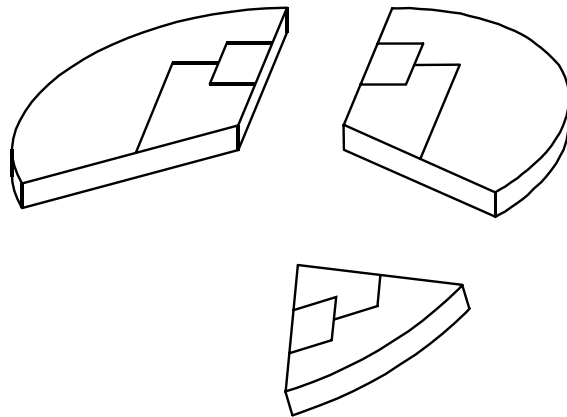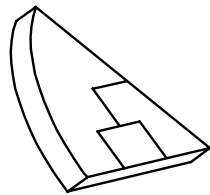# Feedback Oriented Identification for Enhanced and Robust Control

a fractional approach applied to a wafer stage

Raymond de Callafon

Feedback Oriented Identification for Enhanced and Robust Control

a fractional approach applied to a wafer stage

Feedback Oriented Identification for Enhanced and Robust Control

a fractional approach applied to a wafer stage

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 voor Promoties aangewezen,

op dinsdag 13 oktober 1998 te 10:30 uur

door Raymond Arnoud DE CALLAFON
werktuigkundig ingenieur
geboren te Rotterdam

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

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

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. ir O.H. Bosgra,	Technische Universiteit Delft, promotor
Dr ir P.M.J. Van den Hof,	Technische Universiteit Delft, toegevoegd promotor
Prof. M. Gevers,	Université Catholique de Louvain
Prof. dr ir H. Kwakernaak,	Universiteit Twente
Prof. dr Olsder,	Technische Universiteit Delft
Prof. R.E. Skelton,	University of California, San Diego
Dr ir M. Steinbuch,	Philips

Printed in the Netherlands

. . . come follow me
you won't expect the illusion you'll see
it's my imagination
hand me your eyes
I will put them in front of mine
you'll see a little better. . .

Linda Perry, 'drifting'

Voor jou, Marijke

Voorwoord

Eindelijk ligt het er dan. Het werk van een aantal jaren dat je, gek genoeg, afsluit met het eerste hoofdstukje in het proefschrift. En toch is het één van de belangrijkste en leukste hoofdstukken, omdat ik weet dat iedereen graag even het voorwoord van een proefschrift leest. Het onderscheidt zich namelijk van de rest van het proefschrift. Niet alleen omdat ik het in het Nederlands geschreven heb, maar ook omdat het net een stukje persoonlijker is dan de secties, definities, proposities en plaatjes in de opeenvolgende hoofdstukken. De reden hiervoor ligt in het feit dat ik hier graag wat mensen wil bedanken die direct of indirect aan mijn promotie en de afronding van dit proefschrift hebben bijgedragen.

Allereerst wil mijn promotor Okko Bosgra bedanken. Hij heeft ervoor gezorgd dat ik me in Delft ontwikkeld heb tot een promovendus waar hij hopelijk een beetje trots op kan zijn. Per slot van rekening zijn er toch maar weinig AIO's die hun professor op bezoek krijgen in het buitenland om even te laten te weten dat het proefschrift nu toch echt wel eens af moet. Okko, heel erg bedankt.

Verder gaat ook dank uit naar het Netwerk Systeem en Regeltheorie, tegenwoordig bekend onder de naam Dutch Institute of Systems and Control (DISC). Het Netwerk heeft voor het grootste gedeelte mijn werk en de conferentiebezoeken gesponsord en daarvoor ben ik hen zeer dankbaar.

Veel dank gaat ook uit naar Paul Van den Hof. Hij was niet alleen begeleider maar ook een hele goede collega om mee samen te werken. De combinatie van motivatie, creativiteit, scherpzinnigheid, een volgeschreven bord in zijn kamer en een gezonde dosis humor maakte het werken met hem erg plezierig.

Humor is voor mij trouwens altijd een belangrijke factor geweest in de vakgroep Meet- en Regeltechniek. Een bonte samenstelling van mensen in de groep zoals Cor Kremers, Piet Ruinard, Leo Beckers en Guus Bout hebben mijn tijd in Delft zeker veraangenaamd en versoepeld. De combinatie van humor en bittere ernst die bijvoorbeeld Sjoerd Dijkstra, Ton van der Weiden en Piet Teerhuis kunnen inbrengen is iets wat ik zeker zal missen. Ook Els Arkesteijn wil ik bedanken voor haar steun en begrip. Altijd bereid om orde te scheppen in mijn financiële chaos na een conferentiebezoek. Uiteraard kan ik ook Ben Wenneker, Peter Valk, Rolf van Overbeek en Jaap van Dieten niet vergeten voor hun mateloze geduld in geval van (mijn) computer

problemen.

Naast de 'regelgroep' in Delft, wil ik ook de 'regelgroep', die destijds in het Philips Natuurkundig Laboratorium werkte, bedanken. Deze groep, waarin onder andere Budi Sastra, Maarten Steinbuch en Frank Sperling aktief waren, gaf mij veel inspiratie. De regelmatige 'TUD/NATlab' bijeenkomsten waren altijd zeer waardevol, en ik hoop niet alleen voor mij.

Het AIO bestaan in Delft is ook veraangenaamd door de studenten, collega's, AIO's en kamergenoten die mij bleven motiveren. Ik wil Michel Verheijen en Erwin Walgers bedanken voor hun samenwerking en input. Ik dank Carsten Scherer en Peter Heuberger voor hun sterke inhoudelijke discussies. Ook de inhoudelijke discussies met mensen buiten Nederland, hetzij via email of tijdens een conferentie, zoals Bob Bitmead, Ari Partanen, Franky de Bruyne, Yossi Chait, Michel Gevers en Dan Rivera, waren een inspiratiebron. Daarnaast weet ik dat de wekelijkse AIO voordrachten mij niet populair maakten: 'Zal hij mij vragen of niet?'. Maar voor hun bereidwilligheid, hun stipte aanwezigheid en hun gezonde dosis humor wil ik graag David, Dick, Edwin, Gregor, Hans, Marco en Sjirk bedanken. In het bijzonder wil ik Dick de Roover bedanken voor zijn goede gezelschap als kamergenoot en het delen van de experimentele opstelling met mij.

Last, but not least, gaat er veel waardering uit naar mijn familie en vrienden die mij vaak door dik en dun gesteund hebben. Mijn ouders hebben wellicht geen idee wat ik al die jaren toch in Delft heb uitgespookt, maar ik weet zeker dat ze trots op mij zijn. Verder is mijn broer Paul, ik denk zonder dat hij zich dit realiseerde, ook een heel belangrijk voorbeeld voor mij geweest. Ik mag ook zeker mijn 'Schipluidense hospes' niet vergeten, want volgens hem 'staat hij toch aan de basis van al dit'.

Tenslotte is er een familie die mij altijd met open armen en veel warmte hebben ontvangen en begeleid. Marijke, Ayla en Danyal, ik denk niet dat ik ooit de woorden zal kunnen vinden om dat geluk te kunnen omschrijven. Marijke, het opdragen van dit proefschrift aan jou, mijn liefste vriendin, mijn grootste kameraadje en mijn beste steun en toeverlaat, is in verhouding slechts een minuscuul teken van mijn waardering voor jou.

<div style="text-align: right">

Raymond de Callafon

La Jolla, September 1998

</div>

Summary

Feedback Oriented Identification for Enhanced and Robust Control

a fractional approach applied to a wafer stage

R.A. de Callafon, October 1998

Feedback control and system identification both involve the control and prediction or modelling of the dynamical aspects of a system, and yet, they are merged quite sporadically. The use of experimental data to model the dynamics of a system is a powerful tool for providing models that can be used to develop a controller for that system. However, more refined and enhanced robust controllers can be developed when the system identification and the design of the feedback control are done simultaneously. This thesis contributes to the development of such an integrated approach of both feedback control and system identification, with the intent to set up a systematic procedure to design an enhanced and robustly performing feedback control for a dynamical system.

The emphasize of this thesis lies in the field of system identification and provides results and tools for a so-called feedback oriented identification of systems. New results to forward the integration of system identification with robust control design can be found in this thesis. The integrated approach is illustrated and applied successfully to an industrial high accuracy multivariable mechanical positioning system known as a wafer stage. Such a wafer stage is used in wafer steppers for manufacturing integrated circuits.

In order to monitor and ensure the closed-loop performance of the feedback control system, a model-based procedure is proposed. The model-based procedure involves the estimation of a set of models, built up from a nominal model equipped with a characterization of the model uncertainty. The identified set of models is then used in a robust control design method to obtain an enhanced and robust feedback control system. In order to ensure closed-loop performance enhancement while performing the subsequent steps of model set estimation and robust control design, closed-loop validation tests for both the feedback oriented modelling and the robust control design are formulated. The closed-loop validation tests guarantee that an upper bound on

the closed-loop performance can be improved monotonically.

Using system identification to find models for control design typically involves the identification of a system that is operating under feedback controlled conditions. Furthermore, approximate models or models of low complexity are needed in order to set up a manageable low order control design problem. For that purpose, this thesis contains a critical evaluation of so-called closed-loop approximate identification techniques that are used to address the problem of finding approximate models of a (possibly unstable) system on the basis of closed-loop data. A fractional model approach, where the possibly unstable system is represented and identified via stable coprime factorizations is used and is shown to be beneficial to address the closed-loop approximate identification problem.

With the fractional model approach, the nominal model is identified via the estimation of stable coprime factorizations. Frequency domain based identification techniques are used to estimate such a stable coprime factorization of the system. The model uncertainty is characterized by considering a perturbation in a so-called dual-Youla parametrization. The model uncertainty is estimated by existing techniques for probabilistic uncertainty bounding identification. It is shown that, due the fractional model approach and the chosen structure of the model set, the approximate and feedback relevant identification of the set of models can and has to be done entirely on the basis of closed-loop experiments. Furthermore, the proposed structure of the set of models is shown to be particularly useful to evaluate and monitor the closed-loop performance of the feedback controlled system.

The proposed model-based procedure provides a method to integrate system identification and robust control design. It provides a systematic procedure to design an enhanced and robust feedback controller for a possibly unstable, multivariable dynamical system. Each subsequent step of feedback oriented identification and robust control design within the procedure, can be used to guarantee progressive improvement of the closed-loop performance of the dynamical system.

Note to the Reader

In an attempt to guide the reader through the different chapters and sections of this thesis, the contents has been split up in six parts.

- The first part is the **prologue**, where the problem statement is given and the background to the problem field is sketched.

- The second part discusses the **closed-loop identification** problem. In this part, the problems associated to the identification of dynamic systems operating in a feedback connection is outlined and the fractional approach, as used in this thesis, is discussed in more detail.

- The third part constitutes the (mathematical) **procedure** used to address the problem statement of this thesis. This part combines existing techniques and the newly developed techniques of this thesis to complete a procedure that enables a feedback oriented identification for finding models to design an enhanced and robust performing feedback control.

- The fourth part of this thesis discusses the **application** of the procedure introduced in this thesis on a multivariable mechanical servo mechanism called a wafer stage.

- The fifth part is the **epilogue** and concludes the work that has been presented in this thesis.

- Finally, the **addenda** are combined in the last part. In this sixth part, the appendices, list of symbols, list of figures, bibliography, an index, a dutch summary and a curriculum vitae can be found.

To further improve the layout of this thesis, the sections are numbered within each chapter. Cross referencing between different items such as formulas, remarks, definitions, lemmas, theorems or figures is done via numerical labels that include either a chapter or a section number.

- The numerical labels for remarks, definitions, lemmas and theorems are numbered consecutively within each chapter and include the section number. As an example, Definition 4.2-5 refers to a definition given in Section 4.2, while Lemma 4.2-6 refers to the lemma in Section 4.2, following the definition given in Definition 4.2-5.

- The numerical labels for formulas are numbered separately within each chapter, but do not contain a section number. As an example, (4.3) refers to the third numbered formula given in chapter 4.

- The numerical labels for figures are also numbered separately within each chapter, and also do not contain a section number. As an example, Figure 4.3 refers to the third figure depicted in chapter 4.

- References to the bibliography are done by referring to author(s) and year of publication and are placed between brackets. If there are more references of the same author(s) in the same year, a letter will be added to the year to distinguish between the different publications. As an example, Bitmead *et al.* (1990a) refers to the ath publication written by Bitmead and co-authors in 1990 that can be found in the bibliography.

Contents

Part I

Prologue

1

Background and Problem Formulation

1.1 Introduction

Automatic control is a widely accepted and frequently used technique in order to enforce a dynamical system to behave in a satisfactory manner. The broad concept of a dynamical system can thereby refer to many biological or engineering processes in which the behaviour of a process as a function of time has to be conducted. The application of automatic control can be found in many complex industrial processes or sophisticated mechanical systems to attain a properly operating dynamical system.

Next to automatic control, system identification is used repeatedly to elucidate the dynamical aspects of a system. This procedure enables one to predict the dynamical behaviour of an unknown system on the basis of foregoing observations of the system. In this way, knowledge of the dynamical aspects of an industrial process or a mechanical system is acquired on the basis of experiments.

Although both techniques are involved with respectively the control and prediction of the dynamical aspects of a system, still they have been merged sporadically. Only recently the theory and applications within the fields of automatic control and system identification are being treated conjointly, in order to try to accommodate the research and engineering applications in both fields. As a result, more refined and enhanced automatic controllers can be developed to control a dynamical system.

This thesis focuses on both automatic control and system identification, with the aim to set up a systematic way to design a control for a given system on the basis of observations of the system. This approach is illustrated for an industrial wafer stepper system, a high accuracy mechanical positioning system used in chip manufacturing processes. The first chapter of this thesis is used to give a brief overview on the developments in the research area of automatic control and system identification. The usance of this overview is to illuminate their corresponding interrelation, thereby illustrating the contribution and problem formulation of this thesis.

1.2   Model-Based Feedback Control

1.2.1   The need for control

By means of automatic control, operational conditions of a system can be modified in order to satisfy specific performance specifications that may include enhanced accuracy and additional safety requirements. Moreover, many systems such as chemical processes, mechanical servo systems or electronic devices would fail to work properly without additional control. Therefore, control has played and is still playing an important role in the design of industrial and engineering processes that should meet improved performance requirements.

   An essential element in control has been the usage of feedback, (Horowitz 1959). In feedback-based automatic control, a system is equipped with a set of sensors and actuators. The sensors are used to measure specific signals of the system, whereas the feedback control processes these signals and drives the actuators to affect the behaviour of the system. In this way, the sensed signals of the system are processed and fed back, leading to the so-called feedback configuration depicted in Figure 1.1.
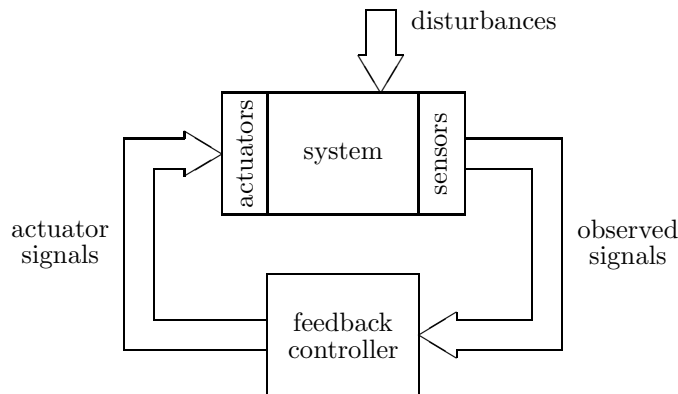
Fig. 1.1: Schematic diagram of a feedback controlled system.

   The application of feedback has numerous advantages and technical implications that can be exploited in controlling a system. One of the most important features is the ability to suppress unknown disturbances acting on the system. Alternatively, tracking of setpoints signals by means of an automatic controller can be labelled too as one of the main motivations to use feedback. One is referred for example to the books by Desoer and Vidyasagar (1975), Maciejowski (1989), Franklin *et al.* (1991), Kwakernaak and Sivan (1991) or Doyle *et al.* (1992) for an overview on the most important features of (linear) feedback controlled systems. It should be mentioned that the use of feedback is merely a special case of an automatic controlled system

as indicated in Willems (1991) or Willems (1992). However, for many engineering processes it is a natural starting point to control a system that is equipped with a distinguishable set of sensors and actuators, see also Boyd and Barrat (1991) or Willems (1995).

