certain frequencies reduces the uncertainty about the system for these frequencies.

The identification results for the open and closed loop experiments can be combined by calculating the *intersection* $\mathcal{P}_{12}(\omega_j)$ of the uncertainty region $\mathcal{P}_1(\omega_j)$ and the uncertainty region $\mathcal{P}_2(\omega_j)$. This is done for each frequency ω_j separately. The resulting intersections are shown in Figure 8.20. Note that the confidence of the intersected uncertainty regions reduces to 99.8%.

With the curve fit procedure described in Chapter 6 a strictly proper fifth order nominal model $\widehat{G}_2(z)$ has been identified, on the basis of the centers $\widehat{G}_C(\omega_j)$ of the uncertainty regions $\mathcal{P}_{12}(\omega_j)$. The weighting $\widetilde{W}(\omega_j)$ has been chosen equal to

$$\widetilde{W}(\omega_j) = \frac{C_1(e^{i\omega_j})}{1 + C_1(e^{i\omega_j})\widehat{G}_C(e^{i\omega_j})},$$

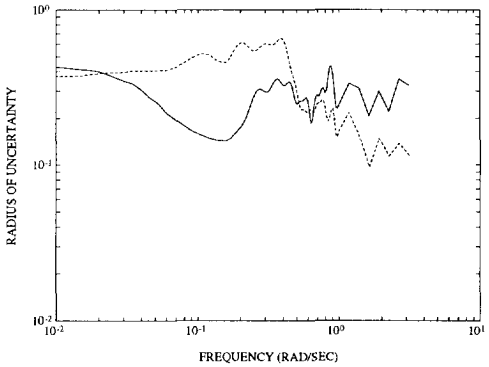

Fig. 8.19: Radii of uncertainty as a function of frequency for the uncertainty regions $\mathcal{P}_1(\omega_j)$ (dashed) and $\mathcal{P}_2(\omega_j)$ (solid).

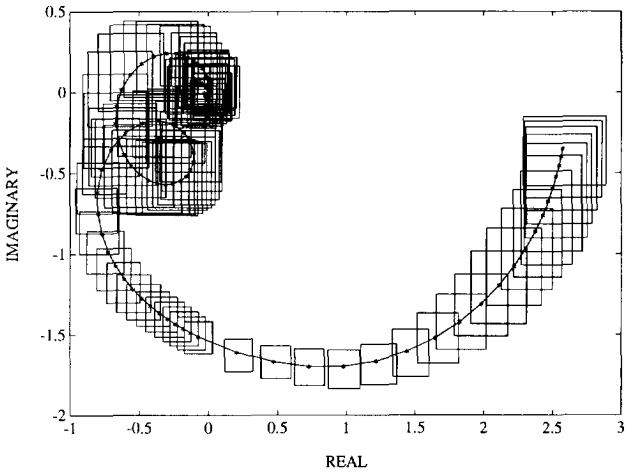

Fig. 8.20: Nyquist diagram system's frequency response $G_0(e^{i\omega_j})$ (solid,$*$) and 99.8%-confidence regions $\mathcal{P}_{12}(\omega_j)$ (rectangulars) which are the intersections of the uncertainty regions $\mathcal{P}_1(\omega_j)$ and $\mathcal{P}_2(\omega_j)$.

which emphasizes the feedback-relevant frequencies. The resulting model is equal to

$$\widehat{G}_2(z) = \frac{0.05694z^4 - 0.19163z^3 + 0.26261z^2 - 0.17346z + 0.04706}{z^5 - 4.53127z^4 + 8.38473z^3 - 7.91576z^2 + 3.81216z - 0.74929}.$$

In Figure 8.21 an amplitude Bode diagram is depicted of the system $G_0(z)$, the central estimate $\widehat{G}_C(\omega_j)$ and the curve fit model $\widehat{G}_2(z)$.

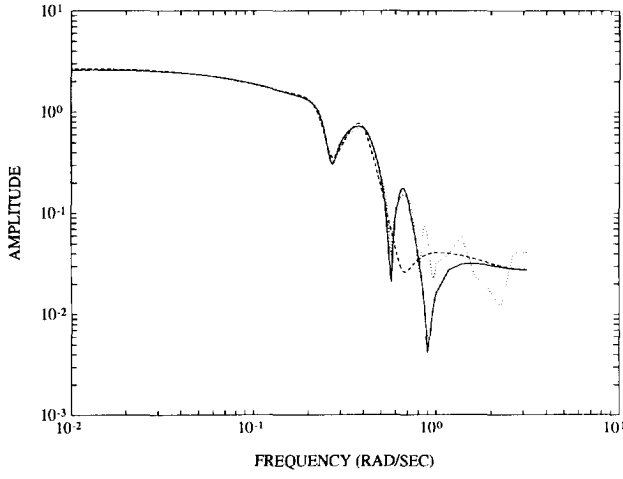

Fig. 8.21: Amplitude Bode diagram of the system $G_0(z)$ (solid), the central estimate $\widehat{G}_C(\omega_j)$ (dotted) and the curve fit model $\widehat{G}_2(z)$ (dashed).

With the same procedure as before, a robust controller $C_2(z)$ is designed on the basis of the model $\widehat{G}_2(z)$ and the uncertainty regions $\mathcal{P}_{12}(\omega_j)$. Due to the reduced size of the uncertainty regions in the relevant frequency range, the control design parameter α can be increased to 0.145. The resulting optimal controller is equal to

$$C_2(z) = \frac{0.2955 \prod_{k=1}^{6} (z - z_k)}{\prod_{k=1}^{6} (z - p_k)},$$

with the zeros being given by

$$z_1 = 0.876 + 0.358i,\ z_2 = 0.876 - 0.358i,\ z_3 = 0.927 + 0.209i,$$
$$z_4 = 0.927 - 0.209i,\ z_5 = 0.893,\ z_6 = 0,$$

and the poles by

$$p_1 = 0.839 + 0.379i,\ p_2 = 0.839 - 0.379i,\ p_3 = 0.927 + 0.25i,$$
$$p_4 = 0.927 - 0.25i,\ p_5 = 0.678,\ p_6 = 1.$$

In Figure 8.22 a Nyquist diagram is depicted of $\widehat{G}_2(z)C_2(z)$, $\mathcal{P}_{12}(\omega_j)C_2(e^{i\omega_j})$ and the $M = 1.3$-circle.

On the basis of the uncertainty regions $\mathcal{P}_{12}(\omega_j)$ upper and lower bounds on the amplitude of the sensitivity function can be calculated. These are shown in the ampli-

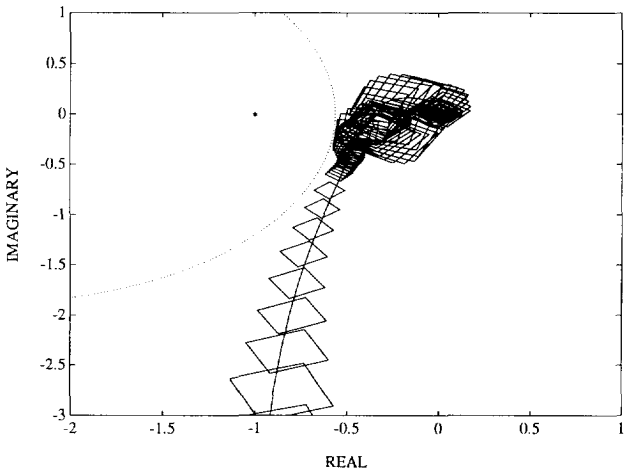

Fig. 8.22: Nyquist diagram of $\widehat{G}_2(z)C_2(z)$ (solid), $\mathcal{P}_{12}(\omega_j)C_2(e^{i\omega_j})$ (boxes), $M = 1.3$-circle (dotted) and the critical point -1 $(*)$.

tude Bode diagram in Figure 8.23, together with the sensitivity function of the system $G_0(z)$.

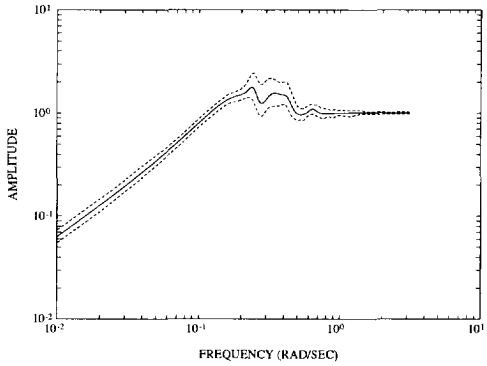

Fig. 8.23: Amplitude Bode diagram of the sensitivity function $\frac{1}{1+C_2(z)G_0(z)}$ (solid) and identified upper and lower bounds (dashed).

A comparison with Figure 8.16 immediately shows the increased bandwidth obtained with the new controller. If desired, the controller performance can be further increased by continuing the iteration of identification and control design. An important feature of the iterative scheme outlined in this example is that the identification

results of previous iterations can be used in a next iteration. This is done by applying the intersection of all uncertainty regions available, as demonstrated in this example. Obviously, then in each iteration the uncertainty about the system is guaranteed to decrease, and for that reason in each iteration the controller performance can be increased. Note that, due to the use of explicit probabilistic uncertainty bounds, the statement can be made that it is very likely that this controller performance is indeed achieved for the system.

8.5 Discussion

In this chapter a practically applicable robust-control-oriented identification procedure has been introduced. The features of the procedure have been demonstrated in two simulation examples. The identification procedure has primarily been developed for industrial processes, but its application is not restricted to this class of systems. The procedure is applicable to all other kinds of systems as well, such as mechanical systems. Essential features of the identification procedure are:

- An accurate nominal model of a user-defined order can be identified.
- An explicit and non-conservative frequency dependent bound on the model error can be identified.
- There is a clear distinction between the error due to undermodelling (bias) and the error due to noise (variance). The bias error can be shaped in a frequency dependent way by a suitable choice of model order and identification criterion. The variance error can be shaped in a frequency dependent way by a suitable experiment design.
- All prior information can be estimated from data.
- There are no restrictions on the input signal, except that it should be sufficiently informative in the frequency ranges of interest.
- The procedure is MIMO applicable, as all steps in the procedure have MIMO extensions.

On the other hand, some drawbacks of the identification procedure are:

- The procedure is computationally demanding and therefore currently not suited for on-line applications, or for processes with many inputs and outputs.
- The identification procedure only works for systems which are almost linear, or at least approximately linear in an operating point. It is not intended for processes with heavy nonlinearities as sometimes encountered in the process industry. On the other hand, it is certainly possible to extend the identification procedure in order to include structured nonlinearities as well.

Chapter 9

Application to a Glass Tube Manufacturing Process

9.1 Introduction

In this chapter an application to a multivariable industrial process, a glass tube manufacturing process, is discussed. The objective is to design a controller for the process in such a way that, before actual implementation of the controller, guarantees can be given about the closed loop performance. In particular, in an industrial environment it is very desirable that prior guarantees of closed loop stability can be given. This objective is realized by collecting measurement data of the process, identifying a nominal model and a bound on the model error, and designing a robust controller.

The process imposes the following requirements on the identification procedure:

- The identification method has to be MIMO applicable, as the process is multivariable.
- The identification method has to be able to handle arbitrary input signals, as there is no complete freedom to do experiment design. During the identification experiments the production is lost, and therefore measurement time is very much restricted. In particular no sine-sweep experiments can be carried out.
- All prior information required has to be gathered from data.
- Because of the control-design step later, and the necessity to implement a low-order controller, it has to be possible to identify a reasonably low-order nominal model.

The identification procedure presented in Chapter 8 satisfies all of these requirements. Therefore this identification procedure is applied to the glass tube manufacturing process.

The identification procedure yields both a nominal model and a probabilistic frequency dependent upper bound on the model error. The nominal model is used to

design an H_∞-controller with the procedure described in Procedure 8.4.2 in Chapter 8. The reasons for using this control design procedure are:

- The controller performance and robustness can easily be tuned by the use of simple (constant, diagonal) input and output weighting matrices.
- Low-order controllers can be designed, which is important for the implementation.
- Fast and reliable software is available to calculate a discrete-time controller.

The model error bound is used in a μ-analysis to prove robust stability of the closed loop, before actual implementation of the controller.

The glass tube manufacturing process has previously been considered in Backx and Damen (1992), Falkus *et al.* (1993), Murad *et al.* (1993) and Van Overschee and De Moor (1993). However, in these papers no model error bounds have been identified, and consequently no robust controllers could be designed, which account for the identification model error.

The outline of the chapter is as follows. In Section 9.2 a description is given of the glass tube manufacturing process. In Section 9.3 the identification results are presented. Next, in Section 9.4 the control design is described and the resulting controller is evaluated. In Section 9.5 a robust stability analysis is performed. Finally, in Section 9.6 the results are discussed.

In this chapter application of the controller to the process is not described, as the actual controller implementation has not yet taken place.

9.2 Process Description

The industrial process under consideration is a glass tube manufacturing process, schematically depicted in Figure 9.1. By direct electric heating, quartz sand is melted and flows down through a ring-shaped hole along the accurately positioned mandrel. Under pressure, gas is led through the hollow mandrel. The glass tube is pulled down due to gravity and supported by a drawing machine.

Shaping of the tube takes place at, and just below the end of the mandrel. The longitudinal shape of the tube is characterized by two important dimensions, which are taken as outputs to be controlled: tube diameter (first output) and tube wall-thickness (second output). Both outputs are influenced by many process conditions such as:

- mandrel gas pressure,
- drawing speed,
- power applied to the furnace (temperature of the glass),
- melting vessel pressure,
- composition of raw materials.

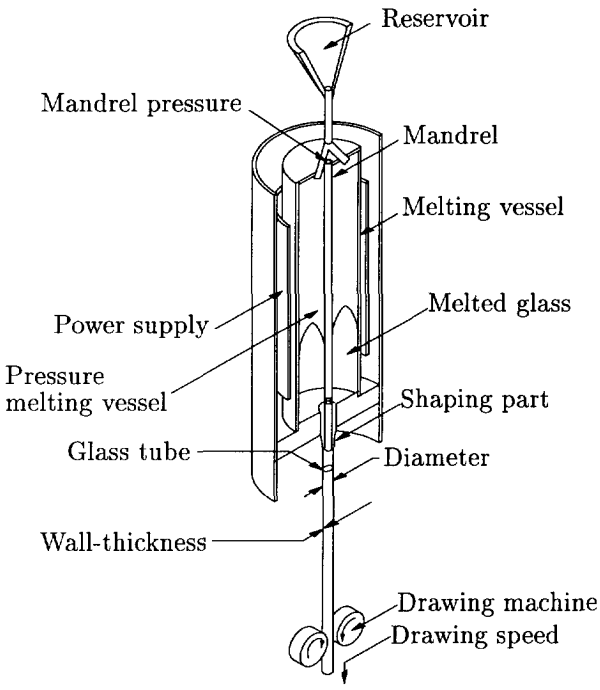

Fig. 9.1: Schematic overview of the glass tube manufacturing process (this picture has been created by Heinz Falkus).

Some of these have a small bandwidth (power and composition of raw materials), poorly influence the glass quality (compositon of raw materials), or have extremely large delay times involved (power, melting vessel pressure and composition of raw materials). Therefore these are not well suited for control of the tube dimensions.

The mandrel pressure and the drawing speed influence the shaping of the tube in a most direct way. Transfers from these inputs to both wall-thickness and diameter have the largest bandwidth, the shortest delay times and permit, to some extent, independent manipulation of the outputs. The permitted ranges of these two process inputs allow a control of the tube dimensions over the full amplitude range of output disturbances and enable the production of a large variety of different products. Consequently, these two are taken as controlling inputs. The drawing speed is denoted as the first input and the mandrel pressure as the second input.

Shaping of the glass tube clearly is a MIMO process with a high degree of interaction. Increase of the mandrel pressure results in an increase of the tube diameter and a decrease of the tube wall-thickness. Increase of the drawing speed causes a decrease of both diameter and wall-thickness. A physical model of this shaping part can be obtained by deriving the physical laws of the shaping process, describing the shaping

of the tube in detail and over the full range of possible operating points, determined by various values of tube diameter and wall-thickness. However, this physical model is very complex and has physical parameters included with numerical values that are unknown for the different operating points. Therefore modelling is performed by means of black-box identification.

Basically the process is nonlinear. However, one operating point is considered. Stair-case experiments, i.e. experiments with the inputs being excited by steps of different amplitudes, indicate that the process can very well be considered linear and time-invariant in the operating point. Therefore there is no problem expected with the use of linear identification and control design techniques.

In the sequel only scaled data are used, such that the original process data cannot be retrieved. This is because of the industrial confidentiality required.

9.3 Identification of a Nominal Model and a Model Error Bound

The problem of identifying the glass tube production process is a MIMO identification problem. To reduce numerical complexity this MIMO problem is split into two MISO identification problems. Models are estimated for each output separately and at the end these are combined into one MIMO model.

For identification purposes the following (open loop) experiments have been carried out on the glass tube production process:

- A free-run experiment, i.e. output measurements without input excitation. This gives an indication of the output noise and can be used to estimate the second order noise statistics, knowledge of which is required for the uncertainty bounding identification procedure, see Chapter 5.
- Stepresponse experiments. These can be used to accurately identify the low-frequent system dynamics.
- A PRBS experiment, i.e. the input signals are Pseudo Random Binary Sequences, and independent of each other. This type of input signal excites the system uniformly in the entire frequency range of interest.
- An experiment where the input signals are filtered PRBS signals. The filter is a bandpass filter with high gain in the medium frequency range, which is considered important for the control application.
- One more unfiltered PRBS experiment, which is not used for identification, but for validation purposes.

The data preprocessing consists of the following steps:

- Peakshaving. Outliers in the data (due to sensor failures) are removed.
- Detrending. Trends and offsets in the data are removed.

- Decimation. The data have been collected at a very fast sampling rate (to reduce aliasing effects as much as possible). Now the sampling rate is reduced to a proper value, which is identical to the sampling rate for the discrete-time controller.

The process contains large time delays, partly due to the physical time it takes before the glass tube reaches the measurement equipment. First, these time delays are estimated and removed from the data. Due to the fact that MISO problems are considered, it is always possible to do this by shifting both input signals over the proper amount of samples. A correlation analysis (Söderström and Stoica, 1989, pp. 42,43) based on the unfiltered PRBS experiments is applied to estimate the time delays. In Table 9.1 the estimated delay times are given.

	input 1	input 2
output 1	5	7
output 2	10	12

Table 9.1: Time delays in number of samples.

Next, the identification procedure of Chapter 8 has been carried out. By means of the iteration of frequency response uncertainty bounding and frequency response curve fitting (procedure 8.2.2) a suitable set of basis functions has been determined. In each iteration step more high-frequent dynamics have become apparent and have been modelled accurately. Altoghether 3 iterations have been performed. The final nominal model $\widehat{G}(z)$ (without delays) is shown in the Bode plot of Figure 9.2.

The undelayed model has been used to generate basis functions for a final identification of frequency response uncertainty regions. Rectangular confidence regions have been derived for the PRBS and the filtered PRBS experiments separately. Next, the intersection of these confidence regions has been calculated. The resulting uncertainty regions are depicted in the Nyquist diagram of Figure 9.3, where the delays have been added again, together with the frequency response of the nominal model $\widehat{G}(z)$. The confidence regions correspond to a 99.8%-confidence level. From this Nyquist diagram it is straightforward to calculate upper and lower bounds on the amplitude and the phase of the system's frequency response. These bounds, corresponding to the undelayed model, are depicted in Figure 9.2 as well.

From Figure 9.3 a frequency dependent upper bound on the additive model error is easily derived as the worst-case distance from the nominal frequency response $\widehat{G}(e^{i\omega_j})$ to the rectangular uncertainty region for that frequency. In Figure 9.4 this upper bound on the model error is shown for each entry i', j', and corresponding to a confidence level of 99.8%.

The undelayed MISO model for the first output is of order 9. The undelayed model for the second output is also of order 9. The combined MIMO model with delays is of order 35. Due to similar dynamics for both outputs, there are redundant orders in this MIMO model. As a low-order model is desired for the control design application, this nominal model is reduced to order 15 with Hankel-norm model reduction

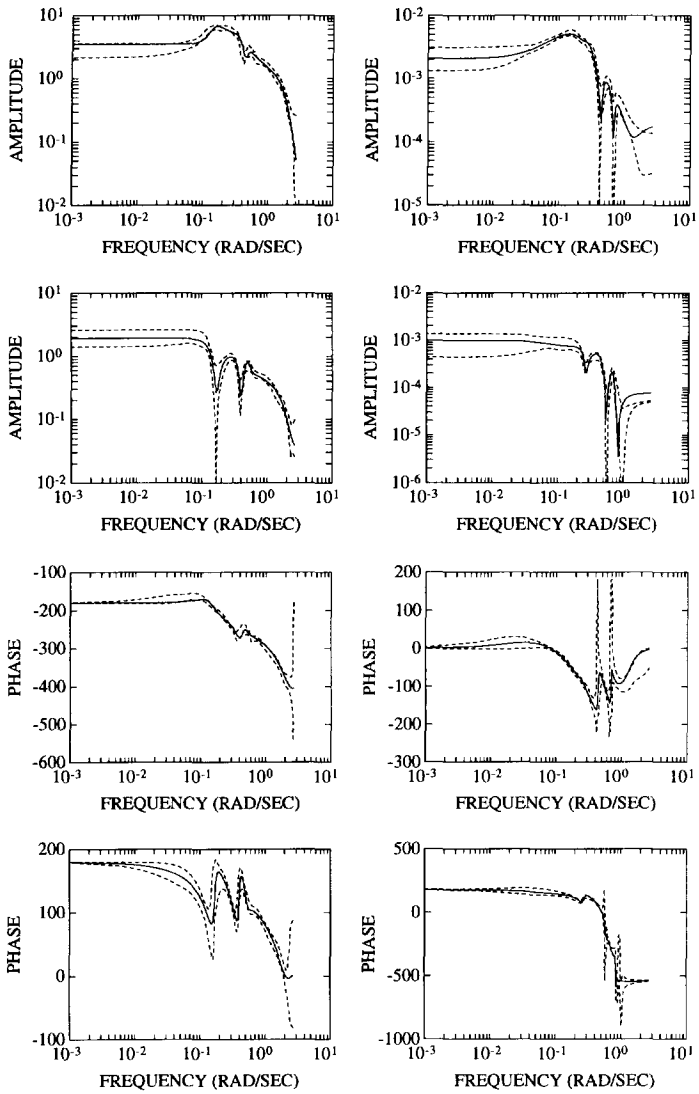

Fig. 9.2: Bode plot of the identified undelayed nominal model $\widehat{G}(z)$ (solid) and 99.8%-confidence intervals for the amplitude and phase (dashed). The entry in row i' and column j' corresponds to the transfer function from input $u_{j'}$ to output $y_{i'}$ (for both the amplitude and the phase part).

(Glover, 1984). This model reduction step has been carried out such that no significant dynamics are lost.

Notice from Figure 9.2 that also high-frequent dynamics present in the process

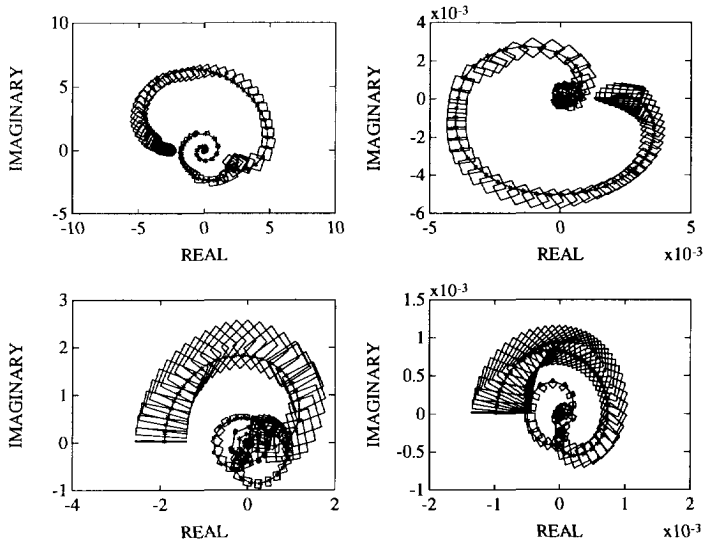

Fig. 9.3: Nyquist plot nominal model $\widehat{G}(z)$ (solid,$*$) and 99.8%-confidence regions (rectangulars) with the delays being added. The entry in row i' and column j' corresponds to the transfer function from $u_{j'}$ to $y_{i'}$.

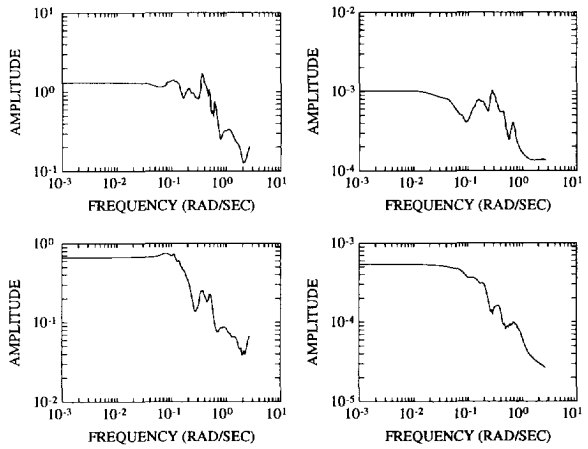

Fig. 9.4: Upper bound on the additive model error corresponding to a confidence level of 99.8%. The entry in row i' and column j' gives the additive model error bound for the transfer function from $u_{j'}$ to $y_{i'}$.

can be accurately modelled, this despite of the bad signal-to-noise-ratio. Other iden-

tification methods, such as standard prediction error and subspace methods, appear not capable of modelling so much detail in the high-frequency range. Compare, for example, Figure 9.2 with the results in Falkus *et al.* (1993) and Van Overschee and De Moor (1993), where the same data sets have been used for identification. If an attempt is made to identify high-order models with a prediction error method, then easily numerical problems in the nonlinear optimization occur because of the bad signal-to-noise-ratio. On the other hand, if a low-order model is identified, only the low-frequent dynamics are accurately modelled. Application of a high-pass filter in order to emphasize the high-frequent dynamics, again easily leads to numerical problems.