1.2.2 Classical approaches

Developments in feedback control emerged by the work of Nyquist (1932) and the analysis of feedback amplifiers by Black (1934). The actual design of feedback controllers was greatly simplified by the introduction of standard components like the lead/lag-, PI- and PID-controller. The simplicity of these components and the accompanying rules to tune them (Ziegler and Nichols 1942) requires only a limited amount of knowledge on the actual system to be controlled. Both aspects broadened the success and popularity of feedback control (Bode 1945, Evans 1950, Savant 1958). Nowadays, the design of feedback controllers based on PID tuning rules is still used extensively in controlling chemical processes or mechanical servo systems.

Despite the success of the classical PID-control, the everlasting demand to control more complex and multivariable systems at a higher performance level required a less heuristic approach to controller design. Furthermore, the availability of powerful computer hardware opened the possibility to design, analyze and implement more complex feedback controllers in a systematic way. During the sixties, the use of state space representations (Horowitz 1963) and the notion of optimal control (Athans and Falb 1966) to compute controllers on the basis of a (linear) model of the system, strongly improved this systematic design. In this so-called modern or state-space control design (Boyd and Barrat 1991), the tuning of a complex controller is done by means of calculation. For that purpose the actual system is represented by a model, possibly in a state-space form. The underlying assumption is the certainty-equivalence principle, implying that the model being used is an exact representation of the system to be controlled, see also Kwakernaak and Sivan (1972) or Anderson and Moore (1990).

A more sophisticated control algorithm can outperform simple control, provided that it has been designed properly. Unfortunately, the certainty-equivalence principle requires the impracticable task to formulate a model that meticulously describes the actual system. The system to be controlled can be highly complex and may contain nonlinear and time varying dynamic phenomena, whereas a model can only be an approximation of the actual system. Usually it is desired that a model is linear, time invariant and of low complexity in order to formulate a control design procedure that is solvable via existing techniques. If the design of the controller is done on the basis of such a simplified model, the mismatch between system and model may cause the performance of the actual controlled system to deteriorate. In some cases the controlled system might not even be stabilized by the designed feedback controller.

1.2.3   Robust control

To circumvent the problems associated with the mismatch between a model and a system to be controlled, the design and analysis of robust controllers emerged. In robust control or so-called post-modern control (Zhou *et al.* 1996), the presence of uncertainty, opposite to the certainty-equivalence principle as discussed before, can be taken into account. To account for the presence of uncertainty, in robust control typically the system is not modelled by a single model. Instead, the system is assumed to lie in a set of models that is built up from a nominal model along with an allowable model perturbation (Doyle 1982, Francis 1987, Maciejowski 1989). Consequently, robust control design methods synthesise a controller that is guaranteed to satisfy stability and additional performance requirements not just for the nominal model, but for the complete set of models of which the actual system is assumed to be an element. As a consequence, the controller is said to be robust and indicated by the term "robust controller" as it satisfies robustness properties such as stability robustness or even stronger, performance robustness (Doyle *et al.* 1992).

Based on the motivations in the work of Zames (1981), the application of the so-called $\mathcal{H}_\infty$ norm was found to be appropriate to describe the allowable model perturbation and to characterize the performance of a feedback controller. As a consequence, the attention in the literature has been focused on the design or synthesis of robust controllers by means of $\mathcal{H}_\infty$ methods. The $\mathcal{H}_\infty$ problem has been studied before but the methods heavily relied on operator-theoretic methods, see for example Adamjan *et al.* (1978). One of the first general solutions that served as a computational tool using state-space methods has been presented in the work by Doyle (1984). Extensions that followed the line of this work have been formulated in Francis (1987) and Francis and Doyle (1987). These extensions use a state space approach that is closely related to the Hankel operators as used in the paper by Glover (1984). Finally, a generalization of the results has been published by Khargonekar *et al.* (1988) and Glover and Doyle (1988) in which a solution for a general $\mathcal{H}_\infty$ problem has been posed in terms of solutions of an associated Riccati equation. A combination of the work of most of the authors mentioned above lead to the frequently cited paper by Doyle *et al.* (1989). For a more comprehensive overview one is also referred to the book by Green and Limebeer (1995) or Zhou *et al.* (1996).

Alternative computational methods for robust controller design have also been presented in the literature. The analysis and design approach as presented in Vidyasagar (1985), Georgiou and Smith (1990) and the work of McFarlane and Glover (1990) or McFarlane and Glover (1992) differ in the way both the allowable model perturbation and the performance specification are incorporated. Alternative optimal robust control procedures, using a so-called l_1-norm either to bound the uncertainty or to characterize performance, has been considered in Dahleh and Boyd Pearson Jr (1987*b*) and Dahleh and Boyd Pearson Jr (1987*a*). Closely related to the $\mathcal{H}_\infty$ methods of Doyle

et al. (1989), the so-called μ-analysis and synthesis (Doyle and Stein 1981, Packard and Doyle 1993) can handle a wider class of allowable model perturbations in order to design robust controllers more thoughtfully. Parallel to these developments, the robust control analysis and synthesis formulated on the basis of linear matrix inequalities (Boyd and Barrat 1991, Boyd *et al.* 1994) is a promising development. The use of linear matrix inequalities enables even more freedom in specifying performance and allowable model perturbations. Although linear matrix inequalities have already been proven to be very appealing and useful in control design (Willems 1971), the availability of powerful algorithms (Nesterov and Nemirovsky 1994) allows the associated convex optimization problems to be solved via a computational solution.

On the whole, the introduction and the continuation in the developments of robust control offers possibilities to design enhanced controllers for systems that should satisfy high-performance objectives and additional robustness properties. Still, the techniques discussed above are model-based and require a model, or even better, a nominal model along with an allowable model perturbation to characterize a set of models.

1.3   Modelling by System Identification

1.3.1   Motivation for experimental modelling

In an engineering sense, a model constitutes a formalism in order to describe the knowledge about a particular system. The systems to be modelled may vary within the wide range of engineering processes such as communication, electrical or mechanical systems. Alternatively, more or less non-technical applications such as environmental or biological systems have the need to come up with models to formalise the properties of an actual system in terms of a mathematical model.

The appearances, purposes and applications of a model are innumerable and probably even more extensive than the number of existing engineering processes. Referring to the previous sections, the attention in this thesis is focused on the purpose of deriving dynamic models that are to be used for model-based (robust) control design. Essentially, for this purpose one can distinguish the following two main approaches to derive a dynamical model.

- Physical modelling.

  The mathematical equations to describe the dynamical behaviour of a system are derived from physical laws or first principles. In this type of modelling, one heavily relies on the theory of the related science, in which the underlying assumptions and modelling aspects have already been accepted.

- Experimental modelling.

In this type of modelling the emphasize lies on the use of experimental data obtained from a system. The modelling is done by characterizing any systematic relations that are present in the data. The systematic relations obtained from the data now constitute the model, sometimes without explicitly taking into account any physical relations that the actual system may emulate.

Due to the lack of information about the internal representation of the actual system used in experimental modelling, this type of modelling is often referred to as "black box" modelling or system identification (Åström and Bohlin 1965, Åström and Eykhoff 1971, Eykhoff 1974). Opposite to "black box" modelling is the full information used in physical or "white box" modelling, in which the model is obtained by investigating the interior parts of the system in detail. This type of modelling is frequently done by the process of tearing and zooming, as mentioned recently in Willems (1995). As a consequence, a combination of the two approaches mentioned above is referred to as "grey box" type of modelling, see for example Jørgensen and Hangos (1995) for a recent reference. In this type of modelling, system identification techniques are used to resolve unknown quantities that appear in the equations obtained from first principles.

Supported by the numerous successful applications of system identification techniques that have been reported in the literature, the most important motivations to use system identification in this thesis can be summarized as follows.

- Complexity of the system or inadequate knowledge of the physical laws underlying the system behaviour limits the use of modelling by first principles in many engineering applications. In that case, system identification is a reasonable alternative for (approximate) modelling purposes and is denoted by the notion of approximate identification (Ljung and Glad 1994).

- As pointed in Section 1.2, it is desirable to formulate a model that is linear, time invariant and of low complexity in order to formulate a control design procedure that is solvable via existing techniques. System identification is able to yield models that satisfy these requirements.

- For feedback control of a system either the feedback system depicted in Figure 1.1 needs to be created, or is already available. Hence, the sensor and actuator signals for observing the dynamical behaviour of the system are readily available in many control applications and can be used for system identification purposes too.

Clearly, the success of a model obtained by system identification does not depend solely on the possibility to observe the system. Referring to Ljung (1987), pp. 8–9, the success of any system identification technique hinges on several aspects that may include the design and excitation of experiments, model structure determination, the

criterion used to determine the model and finally, the validation of the model. For a thorough treatment of these aspects one is also referred to the books by Norton (1986), Söderström and Stoica (1989) or Johansson (1993). These aspects will also reappear in this thesis. However, they will be formulated in the context of feedback controlled systems pointing to models that should be suitable for designing enhanced performing (robust) controllers.

1.3.2   Two main branches

Inevitably, a model obtained by system identification will be an inaccurate or approximate representation of the actual system that needs to be controlled. It is unrealistic to assume that an exact model of a system can be found on the basis of data that has been produced by the underlying system. This is due to the fact that data can only represent a finite time, possibly disturbed, observation of the unknown dynamical system. As a consequence, it is not possible to verify or validate the resemblance between a model and a system on the basis of a finite number of experiments (Smith and Doyle 1989). Therefore, the knowledge on the system remains incomplete and a model that exactly represents the system cannot be formulated without prejudice.

Fortunately, a robust control design paradigm can deal with the inexact knowledge of the system to be controlled. As pointed out in Section 1.2.3, a set of models is used to represent the inexact knowledge on the dynamical system to be controlled. A similar argumentation can also be used in the identification of an unknown dynamical system. For that purpose, the system identification should estimate a set of models that can be used in robust control design. Such an estimated set of models will be denoted by the notion of a model uncertainty set. The estimated set of models or model uncertainty set is then used to represent the incomplete knowledge of an unknown dynamical system, present due to the finite number of experiments being used during the identification.

As pointed out in Section 1.2.3, a model uncertainty set can be built up from an nominal model, along with an allowable model perturbation or bound on the model error[1]. Similar to this construction of a model uncertainty set, the research in the field of system identification can be divided in two distinguishable branches. A branch that focuses on the estimation of a nominal model and a branch that is considered with the quantification of a model error.

- estimation of a nominal model

  In fact, this can be viewed as the "classical approach" to system identification of building and validating (nominal) models based on observed data from a system. Starting from the work by Gauss at the beginning of 1800 and printed

---

[1]From an identification point of view, the allowable model perturbation is described by a "bound on the model(ling) error" to stress the incomplete knowledge of the system to be identified.

in Gauss (1901) or Gauss (1963), several methods have been reported in the literature, some of which have been listed below.

Well established and frequently applied techniques are based on the prediction error framework (Ljung 1987, Söderström and Stoica 1989). Alternative methods based on realization techniques (Ho and Kalman 1966, Zeiger and McEwen 1974, Kung 1978) or the approach used in Willems (1986), (1987) also focuses on the estimation of a nominal (state space) model. More recently and comparable to the realization based approach are the powerful sub-space methods (Viberg 1994) and the method based on canonical variate analysis (Larimore 1990). Recent contributions on sub-space identification can for example be found in the theses written by van Overschee (1995) or McKelvey (1995) or the paper by Verhaegen (1994).

- estimation of a model error

Some of the methods for estimation of a nominal model provide quality measures of the model being estimated by means of variance and bias expressions, see e.g. (Ljung 1987). These quality measures give rise to so-called "soft" bounds on the resulting model error, due to the stochastic or probabilistic nature of the confidence intervals being used in deriving the results (Wahlberg and Ljung 1991). Despite of the stochastic nature of the bound on the model error, the soft bounds provide a quality measure that is useful in many applications (Bayard 1992). Approaches to probabilistic model error bounding can for example also be found in the paper by Bayard (1992) or the work presented by Ninness (1993).

Unfortunately, "hard" error bounds are desirable in robust control to enforce and guarantee stability and performance robustness (Zames 1981). This has been one of the motivations to invoke the research on deriving such non-probabilistic or deterministic error bounds via a worst-case or deterministic approach to system identification (Milanese and Tempo 1985, Tempo 1988). This new approach to system identification is also indicated by various names such as $\mathcal{H}_\infty$ identification or worst-case estimation (Helmicki *et al.* 1989). The line of thinking has been continued by several researchers, see for example (Gu and Khargonekar 1992, Bai and Raman 1994, Chen and Nett 1995*a*). A nice overview on both the probabilistic and deterministic approach to model error bounding can also be found in the papers by Ninness and Goodwin (1995*a*) and Mäkilä *et al.* (1995).

Although the non-probabilistic approaches as discussed above do yield a "hard" error bound, generally the bound on the model error tends to be overestimated (Hjalmarsson 1993, Hakvoort 1994). An overestimation of the model error is mainly due to a conservative nature of the "hard" assumptions being made on the actual system or the noise that might be present on the data. As a result, the bound on

model error can be a (very) conservative estimate of the mismatch between nominal model and the actual system.

As a consequence, mixed deterministic-probabilistic approaches have been developed to diminish the effect of conservatism, see e.g. Hakvoort (1994) or de Vries and Van den Hof (1995). In these approaches the bound on the model error is derived by combining the results on "hard" and "soft" error bounding. At this moment it can be mentioned that the model error estimation of (Hakvoort and Van den Hof 1997) will be employed in this thesis.

To conclude, the availability of a nominal model along with a bound on the model error can be used to construct a model uncertainty set. The model uncertainty set is an estimated set of models that is used to represent the unavoidable incomplete knowledge on the unknown dynamical system to be controlled. As pointed out in Section 1.2.3, a controller can be designed on the basis of a set of models. However, the question remains whether or not a model uncertainty set, found by the techniques listed in the two branches above, is suitable for designing an enhanced robust controller. Clearly, the suitability will depend on the character of the constituents used to construct the model uncertainty set. These constituents include the choice of the nominal model and the way a tight bound on the model error is being characterized in order to capture the unknown dynamical system within the model uncertainty set.

1.3.3   Estimating models for robust control

As pointed out above, system identification techniques can be used to estimate model uncertainty sets on the basis of data obtained from an unknown dynamical system. As such, a model uncertainty set is used to represent the incomplete knowledge of a dynamical system to be controlled. For the purpose of designing a robust controller, such a model uncertainty set is built up from a nominal model and a bound on the model(ling) error. The utilisation of a nominal model and a bound on the model error provides the opportunity to structure the model uncertainty set, so that it can be used in a robust control design paradigm.

From an identification point of view, a model uncertainty set can be considered to consist of all models that are either validated (Ljung 1987) or cannot be invalidated by the available data (Smith and Doyle 1989). Clearly, such a set of models may consists of an uncountable number of possible models. However, the introduction of single nominal model, along with an allowable model perturbation enables one to capture this set of models relatively easily. This has been illustrated in a schematic picture depicted in Figure 1.2, where a set of models that cannot be invalidated by the data is characterized by means of a nominal model along with an allowable model perturbation. Furthermore, characterizing a set of models by means of a nominal model along with an allowable model perturbations enables one to design robust performing controllers that take into account the incomplete knowledge associated

with the availability of only finitely many, possibly disturbed, experiments.

As pointed out before, the main purpose of estimating a set of models in this thesis, is directed towards the design of a controller. It can be observed from Figure 1.2 that there is freedom in choosing a nominal model and the way in which a bound on the modelling error is described. However, both elements will determine the structure of the model uncertainty set and its suitability for the design of an enhanced performing, robust controller.

a possible set of   set of models described by a nominal mo-
unfalsified models   del and an allowable model perturbation
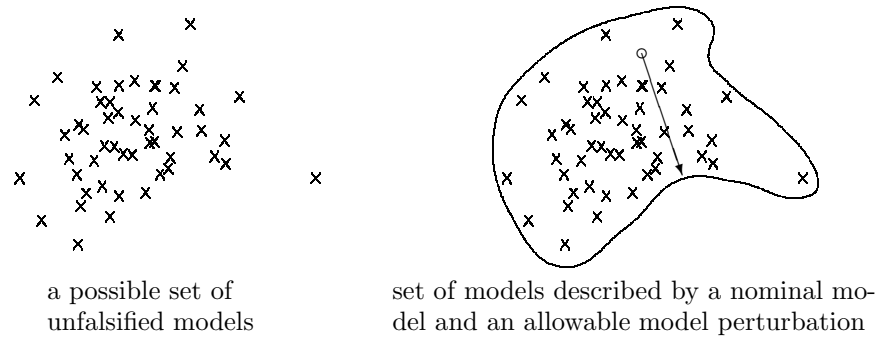
Fig. 1.2: The idea of the construction of a set of models.