The time-domain behaviour of the nominal model $\widehat{G}(z)$ is evaluated. In Figure 9.5 the measured and simulated stepresponses are shown. In Figure 9.6 the measured and simulated responses to the PRBS inputs are shown, for the experiment used in the identification. Finally, in Figure 9.7 measured and simulated responses are shown for the validation PRBS experiment, which has not been used in the identification. The time-domain fit appears very good for all these three cases, which gives confidence in the nominal model.

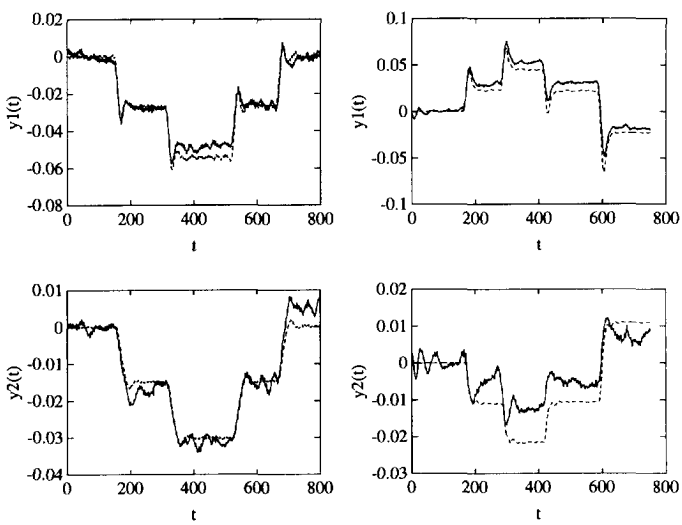

Fig. 9.5: Measured (solid) and simulated (dashed) stepresponses. The entry in row i' and column j' corresponds to the response of $y_{i'}$ to step-excitation of $u_{j'}$.

9.4 Robust Controller Design and Analysis

On the basis of the identified nominal model an H_∞-controller is designed with the procedure described in Subsection 8.4.2 in Chapter 8. When evaluating the controller performance the following items need to be considered.

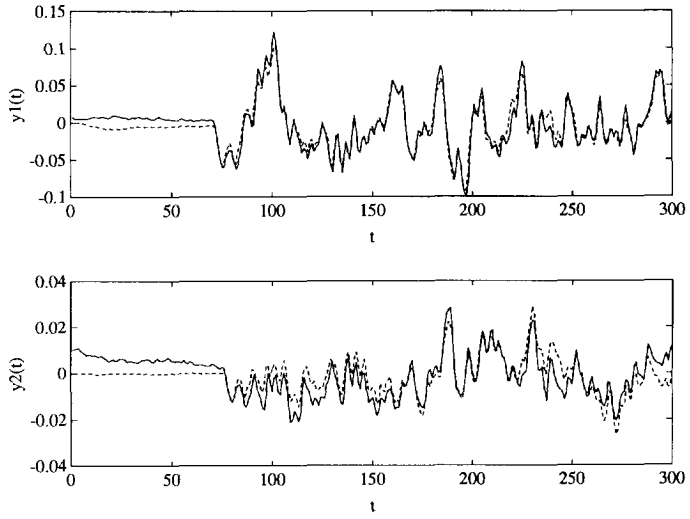

Fig. 9.6: Measured (solid) and simulated (dashed) responses on PRBS-excitation of both inputs for the experiment used in identification.

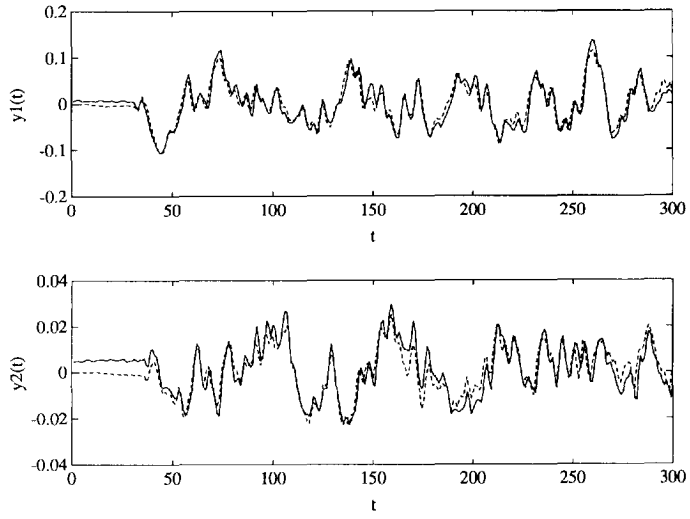

Fig. 9.7: Measured (solid) and simulated (dashed) responses on PRBS-excitation of both inputs for the validation experiment.

- The controller has to be robustly stable, both with respect to model error in one operating point and to changing dynamics for a slightly different operating

point.

- The controller has to track setpoint changes reasonably fast, with small overshoot and without static error.
- Low-frequent and very-low-frequent noise, such as trends, should be removed by the controller as much as possible.
- High-frequent noise should not be amplified too much.
- The input signals to the system should be as smooth as possible, without large overshoots on setpoint changes.
- Both outputs have to be statically and dynamically decoupled as much as possible.

The items above consitute a rather qualitative measure for controller performance. Engineering interpretation is required to actually construct a controller which meets these requirements.

In the H_∞-control design the following input scaling matrix has been used,

$$W_2(z) = \begin{bmatrix} \frac{\alpha z}{z-1} & 0 \\ 0 & \frac{\beta z}{z-1} \end{bmatrix},$$

and the output scaling matrix

$$W_1(z) = \begin{bmatrix} \gamma & 0 \\ 0 & 1 \end{bmatrix}.$$

The presence of integrators in the input scaling matrix provides integrating action for both outputs, which is necessary for both setpoint-tracking and low-frequent noise reduction. The scalars α, β and γ are used to tune the controller, such that acceptable performance and robustness are achieved, in accordance with the items listed above. The parameter γ is used to tune the relative importance of each output. The parameters α and β are used to tune the relative importance of each input and to tune the bandwidth of the closed loop system. A larger value for the mean of α and β generally means a higher bandwidth. Finally, the following values appeared to give a proper trade-off for all the requirements listed above,

$$\alpha = 0.014,\ \beta = 25,\ \gamma = 1.2.$$

In Figure 9.8 an amplitude Bode plot is shown of the designed controller $C(z)$, which is of order 12. In Figure 9.9 an amplitude Bode plot is given of the designed output sensitivity function. Clearly disturbance suppression is realized for frequencies up to 0.02 rad/sec.

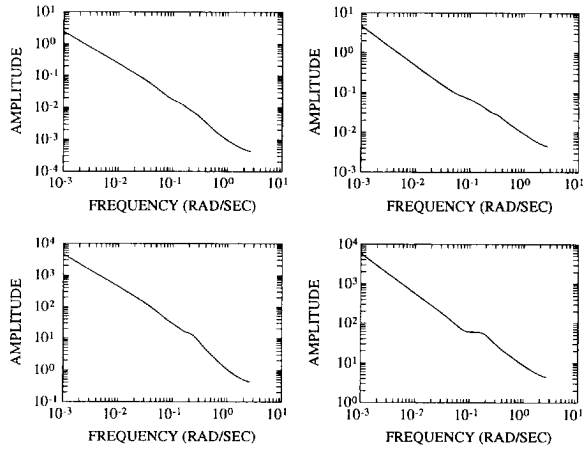

Fig. 9.8: Amplitude Bode plot of the designed controller $C(z)$. The entry in row i' and column j' corresponds to $C_{i'j'}(z)$.

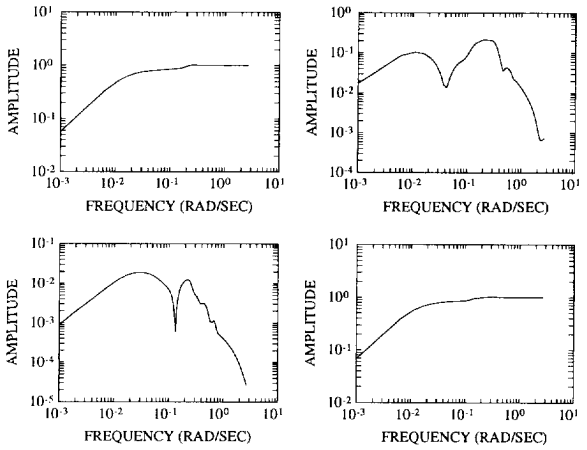

Fig. 9.9: Amplitude Bode plot of the output sensitivity function $S(z) = (I - \widehat{G}(z)C(z))^{-1}$. The entry in row i' and column j' corresponds to $S_{i'j'}(z)$.

9.5 μ-Stability Robustness Analysis

Before implementing the controller on the system, stability guarantees are required. Therefore a robust stability analysis is performed, using the frequency dependent model error bound. Consider the closed loop configuration depicted in Figure 9.10. For each frequency ω_j the weightings $W_{i'j'}(\omega_j)$, $i', j' = 1, 2$, are chosen equal to the worst-case

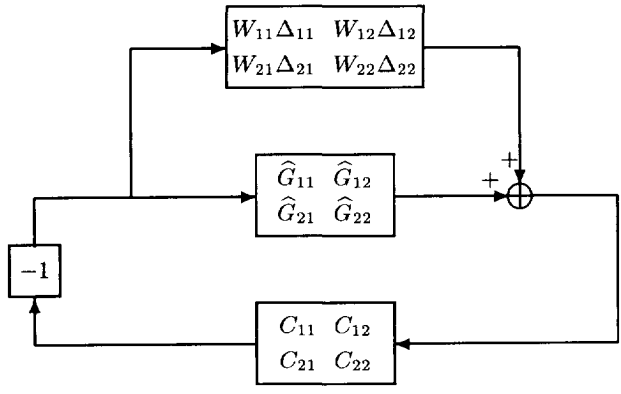

Fig. 9.10: Closed loop configuration with nominal model $\widehat{G}_{i'j'}$, controller $C_{i'j'}$ and weighted additive model error $W_{i'j'}\Delta_{i'j'}$.

error bounds depicted in Figure 9.4. Thus, the perturbations $\Delta_{i'j'}(\omega_j)$ are normalized to 1,

$$|\Delta_{i'j'}(\omega_j)| \leq 1,\ i' = 1,2,\ j' = 1,2,\ \forall \omega_j.$$

Straightforward manipulations show that the closed loop configuration of Figure 9.10 is stable if and only if the closed loop configuration of Figure 9.11 is stable. Here the

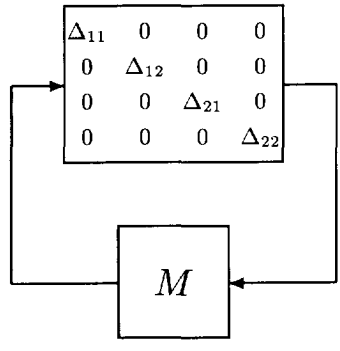

Fig. 9.11: Standard uncertainty configuration for μ-analysis.

matrix M is given by

$$M := \begin{bmatrix} F_{11}W_{11} & F_{11}W_{12} & F_{12}W_{21} & F_{12}W_{22} \\ F_{21}W_{11} & F_{21}W_{12} & F_{22}W_{21} & F_{22}W_{22} \\ F_{11}W_{11} & F_{11}W_{12} & F_{12}W_{21} & F_{12}W_{22} \\ F_{21}W_{11} & F_{21}W_{12} & F_{22}W_{21} & F_{22}W_{22} \end{bmatrix},$$

where $F_{i'j'}$, $i', j' = 1, 2$, is defined by

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} := -[I + C\widehat{G}]^{-1}C.$$

The closed loop configuration of Figure 9.11 is the standard one used in μ-analysis, see Subsection 2.2.3 in Chapter 2. As the $\Delta_{i'j'}(\omega_j)$ have been normalized to 1, the closed loop of Figure 9.11, and thus of Figure 9.10, is stable if and only if the structured singular value of M is smaller than 1,

$$\mu(M(\omega)) < 1, \ \forall \omega \in [0, \pi].$$

For simplicity this requirement is replaced by the requirement,

$$\mu(M(\omega_j)) < 1, \ j = 1, \ldots, l,$$

assuming that the frequency grid $\{\omega_j\}_{j=1,\ldots,l}$ has been chosen suitably, such that the intersample frequency behaviour causes no problems. This approximation is also supported by Packard and Pandey (1993), where it is shown that the structured singular value is a continuous function of frequency in case complex-valued perturbations are considered.

For each frequency ω_j, $j = 1, \ldots, l$, separately, $\mu(M(\omega_j))$ has been calculated. In Figure 9.12 the resulting μ-curve is shown. It appears that for all frequencies $\mu(M(\omega_j))$ is smaller than 1. Hence, robust stability is concluded and it is considered safe to actually implement the controller, as there is no danger of closed loop instability in the specific operating point of the process.

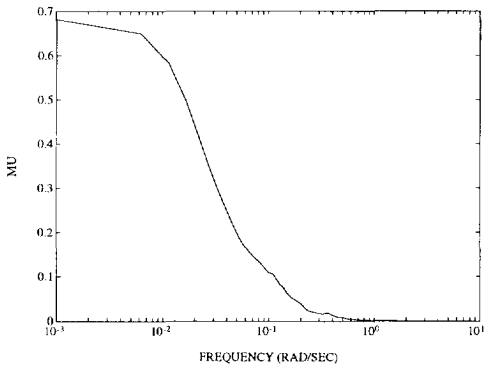

Fig. 9.12: $\mu(M(\omega_j))$ for $j = 1, \ldots, l$.

9.6 Discussion

The identification procedure proposed in Chapter 8 has been applied to measurement data of an industrial glass tube manufacturing process. It appears possible to iden-

tify an accurate nominal model with this identification method. Moreover, reliable frequency response error bounds can be identified, which are not overly conservative. An H_∞-controller has been designed for the nominal model, yielding an acceptable nominal performance. With the model error bounds robust stability has been proven by means of a μ-analysis, before actual implementation of the controller. In the near future the controller will be implemented on the process.

In this chapter is has been shown that the identification procedure developed in this thesis can very well be applied to industrial processes. In fact the procedure can be applied to any system or process, provided the nonlinearities are relatively small. In the case study considered in this chapter this has been ensured by excitation around a fixed operating point. For a heavily nonlinear process the identification can be carried out for a number of different operating points. This leads to a set of linear models and model error bounds, which describe the entire operation range of the nonlinear process. In fact this has been done for the glass tube production process as well, but due to space limitations the results for the other operating points are not reported here.

Epilogue

Chapter 10

Conclusions and Perspectives

10.1 Contributions of the Thesis

The main contribution of this thesis is the development of a robust-control-oriented identification procedure. The identification procedure delivers a nominal model and an explicit model error bound suited for use in high-performance robust process control. The basic steps of the procedure have been outlined in Figure 2.4 in Chapter 2. After the development of new identification tools in the Chapters 3 till 7, a more detailed description of the entire identification procedure has been given in Chapter 8.

An essential step in the identification procedure is the identification of a *system uncertainty set*. This set specifies explicit bounds on the frequency response or weighted pulse response of the system to be identified. The system uncertainty set is constructed on the basis of measurement data and a priori information about the system and the noise which corrupts the data. On the basis of this set both the nominal model and the model error bound are constructed.

In view of industrial applicability an important property of the identification procedure is the fact that it can handle arbitrary input signals and multivariable systems. Also, the resulting nominal model and model error bound are in a format suited for use in modern robust control analysis and design procedures. Moreover, the resulting model and model error bound can be used to design a *high-performing* controller, due to the fact that the model error bound is not unnecessarily conservative and can be tuned towards the control objective. This tuning takes place by means of the following design variables:

- *The identification experiment.* The input signal spectrum determines the size and shape of the system uncertainty set, which corresponds to the *variance* error.
- *The model order and the nominal identification criterion.* The order of the nominal model and the weighting function in the identification criterion determine the size and shape of the *bias* error.
- *The uncertainty structure.* The weighting in the uncertainty structure determines the size and shape of the *upper bound on the sum of variance and bias error.*

This clear distinction between the various error sources greatly facilitates the task of appropriately choosing the design variables in the identification procedure.

For intended use within the framework of the robust-control-oriented identification procedure the following identification tools have been developed:

- A procedure for deterministic uncertainty bounding identification, yielding frequency and pulse response uncertainty regions on the basis of time-domain data and deterministic prior assumptions about the system and the noise (Chapter 4).
- A procedure for probabilistic uncertainty bounding identification, yielding frequency and pulse response confidence regions on the basis of time-domain data, deterministic prior assumptions about the system and stochastic assumptions about the noise (Chapter 5).
- A procedure which solves the problem of identifying an H_∞-optimal nominal model in very good approximation (Chapter 6 and 7).
- A procedure which solves the problem of identifying an ℓ_1-optimal nominal model in very good approximation (Chapter 7).

The procedure for deterministic uncertainty bounding identification has been shown to be optimal in the sense that the smallest possible uncertainty regions are identified on the basis of the given deterministic prior information. Among other things this implies that it is impossible that any other identification procedure yields smaller system uncertainty regions on the basis of the same amount of data and prior information. Nevertheless, the procedure for probabilistic uncertainty bounding identification appears to yield much smaller system uncertainty regions, as shown in the examples in Chapter 5. This is due to the fact that in the probabilistic approach the random nature of the noise is much better taken into account. In particular optimal use is made of the fact that the noise is independent of the input signal or some external reference signal. For practical applications the use of the method for probabilistic uncertainty bounding identification is recommended, provided the noise can (in good approximation) be described as a realization of a stationary stochastic process.

On the basis of the identified system uncertainty set it is possible to construct bounds on the H_∞- or ℓ_1-norm of the weighted additive model error. On the basis of various considerations in Chapter 8 the H_∞-approach to the identification problem has been preferred to the ℓ_1-approach. In fact the recommended identification procedure is Procedure 8.2.1 in Chapter 8. First, frequency response system confidence regions are identified with the procedure described in Chapter 5 and, next, a nominal model is identified with the curve fit procedure developed in Chapter 6. Finally, a frequency dependent upper bound on the model error, or an upper bound on the H_∞-norm of the weighted model error, is established.

The new identification procedure has succesfully been applied to a multivariable industrial glass tube manufacturing process (Chapter 9). The application of the identification procedure in combination with an H_∞-optimal robust control design procedure has made it possible to design a controller with a guaranteed performance for the process. Before implementation of the controller exact statements can be made

about closed loop stability and performance that the controller will achieve. The identification procedure is a very powerful modelling procedure for processes which are approximately linear in an operating point.

As a consequence of the development of new identification tools the following byproducts have resulted:

- *A consistent parameter bounding identification procedure,* in which use is made of alternative types of noise constraints, such as cross-covariance or frequency domain noise constraints (Chapter 3). It has been shown that the new types of noise constraints outperform the classical time-domain amplitude noise constraint.
- *A frequency response curve fit procedure minimizing a maximum amplitude criterion and with guaranteed stability of the model,* both for discrete-time and continuous-time models, SISO and MIMO (Chapter 6). According to the analysis in Section 7.2 the curve fit procedure provides a solution for:
  - The problem of *weighted H_∞-optimal model reduction* for SISO systems.
  - The problem of *identification in H_∞.*
- *An algorithm for weighted ℓ_1-optimal model reduction* for a linearly parametrized reduced-order model set (Section 7.3).

10.2 Recommendations for Future Research

Within the framework of the identification procedure developed in this thesis the following topics deserve further research attention:

- *Experiment design.* What is an optimal choice of the input signal spectrum in view of the control application?
- *Model order selection.* What model complexity is needed in view of the model application?
- *Selection of the nominal identification criterion.* In particular the curve fit procedure developed in Chapter 6 can be extended to minimize a feedback-relevant criterion of the format (2.9) in Chapter 2.
- *Selection of the uncertainty structure.* In particular uncertainty structures other than the weighted additive uncertainty structure in (2.10) may appear useful in the context of feedback-relevant identification.

Moreover, the identification procedure can be extended in order to be applicable to:

- *Nonlinear or time-varying systems.* In particular it is straightforward to include nonlinearities in the uncertainty bounding identification procedure, provided these are of a pre-specified structure and parametrized linearly. The idea is to identify parameter bounds for the nonlinear part and frequency response bounds for the linear part. An important issue then is the estimation of the structure of the nonlinear part of the system to be identified.

- *On-line identification.* It seems possible to construct an on-line version of the identification procedure, provided simplifications are introduced. It needs to be investigated which simplifications are needed in order to obtain a computationally feasible on-line procedure which is still reasonably accurate.

These extensions will make the identification procedure even more tailored towards industrial applications. Next, in the area of robust control the following issues are worthwhile to be explored:

- *Identification-based control.* In particular this concerns the development of a robust control design procedure which takes into account the fact that the model error bounds delivered by the identification are correct with a certain probability. The controller can, for example, be designed such that robust performance is achieved with probability 99% and robust stability with probability 99.999%.

- *Adaptive robust control.* This concerns the integration of on-line uncertainty bounding identification and on-line robust control design.

Moreover, it is strongly recommended that the problem of *identification in* H_∞, as currently investigated in literature, is reformulated. From a theoretical point of view, interesting developments take place in this research area. However, from an engineering point of view, this identification problem is not too interesting, as the tools, which are developed in this area, have a very restricted applicability. A justification for this statement can be found at the end of Subsection 2.4.5. The developments and results in, for example, Chapter 2 and 7 of this thesis certainly indicate a possible direction for future research in this area.

Similarly, it is worthwile to reconsider the standard *parameter bounding* or *bounded error* identification problem. At the end of Subsection 2.4.4 the statement has been made, and motivated, that the tools developed in this area are not applicable to robust control problems. The developments in Chapter 3 of this thesis indicate new lines, along which the bounded error identification may develop, in order to yield industrially applicable identification algorithms. It seems impossible to come up with realistic, i.e. not overly conservative, error bounds, unless, in one way or another, stochastic or averaging properties of the noise are used in the identification.

Addenda

Bibliography

Akçay, H., G. Gu and P.P. Khargonekar (1993). A class of algorithms for identification in H_∞: continuous time case. *IEEE Trans. Autom. Contr.,* Vol. AC-38, pp. 289–294.

Al-Saggaf, U.M. and G.F. Franklin (1988). Model reduction via balanced realizations: an extension and frequency weighting techniques. *IEEE Trans. Autom. Contr.,* Vol. AC-33, pp. 687–692.

Andersen, H.W., K.H. Rasmussen and S.B. Jørgensen (1991). Advances in process identification. *Chemical Process Control, Proc. 1991 Engineering Foundations Conf.,* Am. Inst. of Chem. Eng., New York, pp. 237–269.

Backx, T.C.P.M. and A.A.H. Damen (1992). Identification for control of MIMO industrial processes. *IEEE Trans. Autom. Contr.,* Vol. AC-37, pp. 980–986.

Bai, E. (1991). Adaptive quantification of model uncertainties by rational approximation. *IEEE Trans. Autom. Contr.,* Vol. AC-36, pp. 441–453.

Bai, E. and S. Raman (1993). A linear interpolatory algorithm for robust system identification with corrupted measurement data. *IEEE Trans. Autom. Contr.,* Vol. AC-38, pp. 1236–1241.

Balakrishnan, V. and S. Boyd (1992). On computing the worst-case peak gain of linear systems. *Proc. 31st IEEE Conf. Dec. and Contr.,* Tucson, pp. 2191–2192.

Balas, G.J. and J.C. Doyle (1990). Identification of flexible structures for robust control. *IEEE Contr. Syst. Mag.,* pp. 51–58.

Baratchart, L. and J. Leblond (1993). On robust identification from partial frequency data. *Proc. 2nd European Contr. Conf.,* Groningen, The Netherlands, pp. 2166–2168.

Bayard, D.S. (1992). Statistical plant set estimation using Schroeder-phased multisinusoidal input design. *Proc. Am. Contr. Conf.,* Chicago, pp. 2988–2995.

Bayard, D.S. (1994). High-order multivariable transfer function curve fitting: algorithms, sparse matrix methods and experimental results. *Automatica,* Vol. 30, pp. 1439–1444.

Bayard, D.S. and Y. Yam (1993). Frequency domain identification for robust control design. In *The Modeling of Uncertainty in Control Systems,* Proc. 1992 Santa Barbara Workshop, Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, London, Vol. 192, pp. 303–335.

Bayard, D.S., Y. Yam and E. Mettler (1992). A criterion for joint optimization of identification and robust control. *IEEE Trans. Autom. Contr.,* Vol. AC-37, pp. 986–991.

Belforte, G., B. Bona and V. Cerone (1990). Parameter estimation algorithms for a set-membership description of uncertainty. *Automatica,* Vol. 26, pp. 887–898.

Bitmead, R.R. (1993). Iterative control design approaches. *Prepr. 12th IFAC World Congress,* Sydney, Australia, Vol. 9, pp. 381–384.