Due to the freedom in describing a set of models via a nominal model along with its corresponding model error, the estimation of a model uncertainty set is by far a trivial exercise. To address the suitability of a model uncertainty set for control design, the following aspects have to be considered.

- Model complexity

  Frequently, a relatively low complexity controller is desired due to practical limitations of the hardware used to implement a controller. It is well known that the complexity of both the nominal model and the bound that describes the modelling error will contribute to the complexity of the controller being computed (Boyd and Barrat 1991, Zhou *et al.* 1996, Skogestad and Postlethwaite 1996). Therefore, deliberate simplified modelling of both the nominal model and the accompanying model error is desired in order to formulate a manageable control design problem. A manageable control design problem points to the possibility to design relatively low complexity controllers on the basis of low complexity models.

- Structure of the set of models

  A well known approach to characterize a set of models is by assuming modelling errors or uncertainty in an additive or multiplicative form, see e.g. (Doyle 1979). Additionally, more profoundly defined uncertainties can also be used,

such as the numerator-denominator perturbations in Kwakernaak (1993) or even more structured uncertainties as done in the analysis of Packard and Doyle (1993). Clearly, highly sophisticated modelling error descriptions can be used to construct a set of models thoughtfully. Still, system identification should be able to deliver the information to construct such a set, simply on the basis of observations of the system.

- Conflicting requirements

  The last aspect that has to be taken into account when constructing a set of models is the fact that performance and robustness conditions of a feedback controlled system are conflicting requirements (Doyle *et al.* 1992). This implies that a relatively simple nominal model having a large model error will impose strong robustness conditions at the price to lose performance. A controller being designed will then suffer from decreased performance, whereas the aim is to design an enhanced performance controller. On the other hand, an overall small model error to alleviate the robustness conditions would require the nominal model to be highly accurate and complex. Unfortunately, a low complexity model is required to formulate a manageable control design procedure. During the estimation a trade-off must be made between the accuracy of the nominal model and the corresponding model error.

Clearly, given a selection for the modelling error description, the trade-off can be addressed by estimating a low complexity or approximate nominal model, such that the corresponding modelling error will affect the conflicting robustness conditions as little as possible. In this way, the attention can be focused to finding an approximate nominal model that is said to be relevant for (robust) control design. The role of such a low complexity or approximate nominal model in view of the control application has been recognized and emphasized previously in related topics on modelling, identification, model reduction and adaptive control (Sworder 1966, Farison *et al.* 1967, Åström and Wittenmark 1971, Rivera and Morari 1987, Skelton 1989). The main philosophy in these references is to be aware of the intended use of the model, namely the application of the model in a feedback or closed-loop setting. As such, the desired nominal model and the corresponding model error will highly depend on the controller being used, making the approximate identification and the related control design inseparable (Skelton 1985).

In many applications the controller is still unknown as it needs to be (re)designed on the basis of a nominal model and its corresponding model error. Therefore, the quality of an identified nominal model in view of the control application remains unclear. On the other hand, the design of a controller via a model-based procedure cannot be done without at least the availability of a nominal model. As pointed out in Schrama (1992a), this circular argumentation can be interrupted by an iterative scheme of subsequent identification and model-based control design. Based on the

idea of an iterative scheme, many contributions for control-relevant identification have been developed. A short overview is given in chapter 2 of this thesis, alternatively one can consult the survey papers by Bitmead (1993), Gevers (1993) or Van den Hof and Schrama (1995).

In most of the iterative schemes found in the references listed above, the attention is focused on the estimation of a nominal model and iteratively trying to improve nominal performance specifications. Clearly, the model-based controller must be implemented on the actual system and robustness consideration must be taken into account. Some methods do provide tools to perform at least a stability robustness analysis before implementing the controller on the actual plant, see e.g. (Schrama 1992b, Bayard and Chiang 1993). As pointed out in Section 1.2.3, it is desirable to consider and monitor the performance robustness properties by designing a robust controller directly on the basis of a estimated set of models or model uncertainty set formed by a nominal model along with the corresponding error bound. In this way, it is possible to subsequently design model-based controllers such that the performance improvement of the feedback controlled system can be enforced (Van den Hof *et al.* 1994). As a consequence, performance enhancement can be guaranteed robustly during a step of subsequent system identification and robust control design (de Callafon and Van den Hof 1997). This line of thinking will form the basis of the problem formulation of this thesis.

1.4   Problem Formulation

As pointed out in the previous sections, the problem of designing an enhanced performing and robust controller for an unknown system can be done on the basis of system identification techniques followed by a subsequent robust control design. Although a model will always be a simplified representation of the unknown system, robust control can account for the presence of a modelling error being made. To treat the presence of a modelling error, the system identification procedure is required to estimate both an (approximate) nominal model and an upper bound on the modelling error. In this way a set of models can be estimated, denoted by a model uncertainty set, on which a robust controller is designed. Clearly, such a procedure can be completed only if the appropriate tools are available.

- Tools to estimate nominal models along with an accompanying bound on the model error are needed to estimate a set of models.

- Tools for robust controller design are required to design a controller on the basis of an identified set of models.

Most of these tools are already available in the literature. However, the available identification and control design tools are merged and extended in this thesis

to accomplish a successful result in controlling an unknown system. The success of controlling an unknown system is measured in terms of a performance improvement of the feedback controlled system. As such, the overall problem formulation of this thesis can be summarized as follows.

> Consider an unknown, possibly unstable system that is controlled by a feedback controller. Design a controller that is able to improve the performance of the controlled system by means of subsequent system identification and robust control design.

As mentioned above, the design of a robust and enhanced performing controller is done on the basis of the results of an intermediate step, namely system identification. The system identification should deliver an estimated set of models or model uncertainty set. However, estimating a set of models built up from a nominal model and a bound on the modelling error is not the only requirement for the system identification procedure being used. Due to the conflicting nature of robustness and performance of a feedback system, a trade-off must be made between the accuracy of a nominal model and the corresponding model error. Ignoring this trade-off will squander the possibility to design a robust and enhanced performing controller.

- The model uncertainty set obtained by system identification techniques and used in designing a controller, should allow the design of a robust and high performing controller.

As already mentioned in Section 1.3.3, the above statement points to an identification problem in which a set of models has to be estimated that is suitable for robust control design. Due to the freedom in describing a set of models via a nominal model along with the corresponding model(ling) error, the suitability of the model uncertainty set refers to the following two points:

- A selection for describing the modelling error that is suitable for the design of an enhanced performing and robust controller.

- Estimating a low complexity or approximate nominal model, such that the corresponding modelling error will affect the conflicting robustness conditions as little as possible.

Both items mentioned above will affect the usefulness of the set of models being identified for the design of a robust and enhanced performing controller. Hence, they must be taken into account to set up a proper identification problem.

To improve the general applicability of both the identification and the control design procedure being used, the following additional requirements may be mentioned.

- It is desirable to set up a framework that can handle the identification and control of multivariable unstable systems. Clearly, restricting the identification problem to stable systems would exclude the possibility to deal with unstable systems that can be stabilized very well by means of feedback control.

- The identification procedure should be able to deal with data obtained under closed-loop or controlled conditions. Many engineering processes can not be operated properly without additional feedback control. If indeed a (possibly unstable) system is being controlled, the identification procedure should be able to deal with closed-loop data, in addition to experiments yielding open-loop data.

- Performance specifications used in a model-based robust control design procedure to find enhanced performance controllers should have wide applicability. At the moment, the notion of performance of a controller has still not been characterized precisely. When characterizing the notion of performance, it is desirable to cover a wide range of engineering applications.

Using identification as an intermediate step to solve the above mentioned problem formulation in this thesis is not new. Strongly related problems have also been discussed in the field of adaptive control, see for example Åström and Wittenmark (1989) or Bitmead *et al.* (1990*a*) and the references therein. Compared to the procedure followed in adaptive control[2], the approach followed here mainly focuses on the off-line identification of models and implementation of controllers. Basically, an adaptive controller is created here too, but with the intervention of an engineer, being both a model builder and a control designer.

1.5   Overview of the Thesis

Clearly, the aim of this thesis is to present a framework and tools in order to tackle the verbalized problem formulation presented in Section 1.4. Additionally, the tools being developed will be illustrated on a flexible mechanical positioning system, present in a wafer stepper. For that purpose, this thesis has been divided in 8 chapters. This first chapter has been used to give a brief overview on the developments in the research area of automatic control and system identification, with the aim to formulate the contribution and problem formulation of this thesis.

Chapter 2 is used to present some of the concepts related to feedback systems needed in this thesis. Additionally, a short overview on control-relevant identification will be presented. Based on this overview and using the concepts presented in this

---

[2]Defined as the design of controllers "that can modify its behaviour in response to changes in the dynamics of the process and the disturbances" (Åström and Wittenmark 1989).

chapter, a formal definition of the problem formulation as discussed in Section 1.4 is presented at the end of the chapter.

Chapter 3 is concerned with the problem of identifying models on the basis of experiments obtained under feedback. An identification procedure should be able to deal with such experiments as the presence of feedback is unavoidable in many engineering processes. For that purpose, a fractional approach is introduced that will be used in the sequel of the thesis. This approach is able to deal with feedback or closed-loop data and provides a unified approach to handle the identification of both stable and unstable system.

Chapter 4, 5 and 6 are devoted to the presentation of the main framework of this thesis. The topics discussed in chapter 4 include the characterization of performance and the structure of the set of models built up from a nominal model and a corresponding upper bound. Subsequently, chapter 5 and 6 discuss respectively the robust control design and the identification procedure used to tackle the problem formulation of this thesis.

Chapter 7 is reserved for the practical application of the tools presented in the foregoing chapters and consists of the subsequent identification and robust control applied to the multivariable positioning mechanism present in a wafer stepper system.

Finally, chapter 8 is used to end this thesis and contains the main conclusions and additional remarks, pointing to the possibilities of ensuing research.

2

Identification for Control: Preliminaries and Model-Based Approaches

2.1 Introduction

To start the analysis of the problem formulation mentioned in Section 1.4, first some preliminaries have to be introduced. Especially, the assumptions and definitions associated to the so-called unknown system and performance specification require additional clarification. The notion of an unknown dynamical system is a broad concept and too extensive to present a possible solution of the problem formulation mention in Section 1.4. Furthermore, the notion of performance to characterize the behaviour of a feedback controlled system is still unclear. It is necessary to elucidate both items in more detail as they will be used extensively throughout this thesis.

The problem formulation of Section 1.4 is not entirely new. Several researchers have also focused on similar issues that are linked to the system identification and subsequent model-based control design. As already pointed out in Section 1.3.3, most of the existing techniques make use of iterative schemes to address the inseparability between the identification and the control design.

An overview of most of the preliminary aspects used in this thesis and a discussion of some of the existing so-called iterative schemes of subsequent system identification and (robust) controller design are the main items of this chapter. Most of the preliminaries, definitions and additional assumptions are presented in Section 2.2. Subsequently, the main ideas leading to the application of so-called iterative schemes are outlined in Section 2.3. This section also discusses some of the existing identification and model-based control techniques that are involved in iterative schemes of subsequent system identification and (robust) control design and is ended by a short evaluation. Based on this evaluation, the verbalized problem formulation, as posed in Section 1.4, will be reformulated in a mathematical way in Section 2.5 and the latter

is used to set up the remaining part of the thesis.

2.2   Concepts and Definitions

2.2.1   Main assumptions

To enable the analysis of feedback controlled systems and the identification of an unknown dynamical system, some basic assumptions are needed. The assumptions being introduced here reflect the nature of the unknown dynamical system and the feedback controllers that are being used throughout this thesis.

The unknown dynamical system

Up to now, the notion of a dynamical system has been used to indicate a broad concept that represents many engineering, industrial, biological or chemical processes. Due to this wide applicability of the notion of a system, it is not surprising that it cannot be used directly for specific evaluation purposes. As the main interest in this thesis is concerned solely with the dynamical aspects of a system, the following (verbalized) definition of a dynamical system is adopted from Willems (1991).

Definition 2.2-1 *A dynamical system is defined by a triple $(\mathbb{T}, \mathsf{W}, \mathcal{B})$ that consists of a time axis $\mathbb{T}$ with time instants of interest, a signal space W in which the signals produced by the system take their values and the behaviour $\mathcal{B} \subseteq \mathsf{W}^{\mathbb{T}}$, a family of W-valued time trajectories.*

It should be noted that the above definition of a dynamical system does not discriminate between the inputs and or outputs of a system. Typically the inputs and outputs are associated respectively with the actuator signals and observed signals as depicted in Figure 1.1. Furthermore, the time axis in Definition 2.2-1 can be either discrete or continuous, pointing respectively to a discrete-time or continuous-time dynamical system (Willems 1991). Finally, the dynamical system given in Definition 2.2-1 may exhibit a (non-linear) behaviour, as no restrictions are posed on the actual behaviour of the system.

Clearly, Definition 2.2-1 is a rather general formulation and is still too extensive to have a workable notion of a dynamical system for this thesis. Therefore, additional restrictions and assumptions are posed on the behaviour of the dynamical system. These assumptions have been listed below.

- Firstly, it is assumed that the dynamical system is equipped with a distinguishable set of sensors and actuators. As already mentioned in Section 1.2.1, this opens the possibility to distinguish inputs and outputs of the dynamical system and to construct an automatically controlled system by means of feedback (Willems 1992).

- To clarify the analysis being done in this thesis, the dynamical system is assumed to be finite dimensional, linear and time invariant. Although this assumption can be quite restrictive, many engineering or industrial processes have to be operated around a fixed operating point. Around this operating point the dynamical system may exhibit a behaviour that may be approximated very well by means of a finite dimensional linear and time invariant system.

- Finally, the time axis $\mathbb{T}$ of the dynamical system is assumed to be discrete valued. This assumptions restricts the analysis in this thesis to discrete-time dynamical systems only. However, many of the results presented in this thesis can be carried over to continuous-time dynamical systems too. The motivation to introduce this assumption is due to the fact that data acquisition for both identification and control is usually done by digital signal processing hardware (Chen and Francis 1995). To simplify the analysis involved with sampling continuous-time signals, the effects of the inter-sample behaviour are assumed to be negligible. In fact, this assumption is common once the sampling is done fast enough and is combined with appropriate anti-aliasing filters (Ogata 1987, Åström and Wittenmark 1990).

Given the assumptions listed above, the unknown dynamical system under consideration is denoted by "the plant". The notion of plant is used throughout this thesis to indicate the unknown dynamical system that needs to be identified and controlled subsequently.

Assumption 2.2-2 *An unknown dynamical system to be controlled is assumed to be a discrete-time, finite dimensional, linear and time invariant system having an input u and an output y. The map from the input u to the output y is indicated by notion of plant and is denoted by P_o. Furthermore, the plant P_o is allowed to be multivariable, having m inputs and p outputs.*

Due to the discrete-time nature of the plant P_o, the corresponding input u and output signal y are discrete-time signals, that are assumed to be available either for control or identification purposes. The discrete-time domain dependency of u (or y) is indicated by $u(t)$, where $t \in \mathbb{N}$. In case the signals are obtained by sampling, the integer valued time domain signal t corresponds to $t = k\Delta t$, where $k \in \mathbb{N}$ and Δt denotes the sampling interval. For ease of notation, the sampling interval is set to $\Delta t = 1$, making $t = k \in \mathbb{N}$.

Clearly, the availability of an unlimited amount of possibly undisturbed input and output signals would enable the possibility to reconstruct the plant P_o exactly. From a practical point of view, this possibility is limited due to availability of only finite time, possibly disturbed signals. Summing up, the lack of knowledge of the plant P_o is summarized as follows.