Bitmead R.R., M. Gevers and V. Wertz (1990). *Adaptive Optimal Control, The Thinking Man's GPC,* Prentice-Hall, Englewood Cliffs, N.J.

Bode, H.W. (1945). *Network Analysis and Feedback Amplifier Design,* Van Nostrand.

Bongers, P.M.M. (1994). *Modeling and Identification of Flexible Wind Turbines and a Factorizational Approach to Robust Control,* Ph.D. Thesis, Mech. Eng. Syst. and Contr. Group, Delft Univ. of Techn., The Netherlands.

Bongers, P.M.M. and O.H. Bosgra (1990). Low order H_∞ controller synthesis. *Proc. 29th IEEE Conf. Dec. and Contr.,* Honolulu, pp. 194–199.

Boyd, S. and V. Balakrishnan (1990). A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm. *Syst. and Contr. Letters,* Vol. 15, pp. 1–7.

Boyd, S.P. and C.H. Barratt (1991). *Linear Controller Design: Limits of Performance,* Prentice-Hall, Englewood Cliffs, N.J.

Boyd, S.P. and J.C. Doyle (1987). Comparison of peak and RMS gains for discrete-time systems. *Syst. and Contr. Lett.,* Vol. 9, pp. 1–6.

Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory,* Holden-Day, San Francisco.

Caines, P.E. and M. Baykal-Gürsoy (1989). On the L_∞-consistency of L_2 estimators. *Syst. and Contr. Lett,* Vol. 12, pp. 71–76.

Chen, C.T. (1987). *Linear Control System Design and Analysis,* Holt, Rinehart and Winston.

Chen, J., C.N. Nett and M.K.H. Fan (1992). Optimal non-parametric system identification from arbitrary corrupt finite time-series: a control oriented approach. *Proc. Am. Contr. Conf.,* Chicago, pp. 279–285.

Childers, D.G., ed. (1978). *Modern Spectrum Analysis.* IEEE Press, New York.

Clement, T. and S. Gentil (1988). Reformulation of parameter identification with unknown-but-bounded errors. *Math. and Comp. in Sim.,* Vol. 30, pp. 257–270.

Dahleh, M.A. and J.B. Pearson (1987). ℓ_1-Optimal feedback controllers for MIMO discrete-time systems. *IEEE Trans. Autom. Contr.,* Vol. AC-32, pp. 314–322.

Dahleh, M.A. and J.B. Pearson (1988). Optimal rejection of persistent disturbances, robust stability and mixed sensitivity minimization. *IEEE Trans. Autom. Contr.,* Vol. AC-33, pp. 722-731.

Dahleh, M.A. and M.H. Khammash (1993). Controller design for plants with structured uncertainty. *Automatica,* Vol. 29, pp. 37–56.

Dailey, R.L. and M.S. Lukich (1988). Recent results in identification and control of a flexible truss structure. *Proc. Am. Contr. Conf.,* Atlanta, pp. 1468–1473.

De Moor, B. (1988). *Mathematical Concepts and Techniques for Modelling of Static and Dynamic Systems,* Ph.D. Thesis, Dept. Electrical Eng., Catholic University Leuven, Belgium.

Desoer, C.A. and M. Vidyasagar (1975). *Feedback Systems: Input-Output Properties,* Academic Press, New York.

De Vries, D.K. (1994). *Identification of Model Uncertainty for Control Design,* Ph.D. Thesis, Mech. Eng. Syst. and Contr. Group, Delft Univ. of Techn., The Netherlands.

De Vries, D.K. and P.M.J. Van den Hof (1992). Quantification of model uncertainty from data: input design, interpolation and connection with robust control design specifications. *Proc. Am. Contr. Conf.*, Chicago, pp. 3170–3175.

De Vries, D.K. and P.M.J. Van den Hof (1993). Quantification of uncertainty in transfer function estimation: a mixed deterministic-probabilistic approach. *Prepr. 12th IFAC World Congress,* Sydney, Australia, Vol. 8, pp. 157–160.

De Vries, D.K. and P.M.J. Van den Hof (1994a). Quantification of model uncertainty from data. *Int. J. of Robust and Nonlinear Control,* Vol. 4, pp. 301–319.

De Vries, D.K. and P.M.J. Van den Hof (1994b). Quantification of uncertainty in transfer function estimation: a mixed probabilistic-worst case approach. To appear in *Automatica,* Vol. 31, no. 3.

Doyle, J.C. (1982). Analysis of feedback systems with structured uncertainties. *IEE Proc.,* Vol. 129, pp. 242–250.

Doyle, J.C., K. Glover, P.P. Khargonekar and B.A. Francis (1989). State space solutions to H_2 and H_∞ control problems. *IEEE Trans. Autom. Contr.,* Vol. AC-34, pp. 831–847.

Falkus, H., A.A.H. Damen and A.C.P.M. Backx (1993). Identification of a tube glass production process: point vs. set estimation. *Proc. 2nd European Contr. Conf.,* Groningen, The Netherlands, pp. 2344–2349.

Fogel, E. (1979). System identification via membership set constraints with energy constrained noise. *IEEE Trans. Autom. Contr.,* Vol. AC-24, pp. 752–758.

Fogel, E. and Y.F. Huang (1982). On the value of information in system identification — bounded noise case. *Automatica,* Vol. 18, pp. 229–238.

Francis, B.A. (1987). *A course in H_∞ Control Theory,* Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, Vol. 88.

Franklin, G.F., J.D. Powell and A. Emami-Naeni (1986). *Feedback Control of Dynamic Systems,* Addison-Wesley.

Freedman, B.G. (1977). State of the art and needs in the industrial application of process identification and adaptive control. *Chemical Process Control, Proc. 1976 Engineering Foundations Conf.,* Am. Inst. of Chem. Eng., New York, pp. 206–218.

Gevers, M. (1993). Towards a joint design of identification and control? *Essays on Control: Perspectives in the Theory and its Applications,* Birkhäuser, Boston, pp. 111–151.

Gevers, M. and L. Ljung (1986). Optimal experiment designs with respect to the intended model application. *Automatica,* Vol. 22, pp. 543–554.

Gill, P.E., W. Murray and M.H. Wright (1981). *Practical Optimization,* Academic Press, London.

Glover, K. (1984). All optimal Hankel-norm approximations of linear multivariable systems and their L_∞-error bounds. *Int. J. of Contr.,* Vol. 39, no. 6, pp. 1115–1193.

Goodwin, G.C. and R.L. Payne (1977). Dynamic system identification: experiment design and data analysis. *Math. in Science and Eng.,* Academic Press, Vol. 136, pp. 124–207.

Goodwin, G.C., M. Gevers and B.M. Ninness (1990). Optimal model order selection and estimation of model uncertainty for identification with finite data. *Proc. 29th IEEE Conf. Dec. and Contr.*, Honolulu, pp. 285–290.

Gu, G. and P.P. Khargonekar (1992a). A class of algorithms for identification in H_∞. *Automatica*, Vol. 28, pp. 299–312.

Gu, G. and P.P. Khargonekar (1992b). Linear and nonlinear algorithms for identification in H_∞ with error bounds. *IEEE Trans. Autom. Contr.*, Vol. AC-37, pp. 953–963.

Gu, G., P.P. Khargonekar and Y. Li (1992). Robust convergence of two-stage nonlinear algorithms for identification in H_∞. *Syst. and Contr. Lett.*, Vol. 18, pp. 253–263.

Gustafsson, T.K. and P.M. Mäkilä (1993). Modelling of uncertain systems via linear programming. *Prepr. 12th IFAC World Congress*, Sydney, Australia, Vol. 5, pp. 293–298.

Gustavsson, I. (1975). Survey of applications of identification in chemical and physical processes. *Automatica*, Vol. 11, pp. 3–24.

Hakvoort, R.G. (1990). Optimal experiment design for prediction error identification in view of feedback design. *Sel. Topics in Id., Mod. and Contr.*, Vol. 2, Delft Univ. Press, pp. 71–78.

Hakvoort, R.G. (1991). Identification of an upper bound for the ℓ_1-norm of the model uncertainty. *Sel. Topics in Id., Mod. and Contr.*, Vol. 3, Delft Univ. Press, pp. 51–58.

Hakvoort, R.G. (1992a). Approximate identification for feedback design. *Journal A*, Vol. 33, no. 2, pp. 30–36.

Hakvoort, R.G. (1992b). Worst-case system identification in ℓ_1: error bounds, optimal models and model reduction. *Proc. 31st IEEE Conf. Dec. and Contr.*, Tucson, pp. 499–504.

Hakvoort, R.G. (1993a). Frequency domain curve fitting with maximum amplitude criterion and guaranteed stability. *Proc. 2nd European Contr. Conf.*, Groningen, The Netherlands, pp. 252–257.

Hakvoort, R.G. (1993b). Worst-case system identification in H_∞: error bounds and optimal models. *Prepr. 12th IFAC World Congress*, Sydney, Australia, Vol. 8, pp. 161–164.

Hakvoort, R.G. (1994). A linear programming approach to the identification of frequency domain error bounds. *Prepr. SYSID'94, 10th IFAC Symp. on Syst. Id.*, Copenhagen, Denmark, Vol. 2, pp. 195–200.

Hakvoort, R.G., R.J.P. Schrama and P.M.J. Van den Hof (1992). Approximate identification in view of LQG feedback design. *Proc. Am. Contr. Conf.*, Chicago, pp. 2824–2828.

Hakvoort, R.G., R.J.P. Schrama and P.M.J. Van den Hof (1994). Approximate identification with closed loop performance criterion and application to LQG feedback design. *Automatica*, Vol. 30, pp. 679–690.

Hakvoort, R.G. and P.M.J. Van den Hof (1993). Identification of model error bounds in ℓ_1- and H_∞-norm. In *The Modeling of Uncertainty in Control Systems*, Proc. 1992 Santa Barbara Workshop, Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, London, Vol. 192, pp. 139–155.

Hakvoort, R.G. and P.M.J. Van den Hof (1994a). Frequency domain curve fitting with maximum amplitude criterion and guaranteed stability. To appear in *Int. J. of Contr.*

Hakvoort, R.G. and P.M.J. Van den Hof (1994b). An instrumental variable procedure for the identification of probabilistic frequency response uncertainty regions. To appear in *Proc. 33rd IEEE Conf. Dec. and Contr.*, Lake Buena Vista.

Hakvoort, R.G., P.M.J. Van den Hof and O.H. Bosgra (1993). Consistent parameter bounding identification using cross-covariance constraints on the noise. *Proc. 32nd IEEE Conf. Dec. and Contr.*, San Antonio, pp. 2601–2606.

Hansen, F., G.F. Franklin and R.L. Kosut (1989). Closed loop identification via the fractional representation: experiment design. *Proc. Am. Contr. Conf.*, Pittsburgh, pp. 1422–1427.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1991). Control oriented system identification: a worst-case/deterministic approach in H_∞. *IEEE Trans. Autom. Contr.*, Vol. AC-36, pp. 1163–1176.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1992). Worst-case/deterministic identification in H_∞: the continous-time case. *IEEE Trans. Autom. Contr.*, Vol. AC-37, pp. 604–610.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1993). Least squares methods for H_∞ control oriented system identification. *IEEE Trans. Autom. Contr.*, Vol. AC-38, pp. 819–826..

Heuberger, P.S.C. (1990). *On Approximate System Identification with System-Based Orthonormal Functions,* Ph.D. Thesis, Mech. Eng. Syst. and Contr. Group, Delft Univ. of Techn., The Netherlands.

Heuberger, P.S.C. and O.H. Bosgra (1990). Approximate system identification using system based orthonormal functions. *Proc. 29th IEEE Conf. Dec. and Contr.*, Honolulu, pp. 1086–1092.

Heuberger, P.S.C., P.M.J. Van den Hof and O.H. Bosgra (1994). A generalized orthonormal basis for linear dynamical systems. To appear in *IEEE Trans. Autom. Contr.*, Vol. AC-40, no. 3. Abbreviated version in *Proc. 32nd IEEE Conf. Dec. and Contr.*, San Antonio, 1993, pp. 2850–2855.

Heunis, A.J. (1988). Asymptotic properties of prediction error methods in approximate system identification. *Stochastics,* Vol. 24, pp. 1–43.

Hjalmarsson, H. (1993a). *Aspects on Incomplete Modelling in System Identification,* Ph.D. Thesis, Dept. Electr. Eng. Linköping Univ., Sweden.

Hjalmarsson, H. (1993b). A model variance estimator. *Prepr. 12th IFAC World Congress,* Sydney, Australia, Vol. 9, pp. 5–10.

Hjalmarsson, H. and L. Ljung (1992). Estimating model variance in the case of undermodelling. *IEEE Trans. Autom. Contr.*, Vol. AC-37, pp. 1004-1008.

Horowitz, I. (1979). Quantitive synthesis of uncertain multiple input-output feedback systems. *Int. J. of Contr.*, Vol. 30, pp. 81–106.