Remark 2.2-3 *The incomplete knowledge of the plant P_o to be controlled points to the availability of only finite time, possibly disturbed, input and output signals. Consequently, the order or McMillan degree of the plant and the (possible multivariable) structure needed to characterize the dynamical behaviour of the plant are assumed to be unknown.*

Due to the linearity assumption on the plant P_o, the principle of superposition (Chen 1984, Callier and Desoer 1991) allows possible additive disturbances acting on the input u or output y to be modelled by an additive disturbance v present in the output signal y. This has also been illustrated in Figure 2.1, in which the signals available for control or identification purposes match the (noise free) input signal u and the possibly disturbed output signal y.
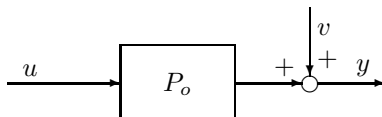
Fig. 2.1: Accessible input u and possibly disturbed output signal y of the plant P_o.

In contrast to the errors-in-variables approach as presented in Deistler and Anderson (1989) or Scherrer *et al.* (1991) where in fact both the input u and output y are allowed to be disturbed, the input signal u is considered here to be exactly known, whereas the output y is assumed to be perturbed by an additive disturbance v. The exact nature and assumption posed on the additive disturbance v is deferred to Section 3.2.1, where the feedback connection of the plant P_o will be discussed.

As a final remark it can be mentioned that the lack of knowledge of the plant P_o may also include the number and location of unstable poles or zeros the plant P_o may have. Although the location of unstable (discrete-time) poles or zeros may highly influence the properties of a feedback controlled plant (Freudenberg and Looze 1985, Looze and Freudenberg 1991, Middleton 1991, Chen 1995), they are not assumed to be known a priori. This knowledge is and should be obtained when estimating models, suitable for the design of a controller.

The feedback controller

In many applications or industrial settings a feedback controller is implemented via a logical switch. In that case, the controller can achieve only two output values and is said to be a boolean controller (Boyd and Barrat 1991). Boolean controllers can for example be found in standard pressure valves or thermal switches like an ordinary thermostat. Although the realization and implementation of such a boolean controller is relatively easy and frequently used, the boolean character may limit the performance of the feedback controlled plant considerably.

By the introduction of special purpose hardware, like digital signal processors, more sophisticated controllers can be implemented. Consequently, the character of a feedback controller does not have to be restricted to a logical switch. In fact, a feedback controller can be an arbitrary dynamical system for which the implementation is limited only by hardware requirements of the signal processor.

Similarly to the nature of the assumptions made on the plant P_o, the feedback controllers to be used in this thesis are designed to be a dynamical system, that maps a control input u_c to a control output y_c. However, in contrast to the knowledge of the plant P_o, the feedback controller is assumed to be known.

Assumption 2.2-4 *A feedback controller is denoted by C and is assumed to be a causal, discrete-time, finite dimensional, linear, time invariant system that maps a control input u_c to a control output y_c.*

As feedback connections of a controller are being considered here only, the notion of "feedback controller" will be abbreviated repeatedly to "controller" or "compensator". Clearly, Assumption 2.2-4 includes the standard components like lead/lag or PID-controllers (Ogata 1990). However, boolean or non-linear controllers will not be considered here. The controller C is restricted to be a discrete-time, finite dimensional, linear and time invariant (FDLTI) controller, as such controllers can and will be designed on the basis of discrete-time FDLTI models obtained by the system identification techniques presented in this thesis. Furthermore, discrete-time FDLTI controllers can be implemented relatively easily using standard digital signal processing hardware and software.

Although both the plant P_o and the controller C are assumed to be FDLTI mappings, there is a subtle difference between P_o and C. As mentioned in Remark 2.2-3, the knowledge of the plant P_o is incomplete and restricted to finite time, possibly disturbed, input and output signals. However, the controller C used to construct a feedback connection does not necessarily have to be unknown. If indeed the controller C is known, this knowledge will be exploited in this thesis in order to estimate models for the unknown plant P_o.

2.2.2   Feedback connections

As mentioned in the previous section, both the plant and the controller are considered to be discrete-time finite dimensional linear time invariant (FDLTI) systems. In this section, first the (standard) notation to represent such FDLTI systems will be summarized. Subsequently, the definition of a well-posed feedback connection will be given.

Representation of FDLTI systems

Let the notation P be used to denote an arbitrary FDLTI mapping. Consequently, P can be represented by its corresponding transfer function (matrix) $P(\xi)$ (Chen 1984, Callier and Desoer 1991). The argument ξ of a transfer function may indicate the Laplace variable s in case of a continuous-time mapping. For the discrete-time mappings considered in this thesis, the argument ξ is replaced by the variable z, being the z-transform (Jackson 1991) of a shift-operator q. The shift-operator q is defined as

$$qu(t) := u(t+1), \; q^{-1}u(t) := u(t-1)$$

where q and q^{-1} respectively denotes the forward and backward shift. Alternative shift operators such as the δ-operator (Middleton and Goodwin 1986) can also be used to describe a discrete-time FDLTI map. However, only the above defined shift operator will be used here.

Due to the analogy between the shift operator q and the z-transform of the shift-operator q, the argument z of a discrete-time transfer function $P(z)$ will be replaced by the variable q frequently. As a consequence, the relation between input and output signals of the plant P_o as depicted in Figure 2.1 and the controller C according to Assumption 2.2-4 can be written in the following (forward) difference equations.

$$
\begin{aligned}
y(t) &= P_o(q)u(t) + v(t) \\
y_c(t) &= C(q)u_c(t)
\end{aligned}
\tag{2.1}
$$

As mentioned before, the controller C used to construct the feedback connection does not necessarily have to be unknown. In case this knowledge is available, a noise free controller output can be constructed. This has been indicated in (2.1) by the noise free signal y_c.

With a slight abuse of notation, the argument of a transfer function and the argument of the signals in (2.1) will be omitted frequently to simplify the notations. According to Assumption 2.2-2, the plant P_o is allowed to be multivariable having m inputs and p outputs. Hence, P_o or $P_o(q)$ will denote the $p \times m$ transfer function matrix of the unknown plant P_o. Supported by the assumption of a distinguishable set of inputs and outputs posed in Section 2.2.1, the transfer functions that are used throughout this thesis are assumed to be proper. For the well known notion of properness of a (discrete-time) transfer function one is referred to Steiglitz (1974), Chen (1984) or Åström and Wittenmark (1990).

For a proper transfer function $P(q)$, an alternative representation by means of state-space realization can be used. Using $x(t)$ to denote the state variable, the discrete-time state-space realization

$$
\begin{aligned}
qx(t) &= \mathsf{A}x(t) + \mathsf{B}u(t), \; x(0) = x_0 \\
y(t) &= \mathsf{C}x(t) + \mathsf{D}u(t)
\end{aligned}
$$

is an alternative representation for the difference equation $y(t) = P(q)u(t)$ by taking $x_0 = 0$ and

$$P(q) = \mathsf{D} + \mathsf{C}(qI - \mathsf{A})^{-1}\mathsf{B} \tag{2.2}$$

Both the transfer function and the state-space representation are used interchangeably here. Due to the ease of notation, preference is given to the representation based on transfer functions. However, most of the operations such as filtering, cascading and interconnection of transfer functions are done on the basis of state-space representations

Finally, it can be mentioned that on the basis of (2.2) a distinction can be made between properness and strictly properness of a transfer function. This has been stated in the following definition.

Definition 2.2-5 *The system P in (2.2) is said to be strictly proper if $\mathsf{D} = 0$, otherwise P is proper and is said to exhibit a feedthrough term D.*

Whether or not a system P or a controller C is equipped with a feedthrough term will play an important role in the interconnection of P and C by means of feedback. Such a feedback connection of P and C is discussed below.

Interconnection and well-posedness

In order to specify the feedback connection of a system P and a controller C, well-posedness of the interconnection of P and C needs to be considered (Boyd and Barrat 1991). The feedback connection of a system P and a controller C is visualized in Figure 2.2.
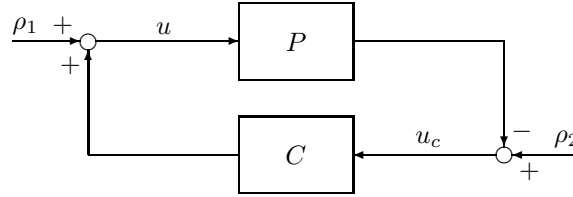
Fig. 2.2: Feedback connection $\mathcal{T}(P, C)$ of system P and a controller C.

Under the condition that a feedback connection of P and C, denoted by $\mathcal{T}(P, C)$, is indeed a well-posed interconnection, the following definition can be given.

Definition 2.2-6 *Let u and u_c be the input signals of respectively P and C. Then $\mathcal{T}(P, C)$ is defined as a well-posed connection of P and C depicted in Figure 2.2 and satisfies*

$$\begin{bmatrix} I & P \\ -C & I \end{bmatrix} \begin{bmatrix} u_c \\ u \end{bmatrix} = \begin{bmatrix} \rho_2 \\ \rho_1 \end{bmatrix} \tag{2.3}$$

where ρ_1 and ρ_2 indicate possible input signals for the feedback connection $\mathcal{T}(P,C)$.

As $u_c = \rho_2 - y$ in Figure 2.2, the above definition refers to a so-called negative feedback connection. The notation $\mathcal{T}(P,C)$ will be used throughout this thesis to describe this negative feedback connection of a system P and a controller C.

Clearly, the well-posedness of the feedback connection $\mathcal{T}(P,C)$ in Figure 2.2 refers to the demand that P and C should at least have compatible sizes in order to be able to set up a well-defined feedback connection. Basically this condition requires the dimension of the signals in (2.3) to be compatible, which will be assumed without mentioning throughout this thesis.

Additionally, well-posedness points to the well-known property that the inverse of the map given in (2.3) should exist and should be proper. As both P and C are proper, the latter condition can easily be reformulated in terms of a condition on the feedthrough terms of P and C.

Proposition 2.2-7 *Let P and C be proper and respectively exhibit a feedthrough term D and D_c according to Definition 2.2-5. Then the feedback interconnection $\mathcal{T}(P,C)$ of Definition 2.2-6 is well-posed if and only if $[I + \mathsf{D}_c\mathsf{D}]$ is non-singular.*

Proof: For well-posedness of $\mathcal{T}(P,C)$, the determinant of $\begin{bmatrix} I & P \\ -C & I \end{bmatrix}$ should not be identically zero. For a proper P and C, this condition is equivalent to the condition of $\begin{bmatrix} I & \mathsf{D} \\ -\mathsf{D}_c & I \end{bmatrix}$ being non-singular (Zhou *et al.* 1996). Application of the matrix inversion lemma (see Appendix A) yields the condition of non-singularity on $(I + \mathsf{D}_c\mathsf{D})$.   □

Many systems are strictly proper and by nature do not exhibit a feedthrough term. With either $\mathsf{D} = 0$ or $\mathsf{D}_c = 0$, the condition mentioned in Proposition 2.2-7 is satisfied trivially. Without loss of generality, it is assumed that Proposition 2.2-7 is satisfied for the feedback connections being discussed here.

The signals ρ_1 and ρ_2 in Figure 2.2 are used to indicate the presence of possible (unknown) signals that act as input signals for the feedback connection $\mathcal{T}(P,C)$. It can be noted that by defining the signals ρ_1 and ρ_2 in (2.3) as follows

$$\begin{aligned} \rho_2 &= r_2 - v \\ \rho_1 &= r_1 \end{aligned} \tag{2.4}$$

a distinction can be made between purposefully applied reference signals $col(r_2, r_1)$ and an unknown noise signal v that may all act as input signals for the feedback connection $\mathcal{T}(P,C)$. With (2.4) it can be observed that the additive noise v is again modelled by an additive disturbance acting on the output y of the system P, similar
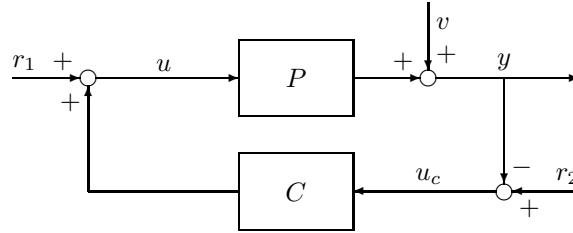
Fig. 2.3: Feedback connection $\mathcal{T}(P,C)$ with distinguishable reference signals $col(r_2, r1)$ and noise signal v.

to Figure 2.1. The signals r_1 and r_2 indicate the reference signals that are (possibly) applied to the feedback connection $\mathcal{T}(P,C)$ by the user and illustrated in Figure 2.3.

Using the distinction between $col(r_2, r_1)$ and v as indicated in Figure 2.3, the following map can be defined.

Definition 2.2-8 *Consider a well-posed feedback connection $\mathcal{T}(P,C)$ where u and y indicate respectively the input and the possibly disturbed output signal of P. Then $T(P,C)$ is defined by*

$$T(P,C) := \begin{bmatrix} P \\ I \end{bmatrix} (I + CP)^{-1} \begin{bmatrix} C & I \end{bmatrix} \tag{2.5}$$

and maps the reference signals $col(r_2, r_1)$ given in (2.4) to the signals $col(y, u)$.

It can be verified that the map $T(P,C)$ in (2.5) reflects the map from the signals $col(r_2, r_1)$ to the signals (y, u) in Figure 2.3. It should be noted that more general feedback connections can be defined by means of sophisticated control architectures such as two degree of freedom controllers (Lunze 1989, Maciejowski 1989, Boyd and Barrat 1991) or interconnections that are based on star products (Doyle *et al.* 1991, Zhou *et al.* 1996). However, the feedback interconnection of Definition 2.2-6 is the most simple (one degree of freedom) representation that is encountered in many typical problems associated to identification and control. Therefore, the (negative) feedback connection $\mathcal{T}(P,C)$ of Definition 2.2-6 and the map $T(P,C)$ given in (2.5) are used throughout this thesis.

2.2.3 Stability and performance

Both stability and performance are relevant when designing and implementing a feedback controller. Although the notion of stability of a feedback connection is a well

known concept, an exact description of the notion of performance may differ in various control applications and is by no means unequivocal. In this perspective, a formal definition of stability and performance of a feedback connection is required here.

In defining stability and performance, the utilisation of a norm or norm function is unavoidable. In this thesis, the notion of stability is said to be equivalent to bounded-input–bounded-output (BIBO) stability which implies that a norm bounded input signal yields a norm bounded output signal (Chen 1984). For ease of notation, $\| \cdot \|_{\mathcal{X}}$ is used to denote such a norm (function) that is defined on an arbitrary normed space $\mathcal{X}$ (Luenberger 1969). The notation $\| \cdot \|$ is used to indicate a norm for which the normed space is yet unknown or irrelevant. As an example of a normed space, the notation $R\mathcal{H}_{\infty}$ is used here to indicate the standard space of all proper, real rational and stable transfer functions (Francis 1987). For a more thorough treatment of norms and the accompanying properties one is referred to Desoer and Vidyasagar (1975), Boyd and Barrat (1991) or Zhou *et al.* (1996).

Internal stability

The notion of stability of a feedback connection used in this thesis refers to the notion of internal stability of feedback interconnected systems (Francis 1987, Boyd and Barrat 1991, Zhou *et al.* 1996). Again using the notation P to denote an arbitrary FDLTI mapping with its corresponding transfer function, internal stability of a feedback connection $\mathcal{T}(P,C)$ reads as follows.

Definition 2.2-9 *The feedback connection $\mathcal{T}(P,C)$ of Definition 2.2-6 is called internally stable if the map from $col(\rho_2, \rho_1)$ to $col(u_c, u)$ is stable.*

The map from $col(u_c, u)$ to $col(\rho_2, \rho_1)$ is given in (2.3). Consequently, the inverse of this map should exist and should be stable in order to satisfy the conditions for internal stability according to Definition 2.2-9. According to Proposition 2.2-7, existence of the inverse of the map in (2.3) is equivalent to well-posedness of the feedback connection $\mathcal{T}(P,C)$. Hence, only stability needs to be verified and the result can been stated in terms of the $T(P,C)$ matrix given in (2.5).