Horowitz, I. (1982). Quantitive feedback theory. *Proc. Inst. of Electr. Eng.*, Part D, Vol. 129, pp. 215–226.

Jacobs, O.L. (1974). *Introduction to Control Theory,* Oxford.

Jacobson, E.W., P. Lundström and S. Skogestad (1991). Modelling and identification for robust control of ill-conditioned plants — a distillation case study. *Proc. Am. Contr. Conf.*, Boston, pp. 242–248.

Jacobson, C.A., C.N. Nett and J.R. Partington (1992). Worst case system identification in ℓ_1: optimal algorithms and error bounds. *Syst. and Contr. Lett.*, Vol. 19, pp. 419–424.

Karloff, H. (1991). *Linear Programming,* Birkhäuser, Boston.

Kavranoğlu, D. (1993). Computation of the solution for the H_∞ model reduction problem. *Proc. Am. Contr. Conf.*, San Francisco, pp. 2190–2194.

Kavranoğlu, D. and M. Bettayeb (1993). Characterization of the solution to the optimal H_∞ model reduction problem. *Syst. and Contr. Lett.*, Vol. 20, pp. 99–107.

Khammash, M. and J.B. Pearson (1991). Performance robustness of discrete-time systems with structured uncertainty. *IEEE Trans. Autom. Contr.*, Vol. AC-36, pp. 398–412.

Khammash, M. and J.B. Pearson (1993). Analysis and design for robust performance with structured uncertainty. *Syst. and Contr. Lett.*, Vol. 20, pp. 179–187.

Kosut, R.L., M. Lau and S. Boyd (1992). Set-membership identification of systems with parametric and nonparametric uncertainty. *IEEE Trans. Autom. Contr.*, Vol. AC-37, pp. 929–941.

Kwakernaak, H. (1993). Robust control and H_∞-optimization — tutorial paper. *Automatica,* Vol. 29, pp. 255–273.

Lamaire, R.O., L. Valavani, M. Athans and G. Stein (1991). A frequency domain estimator for use in adaptive control systems. *Automatica,* Vol. 27, pp. 23–38.

Larimore, W.E. (1993). Accuracy confidence bands including the bias of model underfitting. In *The Modeling of Uncertainty in Control Systems,* Proc. 1992 Santa Barbara Workshop, Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, London, Vol. 192, pp. 275–287.

Lau, M., R. Kosut and S. Boyd (1990). Parameter set estimation of systems with uncertain nonparametric dynamics and disturbances. *Proc. 29th IEEE Conf. Dec. and Contr.*, Honolulu, pp. 3162–3167.

Lee, W.S., B.D.O. Anderson, R.L. Kosut and I.M.Y. Mareels (1992). On adaptive robust control and control-relevant system identification. *Proc. Am. Contr. Conf.*, Chicago, pp. 2834–2841.

Lee, W.S., B.D.O. Anderson, R.L. Kosut and I.M.Y. Mareels (1993). On robust performance improvement through the windsurfer approach to adaptive robust control. *Proc. 32nd IEEE Conf. Dec. and Contr.*, San Antonio, pp. 2821–2827.

Levy, E.C. (1959). Complex curve fitting. *IRE Trans. Autom. Contr.*, Vol. 4, pp. 37–43.

Lin, P.L. and Y.C. Wu (1982). Identification of multi-input multi-output linear systems from frequency response data. *J. Dyn. Syst., Meas. and Contr.*, Vol. 104, pp. 58–64.

Liu, K. and R.E. Skelton (1990). Closed loop identification and iterative controller design. *Proc. 29th IEEE Conf. Dec. and Contr.*, Honolulu, pp. 482–487.

Ljung, L. (1987). *System Identification: Theory for the User,* Prentice-Hall, Englewood Cliffs, N.J.

Ljung, L. (1992). A discussion of model accuracy in system identification. *Int. J. of Adapt. Contr. and Sign. Proc.*, Vol. 6, pp. 161–171.

Ljung, L. and B. Wahlberg (1992). Asymptotic properties of the least-squares method for estimating transfer functions and disturbance spectra. *Adv. Appl. Prob.*, Vol. 24, pp. 412–440.

Lu, L.C. and Y. Wang (1988). A variable transformation method for solving linear absolute-value objective-function problems with linear constraints. *J. of Comp. and Appl. Math.*, Vol. 21, pp. 111–113.

Luenberger, D.G. (1966). Observers for multivariable systems. *IEEE Trans. Autom. Contr.*, Vol. AC-11, pp. 190–197.

Luenberger, D.G. (1984). *Linear and Nonlinear Programming,* Addison Wesley, U.K., 2nd edition.

Lunze, J. (1989). *Robust Multivariable Feedback Control,* Prentice-Hall, Englewood Cliffs, N.J.

Maciejowski, J.M. (1989). *Multivariable Feedback Design,* Addison Wesley, U.K.

Mäkilä, P.M. (1991a). Laguerre methods and H_∞ identification of continuous-time systems. *Int. J. of Contr.*, Vol. 53, pp. 689–707.

Mäkilä, P.M. (1991b). Robust identification and Galois-sequences. *Int. J. of Contr.*, Vol. 54, pp. 1189–1200.

Mäkilä, P.M. and J.R. Partington (1992). Robust identification of strongly stabilizable systems. *IEEE Trans. Autom. Contr.*, Vol. AC-37, pp. 1709–1716.

Manoukian, E.B. (1986). *Modern Concepts and Theorems of Mathematical Statistics,* Springer Verlag, New York.

McDonald, J.S. and J.B. Pearson (1991). ℓ_1-Optimal control of multivariable systems with output norm constraints. *Automatica,* Vol. 27, pp. 317–329.

McFarlane, D.C. and K. Glover (1989). *Robust controller design using normalized coprime factor plant descriptions,* Lecture Notes in Control and Information Sciences, Vol. 138, Springer Verlag, Berlin.

Milanese, M. and G. Belforte (1982). Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: linear families of models and estimators. *IEEE Trans. Autom. Contr.*, Vol. AC-27, pp. 408–414.

Milanese, M. and A. Vicino (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica,* Vol. 27, pp. 997–1009.

Moonen, M., B. de Moor, L. Vandenberghe and J. Vandewalle (1989). On- and off-line identification of linear state space models. *Int. J. of Contr.*, Vol. 49, pp. 219–232.

Moonen, M. and J. Vandewalle (1990). A QSVD approach to on- and off-line state space identification. *Int. J. of Contr.*, Vol. 51, pp. 1133–1146.

Morari, M. and E. Zafiriou (1989). *Robust Process Control.* Prentice-Hall, Englewood Cliffs, N.J.

Murad, G., I. Postlethwaite, D. Gu and J. Whidborne (1993). Robust control of a glass tube shaping process. *Proc. 2nd European Contr. Conf.*, Groningen, The Netherlands, pp. 2350–2355.

Ninness, B.M. (1993). *Stochastic and Deterministic Modelling,* Ph.D. Thesis, Univ. of Newcastle, Australia.

Ninness, B.M. and G.C. Goodwin (1992). Robust frequency response estimation accounting for noise and undermodelling. *Proc. Am. Contr. Conf.*, Chicago, pp. 2847–2851.

Norton, J.P. (1987). Identification and application of bounded-parameter models. *Automatica*, Vol. 23, pp. 497–507.

Norton, J.P. and S.M. Veres (1991). Developments in parameter bounding. In *Topics in Stochastic Systems: Modelling, Estimation and Adaptive Control*, Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, London, Vol. 161, pp. 137–158.

Packard, A. and J.C. Doyle (1993). The complex structured singular value. *Automatica*, Vol. 29, pp. 71–109.

Packard, A. and P. Pandey (1993). Continuity properties of the real/complex structured singular value. *IEEE Trans. Autom. Contr.*, Vol. AC-38, pp. 415–428.

Partington, J.R. (1991). Robust identification and interpolation in H_∞. *Int. J. of Contr.*, Vol. 54, pp. 1281–1290.

Partington, J.R. and P.M. Mäkilä (1994). Analysis of linear methods for robust identification in ℓ_1. *Prepr. SYSID'94, 10th IFAC Symp. on Syst. Id.*, Copenhagen, Denmark, Vol. 2, pp. 79–84.

Payne, P.A. (1970). An improved technique for transfer function synthesis from frequency response data. *IEEE Trans. Autom. Contr.*, Vol. AC-15, pp. 480–483.

Pintelon, R., P. Guillaume, Y. Rolain, J. Schoukens and H. Van hamme (1993). Parametric identification of transfer functions in the frequency domain, a survey. *Proc. 32nd IEEE Conf. Dec. and Contr.*, San Antonio, pp. 557–566.

Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer Verlag.

Poolla, K., P.P. Khargonekar, A. Tikku, J. Krause and K. Nagpal (1994). A time-domain approach to model validation. *IEEE Trans. Autom. Contr.*, Vol. AC–39, pp. 951–959.

Prett, D.M., T.A. Skrovanek and J.F. Pollard (1987). Process identification – past, present, future. *The 1986 Shell Process Control Workshop. Process Control Research: Industrial and Academic Perspectives*, Butterworth Publ., Boston, pp. 79–104.

Pronzato, L. and E. Walter (1989). Experiment design in a bounded error context: comparison with D-optimality. *Automatica*, Vol. 25, pp. 383–391.

Rivera, D.E., X. Chen and D.S. Bayard (1993). Experimental design for robust process control using Schroeder-phased input signals. *Proc. Am. Contr. Conf.*, San Francisco, pp.895–899.

Rivera, D.E., J.F. Pollard and C.E. Garcia (1992). Control-relevant prefiltering: a systematic design approach and case study. *IEEE Trans. Autom. Contr.*, Vol. AC–37, pp. 964–974.

Rivera, D.E., J.F. Pollard, L.E. Sterman and C.E. García (1990). An industrial perspective on control relevant identification. *Proc. Am. Contr. Conf.*, San Diego, pp. 2406–2411.

Safonov, M.G., R.Y. Chiang and H. Flashner (1988). H_∞ robust control synthesis for a large space structure. *Proc. Am. Contr. Conf.*, Atlanta, pp. 2038–2045.

Sanathanan, C.K. and J. Koerner (1963). Transfer function synthesis as a ratio of two complex polynomials. *IEEE Trans. Autom. Contr.*, Vol. AC-8, pp. 56–58.

Scheid, R.E., D.S. Bayard and Y. Yam (1991). A linear programming approach to characterizing norm bounded uncertainty from experimental data. *Proc. Am. Contr. Conf.,* Boston, pp. 1956–1958.

Schoukens, J. and R. Pintelon (1994). Quantifying model errors of identified transfer functions. *IEEE Trans. Autom. Contr.,* Vol. AC-39, pp. 1733–1737.

Schrama, R.J.P. (1992a). *Approximate Identification and Control Design with Application to a Mechanical System,* Ph.D. Thesis, Mech. Eng. Syst. and Contr. Group, Delft Univ. of Techn., The Netherlands.

Schrama, R.J.P. (1992b). Accurate identification for control: the necessity of an iterative scheme. *IEEE Trans. Autom. Contr.,* Vol. AC-37, pp. 991–994.

Schrama, R.J.P. and P.M.J. Van den Hof (1993). Iterative identification and control design: a worked out example. In *The Modeling of Uncertainty in Control Systems,* Proc. 1992 Santa Barbara Workshop, Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, London, Vol. 192, pp. 289–302.

Sidman, M.D., F.E. DeAngelis and G.C. Verghese (1991). Parametric system identification on logarithmic frequency response data. *IEEE Trans. Autom. Contr.,* Vol. AC-36, pp. 1065–1070.

Smith, R.S. and J.C. Doyle (1992). Model validation: a connection between robust control and identification. *IEEE Trans. Autom. Contr.,* Vol. AC-37, pp. 942–952.

Söderström, T. and P.G. Stoica (1983). *Instrumental Variable Methods for System Identification,* Lect. Notes in Contr. and Inf. Sciences, Springer-Verlag.

Söderström, T. and P.G. Stoica (1989). *System Identification,* Prentice-Hall, Englewood Cliffs, N.J.

Spanos, J.T. (1991). Algorithms for ℓ_2 and ℓ_∞ transfer function curve fitting. *Proc. AIAA Guid., Navig. and Contr. Conf.,* New Orleans, pp. 1739–1747.

Stoica, P.G. and T. Söderström (1983). Optimal instrumental variable estimation and approximate implementations. *IEEE Trans. Autom. Contr.,* Vol. AC-28, pp. 757–772.

Szegö, G. (1975). *Orthogonal Polynomials,* American Mathematical Society, Providence, RI, U.S.A., 4th edition.

Theodosopoulos, T.V. (1994). Linear programming algorithms for worst-case identification in ℓ_1. *Proc. Am. Contr. Conf.,* Baltimore, pp. 117–121.

Tischler, M.B. (1991). System identification requirements for high-bandwidth rotorcraft flight control system design. *Proc. Am. Contr. Conf.,* Boston, pp. 2494–2502.

Tse, D.N.C., M.A. Dahleh and J.N. Tsitsiklis (1993). Optimal asymptotic identification under bounded disturbances. *IEEE Trans. Autom. Contr.,* Vol. AC-38, pp. 1176–1190.

Van den Boom, T., M. Klompstra and A. Damen (1991). System identification for H_∞-robust control design. *Proc. 9th IFAC Symp. Ident. and Syst. Par. Est.,* Budapest, Hungary, pp. 1431–1436.

Van den Hof, P.M.J. and R.J.P. Schrama (1994). Identification and control – closed loop issues. *Prepr. SYSID'94, 10th IFAC Symp. on Syst. Id.,* Copenhagen, Denmark, Vol. 2, pp. 1–13.

Van den Hof, P.M.J., R.J.P. Schrama and P.M.M. Bongers (1993a). On nominal models, model uncertainty and iterative methods in identification and control design.

In *The Modeling of Uncertainty in Control Systems, Proc. 1992 Santa Barbara Workshop,* Lecture Notes in Contr. and Inf. Sciences, Springer Verlag, London, Vol. 192, pp. 39–50.

Van den Hof, P.M.J., R.J.P. Schrama, O.H. Bosgra and R.A. de Callafon (1993b). Identification of normalized coprime plant factors for iterative model and controller enhancement. *Proc. 32nd IEEE Conf. Dec. and Contr.,* San Antonio, pp. 2839–2844.

Van Overschee, P. and B. de Moor (1993). N4SID: subspace identification of a glass-tube manufacturing process. *Proc. 2nd European Contr. Conf.,* Groningen, The Netherlands, pp. 2338–2343.

Van Overschee, P. and B. de Moor (1994). N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica,* Vol. 30, pp. 75–93.

Veres, S.M. and J.P. Norton (1989). Structure identification of parameter bounding models by use of noise structure bounds. *Int. J. Contr.,* Vol. 50, pp. 639–649.

Veres, S.M. and J.P. Norton (1991). Structure selection for bounded parameter models: consistency conditions and selection criterion. *IEEE Trans. Autom. Contr.,* Vol. AC-36, pp. 474–481.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica,* Vol. 30, pp. 61–74.

Viberg, M. (1994). Subspace methods in system identification. *Prepr. SYSID'94, 10th IFAC Symp. on Syst. Id.,* Copenhagen, Denmark, Vol. 1, pp. 1–12.

Vidyasagar, M. (1985). *Control System Synthesis: A Factorization Approach,* MIT Press.

Vidyasagar, M. (1986). Optimal rejection of persistent bounded disturbances. *IEEE Trans. Autom. Contr.,* Vol. AC-31, pp. 527–534.

Wahlberg, B. (1991). System identification using Laguerre models. *IEEE Trans. Autom. Contr.,* Vol. AC-36, pp. 551–562.

Wahlberg, B. (1994). System identification using Kautz models. *IEEE Trans. Autom. Contr.,* Vol. AC-39, pp. 1276–1282.

Wahlberg, B. and L. Ljung (1986). Design variables for bias distribution in transfer function estimation. *IEEE Trans. Autom. Contr.,* Vol. AC-31, pp. 134–144.

Wahlberg, B. and L. Ljung (1992). Hard frequency-domain model error bounds from least-squares like identification techniques. *IEEE Trans. Autom. Contr.,* Vol. AC-37, pp. 900–912.

Walter, E. and H. Piet-Lahanier (1990). Estimation of parameter bounds from bounded error data: a survey. *Math. and Comp. in Sim.,* Vol. 32, pp. 449–468.

Whitfield, A.H. (1987). Asymptotic behaviour of transfer function synthesis methods. *Int. J. Contr.,* Vol. 45, pp. 1083–1092.

Wortelboer, P.M.R. and O.H. Bosgra (1992). Generalized frequency weighted balanced reduction. *Proc. 31st IEEE Conf. Dec. and Contr.,* Tucson, pp. 2848–2849.

Wortelboer, P.M.R. (1994). *Frequency Weighted Balanced Reduction of Closed-Loop Mechanical Servo-Systems: Theory and Tools,* Ph.D. Thesis, Mech. Eng. Syst. and Contr. Group, Delft Univ. of Techn., The Netherlands.

Youla, D.C., J.J. Bongiorno and H.A. Jabr (1976a). Modern Wiener-Hopf design of optimal controllers — Part I: the single-input-single-output case. *IEEE Trans. Autom. Contr.*, Vol. AC-21, pp. 3–13.

Youla, D.C., H.A. Jabr and J.J. Bongiorno (1976b). Modern Wiener-Hopf design of optimal controllers — Part II: the multivariable case. *IEEE Trans. Autom. Contr.*, Vol. AC-21, pp. 319–338.

Younce, R.C. and C.E. Rohrs (1992). Identification with nonparametric uncertainty. *IEEE Trans. Autom. Contr.*, Vol. AC-37, pp. 715–728.

Zames, G. (1981). Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms and approximate inverses. *IEEE Trans. Autom. Contr.*, Vol. AC-26, pp. 301–320.

Zang, Z., R.R. Bitmead and M. Gevers (1991). H_2 iterative model refinement and control robustness enhancement. *Proc. 30th IEEE Conf. Dec. and Contr.*, Brighton, pp. 279–284.

Zang, Z., R.R. Bitmead and M. Gevers (1992). Disturbance rejection: on-line refinement of controllers by closed loop modelling. *Proc. Am. Contr. Conf.*, Chicago, pp. 2929–2833.

Zhu, Y. (1989). Estimation of transfer functions: asymptotic theory and a bound of model uncertainty. *Int. J. of Contr.*, Vol. 49, pp. 2241–2258.

Ziegler, J. and N. Nichols (1942). Optimum settings for automatic controllers. *ASME Trans.*, Vol. 64, pp. 759–768.

Glossary of Symbols

In this glossary only symbols have been included which have a certain global meaning throughout the thesis. Locally defined symbols have not been included.

Latin Symbols

a_k	parameter in alternative parametrization of $d(\xi)$
$a(t)$	response of tail $\bar{G}(z)$ or $\bar{G}_0(z)$ to input $u(t)$
$\bar{a}(t)$	upper bound on $\lvert a(t)\rvert$
$\mathcal{A}$	set of ℓ_∞-stable, LTI, causal, discrete-time systems
b_k	parameter in alternative parametrization of $d(\xi)$
$b_k(k')$	pulse response parameter of $B_k(z)$
$b(t)$	response of $G(z)$ or $G_0(z)$ to past input signal $\{u(t)\}_{t=-\infty,\ldots,0}$
$\bar{b}(t)$	upper bound on $\lvert b(t)\rvert$
$B_k(z)$	basis function
c	constant,
	or parameter in alternative parametrization of $d(\xi)$
$\bar{c}(p)$	cross-covariance noise bound
$\bar{c}_e(p)$	extended cross-covariance noise bound $\bar{c}(p)$,
	including effects of undermodelling and unknown initial conditions
$c_l(p)$	lower bound on the sample cross-covariance
$c_u(p)$	upper bound on the sample cross-covariance
$c_w(\tau)$	window function
$C(z)$	controller
$\mathbb{C}$	set of complex numbers
d	denominator degree
d_k	denominator parameter
$d(p)$	error due to truncation (bias), i.e. $\sum_{t=t_s}^{N} r_p(t)a(t)$
$\bar{d}(p)$	upper bound on $\lvert d(p)\rvert$
$d(z)$	denominator
$D_L(\xi)$	multivariable left denominator
$D_R(\xi)$	multivariable right denominator
$e(t)$	output error
$\hat{e}(t)$	estimate of noise signal $e_0(t)$
$\bar{e}(t)$	upper bound on $\lvert e(t)\rvert$

$\bar{e}_e(t)$	extended time-domain noise bound $\bar{e}(t)$, including effects of undermodelling and unknown initial conditions
$e_l(t)$	lower bound on $e(t)$
$e_u(t)$	upper bound on $e(t)$
$e_0(t)$	noise signal
$E(\omega_j)$	DFT of the signal $e(t)$
$E_0(\omega_j)$	DFT of the signal $e_0(t)$
f	parameter in Tukey-window
$f(p)$	error due to past input signals (initial conditions), i.e. $\sum_{t=t_s}^{N} r_p(t)b(t)$
$\bar{f}(p)$	upper bound on $\lvert f(p)\rvert$
$f_{m'}(x)$	function defined in Lemma 3.4.1
$f(\omega_j)$	upper bound on $\lvert E(\omega_j)\rvert$
$\mathcal{F}$	set of feasible systems, i.e. set of systems consistent with data and prior information
$g(k)$	generalized pulse response parameter of $G(z)$, i.e. $G(z) = \sum_{k=0}^{\infty} g(k)B_k(z)$
$\widehat{g}(k)$	generalized pulse response parameter of $\widehat{G}(z)$
$\bar{g}(k)$	upper bound on $\lvert g_0(k)\rvert$ or $\lvert g(k)\rvert$
$g_l(k)$	lower bound on $g(k)$
$g_u(k)$	upper bound on $g(k)$
$g_0(k)$	generalized pulse response parameter of $G_0(z)$
$G(z)$	transfer function of a discrete-time, LTI system
$G(\omega_j)$	frequency response data
$\widehat{G}(z)$	(identified) nominal model
$\widetilde{G}(z)$	first part of $G(z)$, i.e. $\sum_{k=0}^{n} g(k)B_k(z)$
$\bar{G}(z)$	tail part of $G(z)$, i.e. $\sum_{k=n+1}^{\infty} g(k)B_k(z)$
$\widehat{G}_C(z)$	transfer function of the central estimate
$\widehat{G}_C(\omega)$	frequency response of the central estimate
$G_0(z)$	system to be identified and controlled
$\widetilde{G}_0(z)$	first part of $G_0(z)$, i.e. $\sum_{k=0}^{n} g_0(k)B_k(z)$
$\bar{G}_0(z)$	tail part of $G_0(z)$, i.e. $\sum_{k=n+1}^{\infty} g_0(k)B_k(z)$
$\widehat{G}_\Omega(z)$	nominal model constructed on the basis of frequencies in the set Ω
$\mathcal{G}$	weighted additive uncertainty set
h_∞	auxiliary variable in frequency response curve fit procedure
i	imaginary number, $i^2 = -1$
$\mathrm{Im}\,(\cdot)$	imaginary part of $\cdot$
l	number of frequencies in the set Ω, or truncation parameter in ℓ_1-optimal model reduction problem
$\mathcal{L}$	set of systems described by linear inequalities
m	number of vertices of $\mathcal{P}_m(\omega_j)$
M	positive constant used in upper bound $\bar{g}(k) \leq M\rho^k$
$\mathcal{M}$	model set
n	numerator degree, or truncation parameter, i.e. order of $\widetilde{G}(z)$ or $\widetilde{G}_0(z)$
$\bar{n}$	upper bound on $\max_k \lvert n_k\rvert$

n_k	numerator parameter
$n(z)$	numerator
N	number of measurement samples
$\widetilde{N}$	effective number of measurement samples, $N - t_s + 1$
$N(\xi)$	multivariable numerator
$\mathcal{N}(0, \Lambda)$	normal distribution with zero mean and (co)variance Λ
p'	number of output signals of a MIMO system $G(z)$
$p(k)$	pulse response parameter of $W(z)G(z)$, i.e. $W(z)G(z) = \sum_{k=0}^{\infty} p(k)z^{-k}$
$\widehat{p}(k)$	pulse response parameter of $W(z)\widehat{G}(z)$
$\widetilde{p}(k)$	pulse response parameter of $W(z)\widetilde{G}(z)$
$\bar{p}(k)$	pulse response parameter of $W(z)\bar{G}(z)$
$p_{\mathrm{c}}(k)$	$(p_{\mathrm{l}}(k) + p_{\mathrm{u}}(k))/2$
$p_{\mathrm{l}}(k)$	lower bound on $p(k)$
$\widetilde{p}_{\mathrm{l}}(k)$	lower bound on $\widetilde{p}(k)$
$\bar{p}_{\mathrm{l}}(k)$	lower bound on $\bar{p}(k)$
$p_{\mathrm{u}}(k)$	upper bound on $p(k)$
$\widetilde{p}_{\mathrm{u}}(k)$	upper bound on $\widetilde{p}(k)$
$\bar{p}_{\mathrm{u}}(k)$	upper bound on $\bar{p}(k)$
$p_0(k)$	pulse response parameter of $W(z)G_0(z)$
$\widetilde{p}_0(k)$	pulse response parameter of $W(z)\widetilde{G}_0(z)$
$\bar{p}_0(k)$	pulse response parameter of $W(z)\bar{G}_0(z)$
$\mathcal{P}(\omega)$	convex uncertainty region in the complex plane
$\mathcal{P}_m(\omega_j)$	convex polytope in the complex plane with m vertices
q	forward shift operator, i.e. $qu(t) = u(t+1)$
q'	number of input signals of a MIMO system $G(z)$
r	upper bound on real part of roots of $d(\xi)$
$r(t)$	signal correlated to $u(t)$ but independent of $e_0(t)$
$\bar{r}(t)$	external reference signal
$r_p(t)$	signal correlated to $u(t)$ but independent of $e_0(t)$
$R_x(\tau)$	auto-covariance function of the signal $x(t)$
$R_x^N(\tau)$	auto-covariance function of $x(t)$ based on N data samples
$\widehat{R}_x^N(\tau)$	estimated auto-covariance function of $x(t)$ based on N data samples
$R_{xy}(\tau)$	cross-covariance function of the signals $x(t)$ and $y(t)$
$R_{xy}^N(\tau)$	cross-covariance function of $x(t)$ and $y(t)$ based on N data samples
$\mathrm{Re}\,(\cdot)$	real part of $\cdot$
$\mathbb{R}$	set of real numbers
$\mathcal{R}$	(stability) region in the complex plane
s	Laplace variable, or number of cross-covariance noise constraints
$s_{k'}(k)$	pulse response parameter of $S_{k'}(z)$
$S_{k'}(z)$	$W(z)B_{k'}(z)$
$\mathcal{S}$	system uncertainty set
t	time
t_s	starting sample,