Lemma 2.2-10 *Consider a well-posed feedback connection $\mathcal{T}(P,C)$. Then $\mathcal{T}(P,C)$ is internally stable if and only if $T(P,C) \in R\mathcal{H}_{\infty}$.*

Proof: The proof can be found in Schrama (1992*b*) or Bongers (1994) and uses the following argumentation. Let $H(P,C)$ denote the map from $col(\rho_2, \rho_1)$ to $col(u_c, u)$, so

$$H(P,C) := \begin{bmatrix} I & P \\ -C & I \end{bmatrix}^{-1}$$

then $H(P,C)$ exists (is proper) due to well-posedness of the feedback connection $\mathcal{T}(P,C)$. Using algebraic manipulations, it can be verified that

$$H(P,C) = \left[\begin{array}{cc} (I+PC)^{-1} & -(I+PC)^{-1}P \\ (I+CP)^{-1}C & (I+CP)^{-1} \end{array} \right]$$

making

$$T(P,C) = \left[\begin{array}{cc} -I & 0 \\ 0 & I \end{array} \right] H(P,C) + \left[\begin{array}{cc} I & 0 \\ 0 & 0 \end{array} \right]$$

As a result, stability of $H(P,C)$ is equivalent to stability of $T(P,C)$. Application of Definition 2.2-9 proves the result.   □

From the result mentioned in Lemma 2.2-10 it can be seen that it is necessary (and sufficient) to evaluate all four transfer functions in the $T(P,C)$ matrix of (2.5) to test internal stability of a feedback connection $\mathcal{T}(P,C)$. This can be done by formulating a minimal state-space representation of $T(P,C)$ and checking whether or not the state matrix A is Hurwitz (has stable eigenvalues). Such a state-space realization of $T(P,C)$ can be found in Appendix A of this thesis.

The reasoning to check the stability of all four transfer functions in (2.5) is to detect any unstable pole/zero cancellations that might occur between P and C. Designing a controller that cancels any unstable poles in P is not an internally stabilizing controller for the corresponding feedback connection $\mathcal{T}(P,C)$ (Zhou *et al.* 1996). On the other hand, if either P or C is stable, the stability analysis of $\mathcal{T}(P,C)$ can be simplified and the following well-known results can be obtained.

Corollary 2.2-11 *Let a controller C satisfy $C \in R\mathcal{H}_\infty$. Then $T(P,C) \in R\mathcal{H}_\infty$ if and only if $P(I+CP)^{-1} \in R\mathcal{H}_\infty$*

Proof: See e.g. Zhou *et al.* (1996), pp. 124 or Maciejowski (1989), pp. 56.   □

The result mentioned in the above corollary is in fact the basis for classic control theory to check internal stability on the basis of one closed-loop transfer function matrix only. Dually to the result mentioned above, the following holds.

Corollary 2.2-12 *Let a system P satisfy $P \in R\mathcal{H}_\infty$, then $T(P,C) \in R\mathcal{H}_\infty$ if and only if $(I+CP)^{-1}C \in R\mathcal{H}_\infty$*

Proof: See e.g. Zhou *et al.* (1996), pp. 24.   □

In case both the system P and the controller C are stable, the two corollaries can be combined. In that case, internal stability of the feedback connection $\mathcal{T}(P,C)$ can be checked by investigating the input sensitivity function $(I+CP)^{-1}$ only (Zhou *et al.* 1996).

Norm-based performance

The notion of performance or performance cost is used to indicate how well the overall feedback connection is performing. Usually, the performance of a feedback controlled system is awarded to the feedback controller being used. A controller C is said to be a high performance controller if the matching feedback connection $\mathcal{T}(P,C)$ satisfies high performance requirements. Still, the notion of performance should depend at least on both the controller C and the system P that assemble the feedback connection $\mathcal{T}(P,C)$.

To formalize the performance of a feedback connection of a system P and a controller C, a control objective function will be used. Such a control objective function $J(P,C)$ is used to describe the behaviour of the signals present in the feedback connection $\mathcal{T}(P,C)$. As pointed out above, the control objective function depends at least on both the controller C and the system P. Although the characterization of performance may involve the specification of additional weighting functions or the use of time domain constraints (Boyd and Barrat 1991), the performance essentially depends on the controller C and the system P that assemble the feedback connection $\mathcal{T}(P,C)$. Therefore, such a control objective function can be considered to be a function of its two main arguments P and C and is denoted by $J(P,C)$.

Clearly, in order to characterize and compare the performance cost, a numerical value can be assigned to the control objective function $J(P,C)$. For that purpose, the control objective function $J(P,C)$ is chosen to be an element of a (complete) normed space $\mathcal{X}$, where $\|\cdot\|_{\mathcal{X}}$ denotes the norm defined on $\mathcal{X}$. As a result, the value of the performance cost can be characterized by the value of a norm $\|J(P,C)\|_{\mathcal{X}} \in \mathbb{R}$. A smaller value of the norm $\|J(P,C)\|_{\mathcal{X}}$ thereby indicates an enhanced or improved performance. Summarizing, the following notion of performance is used throughout this thesis.

Definition 2.2-13 *Let P and C form a well-posed feedback connection $\mathcal{T}(P,C)$ and let $\mathcal{X}$ be a complete normed space. Then a control objective function $J(P,C)$ is defined as a function of P and C that is an element of $\mathcal{X}$ and $\|J(P,C)\|_{\mathcal{X}} \in \mathbb{R}$ is used to indicate the performance of $\mathcal{T}(P,C)$. Enhanced or improved performance of $\mathcal{T}(P,C)$ is indicated by a decreased value of $\|J(P,C)\|_{\mathcal{X}}$.*

As mentioned above, a smaller value of the norm in Definition 2.2-13 is used to indicate an enhanced or improved performance. This notion of (enhanced) performance provides the possibility either to compare or to compute enhanced controllers on the basis of the performance cost being used. Once a control objective function has been chosen, for a given system P a so-called optimal controller can be found by minimizing the performance cost

$$\|J(P,C)\|_{\mathcal{X}} \tag{2.6}$$

over a set of admissible controllers that is denoted by $\mathcal{C}$. Such an admissible set of controllers $\mathcal{C}$ is needed in order to ensure that $J(P,C)$ is an element of the complete normed space $\mathcal{X}$ while minimizing (2.6). For example, the set of admissible controllers $\mathcal{C}$ can represent the set of controllers C that will form a stable feedback connection $\mathcal{T}(P,C)$.

Minimizing the performance cost $\|J(P,C)\|_{\mathcal{X}}$ for a given system P gives rise to a controller C_{opt} that is optimal in the sense that the norm-based performance $\|J(P,C_{opt})\|_{\mathcal{X}}$ is being optimized. It should be noted that in general the minimization does not have to yield a unique optimal controller as several solutions may exist. However, the minimization of (2.6) provides a useful tool to come up with an optimally performing controllers for a given system P. To illustrate the notion of performance cost and the existence of such an optimal controller, consider the following example.

Example 2.2-14 *Consider a FDLTI discrete-time system P and controller C that form a stable FDLTI discrete-time feedback connection $\mathcal{T}(P,C)$. A particular choice for a control objective function $J(P,C)$ may be defined as follows*

$$J(P,C) = (I + CP)^{-1}$$

where $(I + CP)^{-1}$ denotes the (input) sensitivity function. Obviously, a stable feedback connection $\mathcal{T}(P,C)$ implies $(I + CP)^{-1} \in R\mathcal{H}_\infty$. As both P and C are finite dimensional linear discrete-time systems, $(I + CP)^{-1} \in R\mathcal{H}_\infty$ is equivalent to $(I + CP)^{-1} \in R\mathcal{H}_2$ (Chen and Francis 1995). As a consequence, both the norms $\|J(P,C)\|_\infty$ and $\|J(P,C)\|_2$ can be used to characterize a performance cost.

Now let the FDLTI discrete-time system P be given by

$$P(q) = \frac{q - 1.1}{q - 0.9} \tag{2.7}$$

and consider the design of a simple constant (optimal) controller $C(q) = K$ using the above mentioned control objective function $J(P,K) = (1 + KP)^{-1}$. Using algebraic manipulations it can be verified that

$$\mathcal{C} := \{C(q) = K \mid K \in (-\frac{19}{21}, 1)\} \tag{2.8}$$

denotes an admissible set of controllers that guarantees $(1 + KP)^{-1} \in R\mathcal{H}_\infty$. Evaluating $\|J(P,K)\|_\infty$ and $\|J(P,K)\|_2$ over this set $\mathcal{C}$ results in the graph depicted in Figure 2.4.

From Figure 2.4 it can be observed that there indeed exists a (unique) optimal static controller K for both the performance costs $\|J(P,K)\|_\infty$ and $\|J(P,K)\|_2$. For the performance cost $\|J(P,K)\|_\infty$ the optimal cost is attained when no controller $(K = 0)$ is applied. This is due to the fact that $\|(1 + KP)^{-1}\|_\infty$ measures the peak of the magnitude of the discrete-time sensitivity function $|(1 + KP(z))^{-1}|$, evaluated
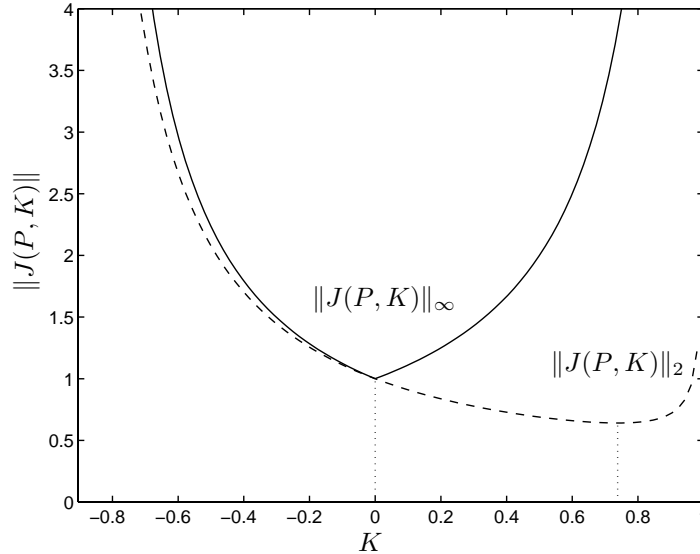
Fig. 2.4: Evaluation of a performance cost over a set of admissible controllers $\mathcal{C}$. In this plot, $J(P, K) = (1 + KP)^{-1}$, where the system P is defined in (2.7) and $\mathcal{C}$ is given in (2.8).

over the unit disk $|z| = 1$. In this example, this peak is minimized when no control is applied. Clearly, to avoid such trivial solutions, the design of a dynamic controller using a more sophisticated control objective function $J(P, C)$ must be exploited. $\diamondsuit$

The characterization of performance given in Definition 2.2-13 limits the analysis done in this thesis to a performance characterization using control objective functions that are norm-based. However, application of a norm-based control objective has wide applicability and may include more sophisticated $\mathcal{H}_\infty$ norm-based performance costs than the one discussed in Example 2.2-14. Well known are the mixed sensitivity problem (Verma and Jonckheere 1984, Kwakernaak 1985, Tsai *et al.* 1992) or the four-block problem as used in McFarlane and Glover (1990) or Bongers and Bosgra (1990). $\mathcal{H}_2$ norm-based objectives such as Linear Quadratic Gaussian (LQG) control (Kwakernaak and Sivan 1972, Anderson and Moore 1990) can also be incorporated, see also the survey paper by Van den Hof and Schrama (1995). In fact, single $\mathcal{H}_\infty$ or $\mathcal{H}_2$ norm-based performance costs that can be represented in a so-called "standard plant description" as used in Doyle *et al.* (1989), Boyd and Barrat (1991) or Zhou *et al.* (1996) cover the characterization of performance given in Definition 2.2-13. Accordingly, in this thesis the attention is focused on $\mathcal{H}_\infty$ norm-based performance costs.

Remark 2.2-15 *The control objective function $J(P, C)$ in Definition 2.2-13 is chosen to be an element of $R\mathcal{H}_\infty$ and the performance cost is denoted by $\|J(P, C)\|_\infty$.*

The motivation to focus on an $\mathcal{H}_\infty$ norm-based performance cost is twofold. Firstly, for a given system P the minimization of (2.6) using an $\mathcal{H}_\infty$ norm can be tackled by using standard techniques for $\mathcal{H}_\infty$ norm model-based controller design, see e.g. (Doyle *et al.* 1989). Clearly, the system P to be controlled is the unknown plant P_o. In order to use such a model-based controller design, a model of the plant P_o must be available. Due to a possible mismatch between the unknown plant P_o and a model used for control design, robustness with respect to stability and performance has to be taken into account during the control design. In an $\mathcal{H}_\infty$ norm-based control design, the robustness issues can be incorporated quite easily, being the second motivation to use an $\mathcal{H}_\infty$ norm-based performance cost. A short evaluation of the aspects associated to robustness are presented in the following section.

2.2.4   Modelling via a set of models

Knowledge of the plant P_o to be controlled is not complete. As pointed out in Remark 2.2-3, incomplete knowledge of the plant is mainly due to the availability of only finite time, possibly disturbed observations of the plant. To represent the incomplete knowledge of the plant P_o, a set of models, instead of one (nominal) model only, can be used to model the plant P_o. To formalize the modelling issues discussed in this section, let $\mathcal{P}$ be used to indicate such a set of models. Typically, such a set of models is built up from a nominal model that is denoted by $\hat{P}$ along with an allowable model perturbation denoted by Δ (Doyle *et al.* 1992).

The composition of such a set $\mathcal{P}$ is useful only if the unknown plant P_o is an element of the set $\mathcal{P}$. Therefore, the allowable model perturbation Δ can be considered to represent the incomplete knowledge of the plant P_o such that $P_o \in \mathcal{P}$. As such, the model perturbation Δ can also account for a possible mismatch or model(ling) error between the plant P_o and the nominal model $\hat{P}$. Such a model error may arise in many practical situations where the nominal $\hat{P}$ is an approximation of the actual plant P_o. Obviously, the presence of only one and exactly known model perturbation Δ would imply exact knowledge of the plant P_o. Therefore, the allowable model perturbation Δ in general is assumed to be unknown, but bounded.

The use of both a nominal model $\hat{P}$ and a model perturbation Δ provides the opportunity to structure a set of models $\mathcal{P}$, so that it can be used in the design of a robust controller. In general, the structure of $\mathcal{P}$ can be characterized by the fairly general framework based on a Linear Fractional Transformation (LFT) (Doyle *et al.* 1991). This LFT framework opens the possibility to rewrite a set of models $\mathcal{P}$ in a standard form, which will be facilitated in this thesis.

On the basis of the LFT framework, each model within $\mathcal{P}$ is represented by a general type of perturbation on the nominal model $\hat{P}$. This general type of perturba-

tion has been visualized in Figure 2.5, where a separate coefficient matrix Q (Zhou *et al.* 1996) is introduced to indicate the global nature of an LFT.

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$
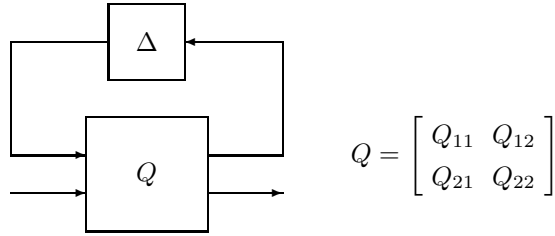
Fig. 2.5: Upper LFT representation of model perturbation.

The entries of the coefficient matrix Q in Figure 2.5 depend on the nominal model $\hat{P}$ and the way in which an allowable model perturbation Δ will affect the nominal model $\hat{P}$. As the allowable perturbation appears at the top of the diagram in Figure 2.5, this general type of perturbation can be represented by a so-called upper LFT $F_u(Q, \Delta)$ with

$$F_u(Q, \Delta) := Q_{22} + Q_{21}\Delta(I - Q_{11}\Delta)^{-1}Q_{12} \tag{2.9}$$

provided that the inverse of $(I - Q_{11}\Delta)$ exists. Opposite to the upper LFT, a lower LFT $F_l(Q, \Delta)$ can be defined provided that the inverse of $(I - Q_{22}\Delta)$ exists and is given by

$$F_l(Q, \Delta) := Q_{11} + Q_{12}\Delta(I - Q_{22}\Delta)^{-1}Q_{21} \tag{2.10}$$

It can be noted that existence of the inverses mentioned above is similar to the well-posedness condition for a feedback connection as discussed in Section 2.2.2. The only difference is the fact that the LFT's mentioned above are equipped with a positive feedback connection instead of the negative feedback connection as depicted in Figure 2.2.