	the first $t_s - 1$ output data samples are not used in the identification
T	period time of a periodic signal
T_{yx}	transfer function from signal $x(t)$ to signal $y(t)$
$\bar{u}$	upper bound on $\max_{t \le 0} \lvert u(t) \rvert$
$u(t)$	input signal
$U(t)$	column-vector with element $k+1$ equal to $w_k(t)$
$v_k(t)$	$B_k(q)r(t)$
$v_k(\omega_j)$	vertex of $\mathcal{P}_m(\omega_j)$
$V(t)$	column-vector with element $k+1$ equal to $v_k(t)$
w	parameter in Tukey-window
$w_k(t)$	$B_k(q)u(t)$
$w(N)$	window parameter specifying the width of the window
$W(z)$	weighting in uncertainty structure
$\widetilde{W}(z)$	weighting in nominal identification (system approximation) criterion
$W_1(z)$	weighting in H_∞-control design procedure
$W_2(z)$	weighting in H_∞-control design procedure
$y(t)$	output signal
z	z-transform variable
$z_{\mathrm{d},k}$	solution to $d(z) = 0$

Greek Symbols

α	probability, confidence level
β	upper bound on $\left\lVert \frac{\mathrm{d}\,\Delta(z)}{\mathrm{d}\,z} \right\rVert_\infty$
β_1	upper bound on $\left\lVert \frac{\mathrm{d}\,G(z)}{\mathrm{d}\,z} \right\rVert_\infty$ or $\left\lvert \frac{\mathrm{d}\,G(\omega)}{\mathrm{d}\,\omega} \right\rvert$
β_2	upper bound on $\lVert G(z) \rVert_\infty$ or $\lvert G(\omega) \rvert$
δ	upper bound on the norm of the model error
$\delta(p)$	upper bound on error contribution of tail $\tilde{G}_0(z)$
$\delta_{\mathrm{l}}(\omega_j)$	lower bound on worst-case weighted model error
$\delta_{\mathrm{u}}(\omega_j)$	upper bound on worst-case weighted model error
Δ	uncertainty block
$\Delta(z)$	model error $W(z)\left(G(z) - \widehat{G}(z)\right)$
θ	parameter vector
θ_0	parameter vector of data-generating system
Θ_N	feasible parameter set
λ_j	worst-case distance between subsequent frequencies, i.e. $\max\{\omega_j - \omega_{j-1}, \omega_{j+1} - \omega_j\}$
Λ	(co)variance matrix
$\widehat{\Lambda}^N$	estimated covariance matrix based on N data samples
$\mu(\cdot)$	structured singular value of $\cdot$
$\mu_k(\omega_j)$	quantities used to characterize $\mathcal{P}_m(\omega_j)$
$\widetilde{\mu}_k(\omega_j)$	part of $\mu_k(\omega_j)$ corresponding to $\widetilde{G}(e^{i\omega_j})$

$\bar{\mu}_k(\omega_j)$	part of $\mu_k(\omega_j)$ corresponding to $\bar{G}(e^{i\omega_j})$
ξ	variable denoting z or s
$\xi_{\mathrm{d},k}$	solution to $d(\xi) = 0$
$\xi(\omega)$	$e^{i\omega}$ (if $\xi = z$) or $i\omega$ (if $\xi = s$)
ρ	constant (between 0 and 1) used in upper bound $\bar{g}(k) \leq M\rho^k$, or upper bound on absolute value of roots of $d(\xi)$
$\rho(\cdot)$	spectral radius of $\cdot$
$\sigma(\cdot)$	singular value of $\cdot$
$\sigma_{\max}(\cdot)$	maximum singular value of $\cdot$
$\phi(t)$	regression vector
$\Phi_x(\omega)$	auto-spectrum of the signal $x(t)$
$\chi^2(n)$	Chi-squared distribution with n degrees of freedom
ω	frequency
Ω	set of frequencies $\{\omega_1, \ldots, \omega_l\}$ with $0 \leq \omega_1 < \ldots < \omega_l \leq \pi$

Abbreviations

DFT	Discrete Fourier Transform
FIR	Finite Impulse Response
IV	Instrumental Variable
LP	Linear Programming
LTI	Linear Time Invariant
LTIC	Linear Time Invariant Causal
MIMO	Multi Input Single Output
MISO	Multi Input Multi Output
NLTVC	Nonlinear Time Varying Causal
PID	Proportional Integral Differential
PRBS	Pseudo Random Binary Sequence
SISO	Single Input Single Output
s.t.	subject to
w.p.	with probability

Samenvatting

Systeemidentifikatie voor Robuuste Procesregeling

Nominale Modellen en Foutgrenzen

Het prestatie-niveau van industriële processen kan vaak verbeterd worden door implementatie van een automatische terugkoppel-regelaar. Bij het ontwerp van een regelaar kan vruchtbaar gebruik gemaakt worden van een model van het te regelen systeem, mits op de juiste manier rekening wordt gehouden met het feit dat een model niet perfekt is. Als een *nominaal model* en een bovengrens op de onvermijdelijke *modelfout* worden gespecificeerd, kan een *robuuste* regelaarontwerpmethode gebruikt worden om een regelaar te ontwerpen met een gegarandeerd prestatie-niveau voor het systeem.

In dit proefschrift wordt het probleem beschouwd van de *identifikatie* van een nominaal model en een expliciete bovengrens op de modelfout op grond van meetdata. Met name is het doel een parametrisch model van een bepaalde orde te schatten, en een bovengrens op de H_∞- of ℓ_1-norm van de gewogen additieve modelfout, geschikt voor gebruik in H_∞- respectievelijk ℓ_1-robuuste regelaarontwerpmethoden. Het is duidelijk dat een al te *conservatieve* bovengrens op de modelfout zou betekenen dat het haalbare prestatie-niveau van de regelaar onnodig beperkt wordt. Daarom dient de identifikatie zo uitgevoerd te worden dat de bovengrens op de modelfout zo klein mogelijk is, en wel met name in de *regelaar-relevante* frequentiegebieden, dat wil zeggen rond de cross-over frequentie. Aan de ene kant vereist dit de identifikatie van een nominaal model, zodanig dat de modelfout klein is in de regelaar-relevante frequentiegebieden. Aan de andere kant dient de bovengrens op de modelfout een strakke bovengrens te zijn voor deze frequentiegebieden.

Om dit probleem van robuust-regelaarontwerp-gerichte identifikatie op te lossen wordt in dit proefschrift een nieuwe identifikatie-procedure ontwikkeld. Eerst wordt op grond van meetdata, en a priori informatie omtrent systeem en ruis, een niet-parametrische *systeem-onzekerheidsverzameling* geïdentificeerd, welke bestaat uit frequentieresponsie of pulsresponsie onzekerheidsgebieden. Een procedure om een dergelijke onzekerheidsverzameling te construeren wordt een *onzekerheidsbegrenzende identifikatie*procedure genoemd. Vervolgens wordt op grond van de systeem-onzekerheidsverzameling een parametrisch model van een door de gebruiker gekozen orde gecon-

strueerd, waarbij de H_∞- of ℓ_1-norm van een bepaald modelfoutkriterium wordt geminimaliseerd. Naar dit probleem wordt gerefereerd als het probleem van *systeembenadering* in H_∞- of ℓ_1-norm. Tenslotte wordt op grond van de systeem-onzekerheidsverzameling en het nominale model een bovengrens op de H_∞- of ℓ_1-norm van de gewogen modelfout bepaald. Het identifikatie-experimentontwerp en de weegfunkties in de identifikatieprocedure kunnen afgestemd worden teneinde de regelaar-relevante frequentiegebieden te benadrukken.

Zowel een procedure voor *deterministische* onzekerheidsbegrenzende identifikatie als een procedure voor *probabilistische* onzekerheidsbegrenzende identifikatie worden in dit proefschrift ontwikkeld. Beide procedures gebruiken deterministische a priori aannames omtrent het systeem. In het bijzonder wordt aangenomen dat het systeem lineair, tijdinvariant en mogelijk oneindig-dimensionaal is, en een bovengrens wordt aangenomen voor de pulsresponsie-coefficienten van het systeem. Verder maakt de methode voor deterministische identifikatie gebruik van deterministische ruisaannames, te weten de onbekend-maar-begrensde-amplitude ruisaanname, welke leidt tot deterministische onzekerheidsgebieden voor het systeem. Bovendien kan een nieuwontwikkelde *kruiscovariantie ruisbegrenzing* worden gebruikt, welke bepaalde uitmiddelingseigenschappen blijkt te bezitten die de standaard tijddomein ruisbegrenzing ontbeert. De deterministische onzekerheidsgebieden worden berekend door een verzameling lineaire programmeerproblemen op te lossen.

De procedure voor probabilistische onzekerheidsbegrenzende identifikatie maakt gebruik van stochastische ruisaannames, in het bijzonder de aanname dat de ruis een realisatie is van een stationair stochastisch proces, en onafhankelijk van het ingangssignaal. Dit leidt tot probabilistische onzekerheidsgebieden voor het systeem, of betrouwbaarheidsgebieden, welke correct zijn met een bepaalde kans. De betrouwbaarheidsgebieden worden berekend door expliciete bias- en variantie-uitdrukkingen van een instrumentele variable schatting te evalueren. Een vergelijking tussen beide procedures laat zien dat de probabilistische benadering in het algemeen tot veel kleinere onzekerheidsgebieden leidt dan de deterministische benadering. Met ander woorden is de stochastische ruisaanname veel minder conservatief dan de deterministische ruisaanname.

Het probleem van systeembenadering in H_∞ wordt opgelost door een nieuwe *frequentieresponsie curvefit*-procedure te ontwikkelen, welke een maximum-amplitude kriterium minimaliseert en stabiliteit van het resulterende model garandeert. Bij de oplossing van het curvefit-probleem wordt gebruik gemaakt van lineaire en niet-lineaire programmeertechnieken. Behalve dat het een oplossing levert voor het probleem van *identifikatie in* H_∞, levert de curvefit-procedure tevens een oplossing voor het H_∞-*optimale modelreduktie*probleem. Het probleem van systeembenadering in ℓ_1 wordt opgelost door een algoritme te ontwikkelen voor ℓ_1-*optimale modelreduktie* voor lineair geparametriseerde gereduceerde-orde modellen. In dit algoritme wordt gebruik gemaakt van lineair programmeren.

Praktische toepasbaarheid van de nieuwe identifikatieprocedure wordt aangetoond door middel van succesvolle toepassing op een multivariabel industrieel glasbuizenfabrikageproces.

Curriculum Vitae

Naam	Richard Gerrit Hakvoort
10 juni 1968	Geboren te Delft
1980–1986	VWO aan het Christelijk Lyceum Delft
1986–1990	Studie Werktuigbouwkunde aan de Technische Universiteit Delft; met lof afgestudeerd; voor het afstudeerverslag *"Benaderende Identifikatie ten behoeve van Regelaarontwerp"* de *Universiteitsfonds Delft Prijs 1990* van het Universiteitsfonds Delft en de *Regeltechniekprijs 1990* van het Koninklijk Instituut van Ingenieurs (KIvI) ontvangen
1990–1994	AIO in de vakgroep Meet- en Regeltechniek van de faculteit Werktuigbouwkunde van de Technische Universiteit Delft
1995–...	Als Associate Research Petrophysicist werkzaam bij het Koninklijke/Shell Exploratie en Produktie Laboratorium te Rijswijk

The fear of the Lord is the beginning of wisdom,
and knowledge of the Holy One is understanding.
Proverbs 9:10