Clearly, if the allowable model perturbation $\Delta = 0$, the LFT of (2.9) should represent the nominal model $\hat{P}$ leading to $Q_{22} = \hat{P}$. However, the remaining entries of Q are still determined by the way the allowable model perturbation affects the nominal model. Additionally, the entries of Q can be used to normalize the bound on the unknown (but bounded) allowable model perturbation Δ (Doyle *et al.* 1991, Zhou *et al.* 1996). As a result, the coefficient matrix Q will contain the necessary information in order to characterize a set of models $\mathcal{P}$. On the basis of this LFT framework, the set of models $\mathcal{P}$ can be formalized as follows.

Definition 2.2-16 *Consider a coefficient matrix Q, where $Q_{22} = \hat{P}$ denotes the nominal model and let Δ be used to denote an unknown but bounded, stable FDLTI map.*

Then a set of models $\mathcal{P}$ is defined as

$$\mathcal{P} := \{P \mid P = F_u(Q, \Delta), \ with \ \|\Delta\|_\infty < 1\} \tag{2.11}$$

where $F_u(Q, \Delta)$ denotes the upper LFT given in (2.9).

It should be noted that Definition 2.2-16 does not include any statement regarding the Q_{11}, Q_{12} and Q_{21} entries of the coefficient matrix Q. Consequently, the way in which the models P are characterized within the set $\mathcal{P}$ given in Definition 2.2-16 is still undefined. However, it can be noted here that it suffices to have an allowable model perturbation Δ that is FDLTI, as both the nominal model $\hat{P}$ and the unknown plant P_o that lie in the set $\mathcal{P}$ are assumed to be FDLTI.

Example 2.2-17 *Typical examples of sets of models that can be captured within the LFT framework include a so-called additive uncertainty set*

$$\mathcal{P}_A(\hat{P}, V, W) := \{P \mid P = \hat{P} + \Delta, \ with \ \Delta \in R\mathcal{H}_\infty, \ \|V\Delta W\|_\infty < 1\} \tag{2.12}$$

for which the coefficient matrix Q of the upper LFT $F_u(Q, \Delta)$ in (2.11) is given by

$$Q = \begin{bmatrix} W & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & I \\ I & \hat{P} \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & I \end{bmatrix}$$

Alternatively, a multiplicative (output) uncertainty set can be formulated by

$$\mathcal{P}_M(\hat{P}, V, W) := \{P \mid P = [I + \Delta]\hat{P}, \ with \ \Delta \in R\mathcal{H}_\infty, \ \|V\Delta W\|_\infty < 1\} \tag{2.13}$$

where the coefficient matrix Q becomes

$$Q = \begin{bmatrix} W & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & \hat{P} \\ I & \hat{P} \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & I \end{bmatrix}$$

In (2.12) and (2.13) the arguments of $\mathcal{P}$ are used to denote that the set depends on the nominal model $\hat{P}$ and some appropriate frequency dependent weighting functions V and W that are used to normalize the $\mathcal{H}_\infty$ norm bound on the allowable model perturbation (Doyle 1979, Doyle et al. 1992). More sophisticated sets of models using an unstructured perturbation Δ can also be defined on the basis of the numerator-denominator perturbations given in Kwakernaak (1993) or the more general (coprime) fractional perturbations discussed in Sefton and Ober (1993) or de Callafon et al. (1996b). $\diamond$

Finally, it should be noted that the requirement on the stability of Δ as mentioned in Definition 2.2-16 or the example mentioned above is a technical condition. This condition is needed in order to be able to apply the small gain theorem (Zames 1963) to evaluate the robustness properties such as stability robustness. The requirement on the stability of Δ can be relaxed for specific sets of models $\mathcal{P}$ (Vidyasagar 1985).

2.2.5   Robustness issues

As pointed out in the previous section, exact knowledge of the plant P_o is not available. This lack of knowledge is replaced by trying to find a set of models $\mathcal{P}$, built up from a nominal model $\hat{P}$ equipped with an allowable model perturbation Δ such that $P_o \in \mathcal{P}$. As a consequence, the stability or performance can (only) be evaluated for the set of models $\mathcal{P}$. If a property can be guaranteed for the set of models $\mathcal{P}$, the controller C is said to be robust with respect to this property. Consequently, for the notion of stability the following definition can be formulated.

Definition 2.2-18 *Let $\mathcal{P}$ denote a set of models as given in Definition 2.2-16. A controller C is said to achieve*

- *nominal stability if $\mathcal{T}(\hat{P}, C)$ is internally stable*

- *stability robustness if $\mathcal{T}(P, C)$ is internally stable for all $P \in \mathcal{P}$*

where the notion of internal stability is given in Definition 2.2-9.

With respect to the notion of norm-based performance as given in Definition 2.2-13 and Remark 2.2-15, a similar definition can be given and is summarized below.

Definition 2.2-19 *Let $\mathcal{P}$ denote a set of models as given in Definition 2.2-16 and consider a performance level $\gamma > 0$. Then a controller C is said to satisfy*

- *nominal performance if $\|J(\hat{P}, C)\|_\infty \leq \gamma$*

- *robust performance if $\|J(P, C)\|_\infty \leq \gamma$ for all $P \in \mathcal{P}$*

where $J(P, C)$ is a control objective function according to Remark 2.2-15.

The value of γ in Definition 2.2-19 can be any positive real value. The smaller the value of γ, the stronger are the performance requirements, see Definition 2.2-13. Consequently, the value of γ can be regarded as the performance level, as indicated in Definition 2.2-19. In this way the notion of nominal or robust performance depends on the performance level γ.

It can be observed from Definition 2.2-19 that lowering the value of γ to impose stronger performance, generally will require (the size of) the set $\mathcal{P}$ to be scaled down in order to satisfy the performance robustness conditions. Consequently, performance and robustness are conflicting requirements (Doyle *et al.* 1992). Furthermore, performance robustness is a stronger requirement than stability robustness. This is due to the fact that stability robustness is required in order to have $J(P, C) \in R\mathcal{H}_\infty$ for all models $P \in \mathcal{P}$.

2.3 Control and Identification

2.3.1 Model-based control design

The performance of a feedback connection has been characterized in Definition 2.2-13 by means of a norm, applied to a control objective function. Due to this norm-based approach, the performance of the feedback connection $\mathcal{T}(P_o, C)$ constructed from a plant P_o and a controller C can be characterized by the numerical value of a norm $\|J(P_o, C)\|$. Ideally, a minimization of the norm-based performance cost $\|J(P_o, C)\|$, similar to (2.6), can be used to design an optimal controller directly on the basis of the (unknown) plant P_o.

As the plant P_o is unknown, measurements from the plant P_o can be used to acquire knowledge of the control objective function $J(P_o, C)$. If indeed (the norm on) $J(P_o, C)$ can be accessed directly on the basis of (time domain) observations coming from either the plant P_o or a feedback connection $\mathcal{T}(P_o, C)$, the possibility to tune or optimize a controller C directly, can be exploited to construct an optimal controller. Basically, this constitutes a model-free tuning of a controller, as no model is formed explicitly[1] on the basis of observations of the plant P_o in order to construct a controller.

This idea is used in Hjalmarsson *et al.* (1994) to perform such a model-free tuning of a controller on the basis of an $\mathcal{H}_2$ norm-based performance specification. The tuning rules of Ziegler and Nichols (1942) can also be regarded as a model-free tuning approach to controller design. Similar ideas to tune PI or PID controllers, without explicitly estimating a (nominal) model, can be found in Åström and Hägglund (1984) or the more recent publication by Voda and Landau (1995). Unfortunately, these approaches provide tools to tune or calibrate only relatively simple feedback compensators. Furthermore, such a model-free approach restricts the tuning of a controller C to those control objective functions $J(P_o, C)$ that indeed can be accessed directly on the basis of (time domain) observations. The control objective function $J(P_o, C)$ will be subjected to noise and finite time approximation if measurements are used to evaluate the performance specification directly. Measuring a perturbed control objective function $J(P_o, C)$ may alter the tuning of the controller C implicitly for which the effect might be unknown on beforehand.

In this model-based control approach, a set of models $\mathcal{P}$ is used both to represent the incomplete knowledge of the plant P_o and to design a robust performing, enhanced feedback controller. Additionally, such a model-based approach has several advantages, compared to a model-free tuning.

- As indicated in Remark 2.2-15, an $\mathcal{H}_\infty$ norm-based performance criterion will be used. In a model-based approach, the design of a controller can be done for

---

[1] Although a FDLTI model is not being estimated explicitly, tuning a controller on the basis of data usually implies an inherent and hidden use of a model.

a wide class of $\mathcal{H}_\infty$ norm-based performance objectives functions $J(P_o, C)$ that cannot be accessed directly by means of (time domain) observations coming from the plant P_o or the feedback connection $\mathcal{T}(P_o, C)$.

- Robustness considerations such as stability or performance robustness can be evaluated and taken into account when designing a controller. This is advantageous if indeed a set of models is being estimated to represent the knowledge of the unknown plant P_o.

- Design trade-offs imposed by linear feedback control (Middleton 1991) can be taken into consideration. Trade-offs present during the design of a controller heavily rely on the properties of the plant to be controlled, see e.g. (Freudenberg and Looze 1985, O'Young and Francis 1985, Freudenberg and Looze 1987, Mohtadi 1990, Chen 1995, Chen and Nett 1995b). The introduction of a model of the plant can help in comprehending these properties while designing a controller.

- Last, but certainly not least, available tools that are model-based (Boyd and Barrat 1991, Zhou *et al.* 1996) can be exploited to design robust controllers. In this way, sophisticated feedback controllers can be designed on the basis of a set of models that take into account design trade-offs and robustness considerations.

Supported by the advantages mentioned above, a model-based approach via the estimation of a set of models $\mathcal{P}$ will be exploited to design a robust and enhanced performing controller for an unknown plant P_o. In order to estimate such a set of models, system identification techniques are used.

2.3.2   System identification

A modelling approach that emphasizes the need of experimental data of a plant in order to characterize any systematic dynamic relations that are present in the plant can be labelled as a system identification technique. In fact, "system identification deals with the problem of building mathematical models of dynamical systems based on observed data from the system" (Ljung 1987). Referring to Figure 2.1, the observations of the unknown plant P_o are reflected by the "data" $\{u(t), y(t)\}$, that consist of the applied input signal u and the (possibly) disturbed output signal y. The additive disturbance v acting on the output of the plant is used to model effects on the data that cannot be described by the input u applied to the plant (Norton 1986, Söderström and Stoica 1989, Ljung and Glad 1994).

The main interest in the usage of system identification techniques in this thesis is the ability to construct models to describe the map from the input u to the output y of the plant P_o and to use this map in the design of a feedback controller. As the plant P_o is a crucial element in the construction of a feedback connection $\mathcal{T}(P_o, C)$,

the main attention is focused on modelling the plant P_o. Still, most of the results
on system identification in this thesis can also be applied to characterize the additive
disturbance v acting on the output y of the plant P_o.

Identification of a set of models

As indicated in the previous sections, a set of models can be used to represent the
incomplete knowledge of the plant P_o. The unavoidable incomplete knowledge of the
plant P_o is due to the availability of only finite time, possibly disturbed, observations
of the plant. Such a set of models can be considered to consist of all models that are
validated by the data (Ljung 1987).

Alternatively, in Smith *et al.* (1997) validation of models is considered to be a
misnomer. It is never possible to validate models solely on the basis of a finite
number of experiments. As a result, a set of models will consist of models that
cannot be invalidated by the data coming from the plant P_o. Although there is a
subtle difference between validated and not invalidated models, a common property
is the ability to state that "there is no evidence in the particular data set we worked
with" (Wahlberg and Ljung 1992) that indicates that the plant P_o is not an element
of the set of models being constructed.

To estimate such a set of models and in order to be able to guarantee that P_o is
an element of the set being estimated, the data $\{u(t), y(t)\}$ coming from the plant
P_o does not suffices. Instead, additional information on the plant P_o (and the data
$\{u(t), y(t)\}$) needs to be introduced in order to be able to formulate a possible set of
models (Wahlberg and Ljung 1991, Hjalmarsson 1993, Hakvoort 1994, de Vries 1994,
Ariaans 1997). Such additional information might include the (approximate) location
of the poles of the plant P_o, initial conditions or assumptions on the nature of the
noise v present on the output y of the plant P_o.

Consequently, a system identification procedure can be formulated as a procedure
that estimates a set of models on the basis of data and additional information coming
from a unknown plant P_o. Schematically, such an identification procedure can be
characterized as a map

$$\text{identification} : \left\{ \begin{array}{c} \text{data } \{u(t), y(t)\} \\ \text{prior assumptions} \end{array} \right\} \mapsto \text{set of models} \qquad (2.14)$$

where the *prior assumptions* are used to indicate any additional prior information
being introduced to estimate the set of models. The exact nature of the prior as-
sumptions being introduced is postponed until chapter 6, where the identification of
a set of models based on the work by Hakvoort (1994) is discussed in more detail. In
the following section, the usage of the set of models for the purpose of designing a
robust controller will be discussed.

Estimating a set of models for control design

The data $\{u(t), y(t)\}$ along with the prior assumptions in (2.14) gives rise to a so-called set of feasible models (Hakvoort 1994). Formally, a set of feasible models $\mathcal{F}$ can be defined as

$$\mathcal{F} := \{P(q) \mid y(t) = P(q)u(t) + v(t), \text{ where} \\ P(q) \text{ and } v(t) \text{ satisfy prior assumptions}\} \tag{2.15}$$

and captures all possible models P that are consistent with the data $\{u(t), y(t)\}$ and the prior assumptions being introduced. If indeed the prior assumptions are correct, then $P_o \in \mathcal{F}$ and the set of models $\mathcal{F}$ reflects the unavoidable incomplete knowledge that is available on the plant P_o.

It would be preferable to estimate the set of feasible models $\mathcal{F}$ meticulously. In this way, the limited knowledge available on the plant P_o in terms of the data $\{u(t), y(t)\}$ and the prior assumptions in (2.14) can be represented in the set $\mathcal{F}$ exactly. Unfortunately, the estimation of such a set $\mathcal{F}$ is impractical as the set can be highly unstructured and is unusable for the design of a robust controller.

In order to formulate a set of models that is suitable for control design, a set of models that is structured according to Definition 2.2-16 would be preferable. Clearly, the estimation of such a set $\mathcal{P}$ that is imposed to be structured can only approximate the set $\mathcal{F}$. Furthermore, the approximation of $\mathcal{F}$ should be done in such a way that $\mathcal{P}$ outerbounds $\mathcal{F}$. Characterizing a set $\mathcal{P}$ that satisfies $\mathcal{F} \subset \mathcal{P}$ also guarantees that $P_o \in \mathcal{P}$.

Outerbounding the set of feasible models $\mathcal{F}$ will enlarge the set of models for which a robust controller needs to be designed. As a result, the design of the controller on the basis of the set $\mathcal{P}$ tends to be conservative. This is due to the fact that the controller should satisfy the robustness conditions as mentioned in Definition 2.2-18 or Definition 2.2-19 for a set of models $\mathcal{P}$ that outerbounds $\mathcal{F}$.

Performance and robustness are conflicting requirements (Doyle *et al.* 1992) that causes the performance of a designed controller to deteriorate in case of a conservative control design. Although outerbounding of $\mathcal{F}$ by a set of models $\mathcal{P}$ is unavoidable, $\mathcal{P}$ should be estimated in such a way, that the performance deterioration of a controller designed on the basis of $\mathcal{P}$ is as small as possible. In that case, an enhanced performing controller can be designed on the basis of a set of models $\mathcal{P}$. Consequently, the identification and construction of the coefficient matrix Q mentioned in (2.11) will play a crucial role in the design of an enhanced performing robust controller.

Approximate identification

In trying to model the plant P_o, a distinction can be made between exact and approximate modelling. In case of exact modelling the attention is focused on trying to model the plant meticulously by trying to capture the dynamical behaviour of plant

P_o exactly. Approximate identification is concerned with system identification problems in which the identification technique is used to find approximative models of the plant P_o. In the latter, the effect of undermodelling, i.e. models that are too simple to describe the plant P_o completely, is of main concern. As a result, a mismatch between the plant and a model is unavoidable.

In most situations, exact modelling of the plant P_o is either too costly or impractical to perform. Especially if models have to be used for control design, deliberate undermodelling is often required. In general, model-based control design procedures tend to yield controllers that have the same complexity as the set of model $\mathcal{P}$ used to compute the controller (Boyd and Barrat 1991, Zhou *et al.* 1996). As a consequence, it is desired that a set of models $\mathcal{P}$ has low complexity to set up a manageable controller design and to find possibly low complexity controllers.

As indicated in Definition 2.2-16, a set of models $\mathcal{P}$ is constructed using an LFT based on a coefficient matrix Q and an unknown, but bounded allowable model perturbation Δ. In light of the possibility to compute low complexity controllers it is preferable to construct a low order coefficient matrix Q as the complexity of Q will directly influence the complexity of the controller (Boyd and Barrat 1991). Low complexity modelling of the coefficient matrix Q requires the entries of Q to be modelled by possibly low order models.

To illustrate this concept one is referred to Example 2.2-17. From this example it can be seen that the nominal model $Q_{22} = \hat{P}$ serves as a "center" for a set of models $\mathcal{P}$. As a low complexity coefficient matrix Q is preferable, the estimation of a low complexity nominal model $\hat{P}$ must be emulated. Additionally, the weighting functions V and W that are used to normalize the $\mathcal{H}_\infty$ norm bound on the allowable model perturbation Δ should be modelled by low complexity stable and stably invertible filters.

2.4   Obtaining Models for Control

2.4.1   Modelling for control

As indicated in the previous sections, inexact knowledge of the plant P_o is replaced by an estimate of a set of models $\mathcal{P}$, of which the plant P_o should be an element. Whether or not such an estimated set of models $\mathcal{P}$ is actually suitable for the design of a (robust) controller that is able to improve the performance of the controlled system, has not been addressed yet.

In this section some of the important issues are discussed, that determine the usefulness of the set of models being estimated for control design purposes. These issues include the choice of the nominal model $\hat{P}$, as it plays an important role in the characterization of nominal properties as indicated in Definition 2.2-18 and Definition 2.2-19. Furthermore the connection between (approximate) modelling and

model-based control design is outlined and possible solutions in terms of iterative schemes are summarized. Finally, some existing techniques are mentioned that can be used to estimate models with uncertainty bounds. These techniques can be used to complete the characterization of a model uncertainty set.

The effectiveness of the existing techniques to find a set of models $\mathcal{P}$ usefull for robust controller design is evaluated in the last section. It is made clear that in the currently available techniques for the estimation of a set of models, the intended (robust) control application of the set $\mathcal{P}$ is hardly taken into account. Acknowledging a link between uncertainty set modelling and the intended robust controller design opens the possibility to address the usefulness of the set of models that is being estimated.

2.4.2   The role of a nominal model

The choice of the nominal model $\hat{P}$ within the set is crucial in constructing the set. A plant P_o that is being controlled successfully by a controller C for which nominal stability or nominal performance on the basis of a nominal model $\hat{P}$ cannot be guaranteed will certainly contest the quality of the nominal model. Alternatively, a controller C that is not able to satisfy nominal stability or a nominal performance specification is not going to be a promising controller for the actual plant. As such, the nominal model $\hat{P}$ plays an important role in both the system identification and the subsequent model-based controller design.

The role of a nominal model has been recognized by several authors in the field of system identification. Some recent contributions in this field can for example be found in the work by Rivera and Gaikwad (1992), Schrama and Bosgra (1993), Hakvoort *et al.* (1994), Zang *et al.* (1995) or Lee *et al.* (1995) and many of the reference listed therein. Although most of the work presented in these reference focuses on the estimation of a nominal model, instead of characterizing a set of models, the approximative nature of the nominal model is being raised and questioned (de Bruyne and Gevers 1994). In the case the nominal model $\hat{P}$ is considered to be a consistent or accurate estimation of the plant P_o (Ljung 1987), the analysis on the quality of the nominal model has been done from a variance point of view (Gevers and Ljung 1986, Hjalmarsson *et al.* 1996). In the more realistic situation where $\hat{P}$ is considered to be an approximation of P_o, the analysis has been limited towards the aspects associated to the bias (Bitmead 1993, Gevers 1993, Van den Hof and Schrama 1995). The bias thereby refers to the difference between the plant P_o and the nominal model $\hat{P}$ that is unavoidable due to the approximative nature of the nominal model.

A common idea in the references listed above is the observation that an approximate identification of a nominal model is allowed, as long as the approximate model $\hat{P}$ takes into account its intended use, namely the design of a high performing controller for the plant P_o. In case a norm function is used to characterize the performance of

a feedback connection, the performance of the controller applied to the plant P_o can be delineated by $\|J(P_o, C)\|$. Even if a controller C is available, a characterization of $\|J(P_o, C)\|$ is not directly possible, as the plant P_o is unknown.

In case $\|J(P_o, C)\|$ can be expressed in terms of data coming from the feedback connection $\mathcal{T}(P_o, C)$, measurements can be used to estimate the performance level $\|J(P_o, C)\|$. In that case, only a (possibly disturbed) estimate of $\|J(P_o, C)\|$ is obtained, while the choice of the control objective function $J(P_o, C)$ is restricted to those performance criteria that actually can be measured directly by data. However, the introduction of a nominal model $\hat{P}$ can be used to bound any norm of the control objective function by the following inequalities.

Proposition 2.4-1 *Consider a plant P_o, a controller C and a nominal model $\hat{P}$ for which $J(P_o, C)$ and $J(\hat{P}, C)$ are well defined control objective functions according to Definition 2.2-13. Then $\|J(P_o, C)\|$ can be bounded by*

$$
\begin{aligned}
\left| \|J(\hat{P}, C)\| - \|J(P_o, C) - J(\hat{P}, C)\| \right| &\leq \|J(P_o, C)\| \\
\|J(P_o, C)\| &\leq \|J(\hat{P}, C)\| + \|J(P_o, C) - J(\hat{P}, C)\|
\end{aligned}
\tag{2.16}
$$

Proof: Application of the triangular inequality (Luenberger 1969) on the norm $\|J(P_o, C)\|$, see also Schrama (1992*b*). □

As indicated by (2.16), a nominal model $\hat{P}$ can be used to formulate both an upper and lower bound for $\|J(P_o, C)\|$, using $\|J(\hat{P}, C)\|$ as a nominal performance cost. A tight upper and lower bound can be found by minimizing $\|J(\hat{P}, C) - J(P_o, C)\|$. For a given controller C, this minimization constitutes a so-called control relevant identification problem (Gevers 1993, Van den Hof and Schrama 1995), where a nominal model $\hat{P}$ is found by minimizing the difference between the norm-based performance of the feedback connections $\mathcal{T}(P_o, C)$ and $\mathcal{T}(\hat{P}, C)$.

Remark 2.4-2 *Consider a fixed controller C and a fixed, but unknown, plant P_o. Then the notion of (nominal) control relevant identification denotes the estimation of a nominal model $\hat{P}$ by minimizing the difference $\|J(P_o, C) - J(\hat{P}, C)\|$ for a particular norm $\|\cdot\|$. The difference $\|J(P_o, C) - J(\hat{P}, C)\|$ denotes the performance degradation due to inexact or approximate modelling.*

Estimating a nominal model $\hat{P}$ that minimizes $\|J(\hat{P}, C) - J(P_o, C)\|$ for a fixed controller C and an unknown plant P_o is an integral part of a so-called norm-based identification and control evaluation[2]. As can be seen from the upper bound on $\|J(P_o, C)\|$ in (2.16), both the nominal model $\hat{P}$ and the controller C can be used to minimize the performance cost $\|J(P_o, C)\|$.

---

[2]Abbreviated and stigmatised to NICE in (Schrama 1992*b*).

2.4.3   Iterative schemes

The minimization of $\|J(\hat{P}, C)\|$ using the controller C is similar to the minimization of (2.6) and corresponds to the design of an optimal controller on the basis of the nominal model $\hat{P}$. Minimizing $\|J(\hat{P}, C) - J(P_o, C)\|$ using the nominal model $\hat{P}$ for a fixed controller C again corresponds to finding a nominal model $\hat{P}$ according to the control relevant identification mentioned in Remark 2.4-2.

Alternately minimizing $\|J(\hat{P}, C) - J(P_o, C)\|$ as an identification problem and $\|J(\hat{P}, C)\|$ as a control design problem provides an iterative scheme of subsequent system identification and control design. By employing such an iterative procedure, it is hoped that $\|J(P_o, C)\|$ decreases in order to find an enhanced performing controller for the (unknown) plant P_o.

The idea of alternately minimizing $\|J(\hat{P}, C) - J(P_o, C)\|$ via system identification and $\|J(\hat{P}, C)\|$ via a control design problem forms a basis for many of the iterative schemes or control relevant identification approaches listed in the literature (Van den Hof and Schrama 1995). Especially the control relevant identification of a nominal model $\hat{P}$ as mentioned in Remark 2.4-2 has gained considerable attention in the field of system identification. In such an iterative scheme, the control relevant identification of a nominal model $\hat{P}$ and the design of a model-based controlled C are applied iteratively with the aim to minimize the performance cost $\|J(P_o, C)\|$ in (2.16). To illustrate the work that has been done in this field, a short overview is presented here. For reasons of clarity, the overview is split in three main parts.

Exact identification and $\mathcal{H}_2$ norm-based control

One of the first steps towards the interaction between identification and control in case of exact modelling has been made in Åström and Wittenmark (1971) and Gevers and Ljung (1986). In Gevers and Ljung (1986) it is mentioned that in the case of exact modelling an $\mathcal{H}_2$ norm-based performance degradation, as mentioned in Remark 2.4-2, can be minimized. For that purpose, a Prediction Error (PE) estimation method (Ljung 1987) can be applied that uses closed-loop experiments and appropriate data filters. As such, the usefulness of closed-loop experiments opposite to open-loop experiments to attain information on the plant P_o was shown to be fruitful.

Unfortunately, the appropriate data filter contains knowledge of the controller yet to be designed and to be used for the closed-loop experiments. To circumvent this circular argumentation, an iterative procedure of identification and model-based control design can be used to update the knowledge of the data filters. One decade later, this conclusion is reconfirmed in the paper by Hjalmarsson *et al.* (1996). Additionally, this paper illustrated very well how the above mentioned circular argumentation can be interrupted by first estimating a model in an open loop way, followed by a model-based control design. Subsequently, the resulting feedback connection can be used for a renewed identification of a model, based on data obtained under feedback.

This approach is proven to outperform an identification based on the same amount of data gathered via open-loop experiments.

The above mentioned results illustrate the usefulness of feedback relevant identification by means of well designed closed-loop experiments. Although the results are powerful, only variance aspects have been analyzed. Furthermore, the following two drawbacks can be mentioned.

- The variance results are valid only in case of exact modelling of the plant P_o to be controlled.

- The variance expressions are valid only in the asymptotic case and based on the assumption of having a infinite number of data points.

The latter is not a severe drawback, as in many practical situations (more than) enough data points are available in order to rely on asymptotic variance expressions, as also motivated in Zhu (1990) and Zhu and Backx (1993). However, the demand on exact modelling of the plant P_o for the variance expression to hold is indefensible in case of identifying a model for control. As indicated in the previous sections, approximate identification is required in many practical situations in order to be able to compute possibly low complexity controllers. Exact identification of a complicated plant P_o inevitably leads to the requirement of estimating a high order nominal model.

Approximate identification and $\mathcal{H}_2$ norm-based control

The results presented in Wahlberg and Ljung (1986) opened the possibility to characterize the bias in case of approximate identification using a least squares PE-method. The paper demonstrates that an $\mathcal{H}_2$ norm-based (implicit) expression can be used to characterize the bias of a (nominal) model $\hat{P}$ being estimated. Furthermore, the bias was shown to be explicitly tunable, provided that a proper model structure is used in order to parametrize the model (Ljung 1987).

Based on the work by Wahlberg and Ljung (1986), in Liu and Skelton (1990) the motivation to use closed-loop experiments from a bias point of view is mentioned. Approximate identification of a model for the purpose of control design is highly facilitated by the usage of closed-loop experiments. In that case, closed-loop experiments are utilized to provide the proper weightings filters in an explicitly tunable bias expression as formulated by Wahlberg and Ljung (1986). A such, closed-loop experiments are regarded to be beneficial for estimating an approximate model suitable for control design (Liu and Skelton 1990).

However, in both Liu and Skelton (1990) and Hakvoort (1990) the conclusion is drawn that the controller to be used for the closed-loop experiments is (yet) unknown and to be designed. From a modelling and model reduction point of view, similar conclusions can also be found in Skelton (1989). In this line of thinking, the possibility to use an iterative scheme of identification and subsequent control design is proposed

in Liu and Skelton (1990). Based on a $\mathcal{H}_2$ norm performance cost criterion, such an iterative scheme can be found in Bitmead *et al.* (1990*b*), while simultaneously similar results are presented in the work by Hakvoort (1990). Primarily, least squares estimation and $\mathcal{H}_2$ norm-based control design techniques are used to set up these iterative schemes.

Based on the work done by Bitmead *et al.* (1990*b*), the so-called Zang-scheme is developed in Bitmead and Zang (1991), Zang *et al.* (1992) or Zang *et al.* (1995). Slight modifications, improvements and applications based on the Zang-scheme can be found in the work by Partanen and Bitmead (1993), Partanen *et al.* (1994*b*) or Partanen (1995). The work by Hakvoort (1990) was further developed in Hakvoort *et al.* (1992) and Hakvoort *et al.* (1994).

The idea of providing the proper weightings filters in an explicitly tunable bias expression during the identification of a (nominal) model has also been used in the work of Rivera *et al.* (1992) or Rivera and Gaikwad (1992). However, in these contributions it is presumed that a prefiltering of open-loop data coming from the plant is able to replace the benefits of closed-loop experiments.

Although the above mentioned methods provide a solution to the interrelation between identification and control design, again some limitations can be mentioned.

- The analysis is limited to bias considerations only and does not include any statements regarding the variance of the models being estimated.

- Although approximate identification is performed, hardly any statements with respect to stability or performance robustness due to the possible mismatch between model $\hat{P}$ and plant P_o have been incorporated.

- The bias results are valid for the asymptotic case, in which an infinite number of data points is assumed to be available.

As mentioned before, the latter does not have to be a severe drawback. Although $\mathcal{H}_2$ norm-based or LQG controller design has proven to be very successful in many applications, LQG controllers may exhibit arbitrarily bad robustness properties (Doyle 1978). Robustness with respect to stability can be regained via a Loop Transfer Recovery (LTR) technique (Stein and Athans 1987). Still, a cautious control design is required to avoid any unpredictable performance deterioration (Partanen *et al.* 1994*a*) as performance robustness is not taken into account. The lack of performance robustness is due to the bias and variance aspects encountered in the modelling of the plant. Although the bias of the model during the approximate identification is tuned towards the intended control application of the model, the model mismatch is not taken into account during the (nominal model-based) control design.

Fractional approaches

The observation that many engineering process cannot be operated without additional feedback due to operational or safety conditions, unleashes the need to estimate models on the basis of experiments obtained in feedback. Additionally, the variance and bias analysis results listed in the references mentioned above indicate the usefulness of closed-loop experiments when estimating models for control.

To deal with the problem of identifying models on the basis of data obtained under feedback, in Hansen and Franklin (1988) a framework is presented that is based on fractional model representations. In this framework a model of the plant is described by the quotient of two stable factors that are parametrized via a so-called dual-Youla Kucera parametrization. This approach is able to deal with the estimation of both stable and unstable plants, operating under feedback controlled conditions (Hansen 1989).

Based on the framework presented by (Hansen and Franklin 1988), an identification of a fractional model representation accompanied by an Internal Model Control (IMC) type of control design is used in Lee *et al.* (1992) or Lee *et al.* (1993*a*) to set up an iterative scheme. Similar to the $\mathcal{H}_2$ norm-based approaches mentioned above, PE methods are used to estimate a nominal model. The iterative procedure of subsequent system identification and IMC are known as the "windsurfer approach", (Lee *et al.* 1993*b*, Lee *et al.* 1995).

Similar to the framework of (Hansen and Franklin 1988) in Schrama (1991) a similar approach is proposed that directly estimates a stable fractional representation of a model. Using the control design procedure of McFarlane and Glover (1990) or Bongers and Bosgra (1990), an iterative procedure of approximate identification of fractional model representations and an $\mathcal{H}_\infty$ norm-based control design is used in Schrama and Van den Hof (1992) and Schrama and Bosgra (1993). Although an $\mathcal{H}_\infty$ norm design opens the possibility to incorporate robustness issues of the controller being designed quite easily (Doyle 1979), only a nominal performance specification is taken into account. This is due to the fact that still only a nominal model is being estimated, instead of a set of models.

The fractional framework provides a unified approach to handle the estimation of both stable and unstable plants, operating under closed-loop conditions. Unfortunately the "windsurfer" approach and the approach mentioned in Schrama (1991) still exhibit the same limitations as mentioned earlier, as only bias considerations are taken into account. Furthermore, the estimation of a model or factorization is limited to the estimation of a nominal model (factorization) only.

The estimation of a set of models (using a fractional approach) would open the possibility to incorporate stability or performance robustness in the controller design. Furthermore, such a fractional approach would be able to deal with closed-loop experiments and provides a unified approach to deal with stable and unstable systems.

Given these possibilities, the fractional model approach is an important constituent in this thesis to address the problem of estimating a set of models on the basis of closed-loop experiments.

2.4.4   Estimating models with bounds

The various iterative schemes discussed in the previous section have emphasized the merit of a good nominal model for controller design. Unfortunately, the availability of only finite time, possibly disturbed, observations of the plant P_o will limit the knowledge of the plant. Hence, even in the case of performing a highly accurate identification, a single nominal model cannot be considered to be a representative of the knowledge available on the plant P_o.

To incorporate robustness in the design of a model-based controller, a set of models $\mathcal{P}$ must be estimated for the design of a robust performing controller, instead of using a nominal model $\hat{P}$ only. Although most of the iterative schemes discussed in the previous section focus on nominal performance specifications and the estimation of a nominal model only, alternative approaches to estimating models for (robust) control can be found in the literature. These approaches are concerned with the estimation of a nominal model $\hat{P}$ such that an explicitly computable bound on the modelling error Δ can be formulated. In this way, a set of models $\mathcal{P}$ similar to (2.11) can be formed. As already indicated in Section 2.3.2, additional prior assumptions on the plant P_o (and the data) need to be introduced in order to be able to formulate a set of models $\mathcal{P}$ such that $P_o \in \mathcal{P}$. Such prior information can include the (approximate) location of the poles of the plant P_o, initial conditions or assumptions on the nature of the noise present on the data coming from the plant P_o.

Estimation of nominal models along with an accompanying error bound can be found in contributions to the so-called identification in $\mathcal{H}_\infty$ (Helmicki *et al.* 1990, Helmicki *et al.* 1991). Methods for estimating a nominal model along with an explicitly computable $\mathcal{H}_\infty$ norm bound on the modelling error Δ can for example be found in Partington (1991), Wahlberg and Ljung (1992), Goodwin *et al.* (1992), Gu and Khargonekar (1992) or Helmicki *et al.* (1993). Estimation of modelling error bounds using a fractional model approach have been reported in van den Boom (1992) and Mäkilä and Partington (1995) or Mäkilä *et al.* (1995).

In the approaches mentioned above, a distinction can be made on the basis of the nature of the assumptions being made or the error bounds being derived. In this way, stochastic or so-called "soft" error bounding approaches can be distinguished in which the prior assumptions have a stochastic nature. The stochastic nature of the assumptions leads to statistical or probabilistic error bounds and examples of "soft" error bounding can be found in e.g. (Ljung 1987, Goodwin *et al.* 1992, Bayard 1992, Rivera *et al.* 1993, Ninness and Goodwin 1995*b*). An alternative branch of model error bounding approaches can be labelled as deterministic or "hard" error bounding. Opposite

to the "soft" error bounding approaches, deterministic assumptions on the data or the plant P_o are used. As a result, non-probabilistic error bounds are derived. Examples can be found in e.g. (Wahlberg and Ljung 1992, Chen *et al.* 1992, Gu and Khargonekar 1992, Helmicki *et al.* 1993, Böling and Mäkilä 1995). A combination of "hard" and "soft" model error bounding may benefit from the advantages associated to both the statistical and the non-probabilistic error bounding procedures as mentioned above. Such combinations can be found in the work of de Vries (1994), de Vries and Van den Hof (1995) or Hakvoort and Van den Hof (1994b), Hakvoort (1994).

Although the above mentioned techniques are able to estimate a nominal model $\hat{P}$ and characterize a bound on the modelling error Δ, most of the methods ignore the merit of a good nominal approximation of the plant P_o for controller design. Consequently, the following limitations can be summarized.

- The selection of a nominal model is primarily directed towards minimization of the worst case modelling error.

- In general, the nominal model is not restricted in complexity and is restricted to be a stable transfer function.

Usually, a performance cost $\|J(P,C)\|$ is not taken into account in the selection of a nominal model $\hat{P}$. Instead, a nominal model is selected for which the worst case $\mathcal{H}_\infty$ norm bound on the modelling error Δ is being minimized (Helmicki *et al.* 1990). In the case that the modelling error is characterized as an additive uncertainty, this implies that the open-loop worst case difference between the (data coming from the) plant and a nominal model will determine the selection of a nominal model, instead of taking the control application into account.

In minimizing the worst case modelling error, the nominal model is usually not restricted in complexity, which might give rise to high order nominal models (Khargonekar *et al.* 1996). The high complexity of the nominal model is mainly due to a linear regression used in the model error bounding (Mäkilä *et al.* 1995). As mentioned before, limited complexity nominal models are needed in order to set up a manageable control design and to find possibly low order controllers. In case of the estimation of a bound on the additive model error, the nominal models are restricted to be stable in order to use an $\mathcal{H}_\infty$ norm bound for the additive model error. In view of the control application such a restriction limits the application of the model error bounding approach, as control is required for unstable plants especially.

Fortunately, some of the error bounding approaches tend to take the control application into account by at least estimating possibly low order, control relevant nominal models along with an additive (or multiplicative) bound on the modelling error Δ (Hakvoort 1994, de Vries 1994, Bayard 1992). As a result, a set of models $\mathcal{P}$ as mentioned in Example 2.2-17 can be formed and used for subsequent robust control design.

Although the merits of a good nominal model have been recognized, the question remains whether or not a set of models based on an additive or multiplicative bound on the modelling error is the most suitable set of models for robust control design. As mentioned before, estimation of a set of models $\mathcal{P}$ involves the outerbounding of the set of feasible model $\mathcal{F}$ as given in (2.15). The choice for the structure of the set of models $\mathcal{P}$, reflected by the coefficient matrix Q in (2.11), is still an open topic. A coefficient matrix Q that is based on an additive or multiplicative modelling error does not have to yield a set of models that allows the design of a high performing, robust controller. Consequently, a stronger motivation on the structure of the set of models $\mathcal{P}$ for the purpose of control design is required.

2.4.5   Evaluation

The various iterative schemes of identification and model-based control design have illustrated the usefulness of estimating models in view of the control application. Especially the use of closed-loop experiments for the purpose of approximate identification of nominal models for control design has been motivated frequently. Furthermore, the estimation of models along with error bounds provide tools to estimate a nominal model that is equipped with a measure to qualify the quality of the model. Unfortunately, both contributions have limitations that need to be dealt with.

Some of the current limitations

- Convergence of an iterative scheme to a specific nominal model $\hat{P}$ and a controller C has not be proven yet.

The main purpose of an iterative scheme, namely alternately minimizing $\|J(\hat{P}, C) - J(P_o, C)\|$ and $\|J(\hat{P}, C)\|$ in (2.16) respectively via a system identification and a model-based control design does not have to convergence. In many practical applications the lack of convergence implies that a finite number of steps of subsequent system identification and model-based control design are performed until either no performance improvement or performance deterioration of the feedback controlled plant P_o is observed.

- A step of subsequent system identification and model-based control design is not guaranteed to attain actual performance improvement of the feedback controlled plant.

In general, gathering (closed-loop) measurements from a plant may be a time-consuming and expensive activity. Subsequent system identification and model-based control design may result in a controller that is inferior to the previous controller being designed, making all effort superfluous. Even if taking measurements from a plant P_o can be done relatively easy and many iterations of subsequent system identification

and control design can be performed, the resulting controller does not have to be optimal.

- Iteratively minimizing $\|J(\hat{P}, C) - J(P_o, C)\|$ and $\|J(\hat{P}, C)\|$ in (2.16) by an iterative scheme does not imply that (the upper bound on) $\|J(P_o, C)\|$ is actually being minimized.

It has been illustrated in Hjalmarsson *et al.* (1995) that minimality of the performance cost $\|J(P_o, C)\|$ is questionable, even if the iterative scheme has converged successfully. Especially in the case of undermodelling, where the controller C is designed on the basis of a biased nominal model $\hat{P}$ that only approximates the unknown plant P_o. As a result, subsequent steps of approximate identification of a nominal model $\hat{P}$ and the design of a model-based controller C will not lead to an optimal controller that actually minimizes $\|J(P_o, C)\|$.

- The identification and controller design in an iterative scheme restricts attention to the nominal model and nominal model-based controller design.

In most of the iterative schemes discussed in Section 2.4.3, only nominal models are being estimated for the design of a controller. Robustness towards stability or performance robustness is taken into account by performing a cautious controller design, allowing only small modifications of the controller (Schrama 1992b). As indicated before, this problem can be circumvented by estimating a set of models $\mathcal{P}$, instead of a nominal model $\hat{P}$ only, and using the set $\mathcal{P}$ in a robust control design procedure.

- Techniques for the estimation of a set of models $\mathcal{P}$ in general do not take into account the intended control application of the set $\mathcal{P}$.

The issue of control relevant identification as mentioned in Remark 2.4-2 has been addressed frequently for the estimation of a nominal model. However, estimating a set of models $\mathcal{P}$ specifically tuned towards a control application that takes into account a performance cost $\|J(P, C)\|$ for each model $P \in \mathcal{P}$ has not been reported in the literature.

Dealing with the limitations

Although the important notions of convergence and optimality are not guaranteed in the iterative schemes listed above, countless numerical simulation examples presented in the literature show promising results, see e.g. Bayard and Chiang (1993), Hakvoort *et al.* (1994), Lee *et al.* (1993a), Schrama and Bosgra (1993) or Zang *et al.* (1995). Successful implementations of iterative schemes of identification and model-based control are less plentiful and have been reported in e.g. Partanen *et al.* (1994a), Partanen and Bitmead (1995) and Schrama and Bosgra (1993). The issues of convergence and

optimality are addressed in these references by performing an adequate number of iterations so that the performance cost $\|J(P_o, C)\|$ has reached a satisfactory level.

Inevitably, it is important to have convergence of an iterative scheme in terms of a performance improvement of the feedback controlled plant. Debatable is the question, whether or not optimality of a restricted complexity controller applied to the (unknown) plant is the key issue. From a practical point of view, it is more valuable to have at least a guaranteed improvement of the performance cost $\|J(P_o, C)\|$, while executing a step of subsequent identification and control design. In this way, any effort put into a step of an iterative scheme is assured to give an improvement of the performance of the feedback controlled plant. Subsequently repeating the executing of such a step of identification and control design is then able to improve the performance of the feedback controlled plant P_o robustly.

To monitor and improve the performance cost $\|J(P_o, C)\|$ of the feedback controlled plant P_o robustly, the design of a model-based controller C should not be focused on a nominal model $\hat{P}$ only. As mentioned before, the incomplete knowledge of the plant P_o must be taken into account, which can be done by estimating a set of models $\mathcal{P}$. Estimating such a set of models and the subsequent design of a robust controller on the basis of the estimated set that improves the performance robustly are the main items in this thesis.

Although the variance and bias aspects are currently treated separately in the estimation of models for control, the estimation of a set of models $\mathcal{P}$ will include both the variance and bias aspects. Basically, the estimated set of models or model uncertainty set represents the limited knowledge of the plant P_o. However, the estimated set of models $\mathcal{P}$ must be suitable for robust control design in order to enable a performance improvement of the feedback controlled plant P_o. For that purpose, the structure and the estimation of the set $\mathcal{P}$ as mentioned in Definition 2.2-16 should be tuned towards the intended application of the set and should take into account the performance cost $\|J(P_o, C)\|$ that needs to be minimized.

2.5 An Approach to Suboptimal Design

2.5.1 Problem formulation reformulated

The performance of a feedback connection $\mathcal{T}(P, C)$ has been characterized by the norm of a control objective function $J(P, C)$ in Definition 2.2-13. Referring to the problem formulation mentioned in Section 1.4, the aim is to find a controller C for an unknown plant P_o that is able to improve the performance of a controller currently implemented on the plant P_o. With the assistance of the notion of performance of Definition 2.2-13 and the restriction on the control objective function $J(P, C)$ mentioned in Remark 2.2-15, the problem formulation given in Section 1.4 can be reformulated as follows.

Problem 2.5-1 *Let a plant P_o and a controller C_i form a stable feedback connection that satisfies the performance specification $\|J(P_o, C_i)\|_\infty \leq \gamma_i$. Design a controller C_{i+1} such that the performance $\|J(P_o, C_{i+1})\|_\infty$ satisfies*

$$\|J(P_o, C_{i+1})\|_\infty \leq \gamma_{i+1} < \gamma_i. \tag{2.17}$$

In the problem formulation mentioned above, the subscripts i and $i + 1$ are used to indicate a repetitive nature.

Remark 2.5-2 *In Problem 2.5-1, the controller C_i corresponds to the known controller that is currently implemented on the unknown plant P_o. The controller C_{i+1} indicates the controller to be designed and yet to be implemented on the plant P_o.*

Correspondingly, γ_i and γ_{i+1} are used to denote the performance level of the controller C_i and C_{i+1} applied to the plant P_o, similar to the notion of performance level γ used in Definition 2.2-19.

In general, a control objective function $J(P, C)$ includes additional weighting functions that are used to represent the performance specifications to be attained. Clearly, in order to be able to compare $J(P_o, C_i)$ and $J(P_o, C_{i+1})$, or respectively the upper bounds γ_i and γ_{i+1} in (2.17), the control objective function $J(P, C)$ should not be altered during the subsequent design of controllers. For reasons of clarity, this has been summarized in the following remark.

Remark 2.5-3 *In Problem 2.5-1, the control objective function $J(P, C)$ is unaltered and is assumed to depend on the variables P and C.*

As mentioned above, the control objective function $J(P, C)$ may include weighting functions or additional design specification. However, the alteration of weighting functions or design specifications that might be present in a control objective $J(P, C)$ are assumed to be fixed during the subsequent design of controllers in Problem 2.5-1.

It should be noted that Remark 2.5-3 does not imply that the control objective function is not allowed to be changed. One can think of an additional iteration around Problem 2.5-1 in order to modify the control objective function $J(P, C)$ by means of the weighting functions that are included in $J(P, C)$. In that case, Problem 2.5-1 can still be applied by considering some (fixed) control objective function $J(P, C)$ that is used a reference to compare the different controller being designed. Therefore, Problem 2.5-1 is considered in an straightforward manner by considering $J(P, C)$ to be fixed according to Remark 2.5-3.

The repetitive nature of Problem 2.5-1 indicates that performing consecutive steps of control design, the performance level γ can be improved progressively. It should be noted that by lowering the value of γ in Problem 2.5-1, the successive design of

controllers C does not necessarily give rise to an optimal controller C that actually minimizes $\|J(P_o, C)\|_\infty{}^3$. However, the approach delineated in Problem 2.5-1 complies with the problem formulation mentioned in Section 1.4, as the newly to be designed controller C_{i+1} is required to improve the performance of the (existing) feedback connection $\mathcal{T}(P_o, C_i)$. For that purpose, a controller found by the the subsequent design mentioned in Problem 2.5-1 is called sub-optimal.

The same philosophy of Problem 2.5-1 is used also in the general framework of $\mathcal{H}_\infty$ control design to compute sub-optimal controllers, see e.g. Doyle *et al.* (1992) or Zhou *et al.* (1996). In this approach, a controller can be computed that is guaranteed to satisfy an upper bound similar to (2.17). Next, a sub-optimal controller is found by lowering the upper bound on $\|J(P_o, C)\|_\infty$. In this way, a sub-optimal controller can be computed, provided that the plant P_o is known. In case the plant P_o is unknown, the $\mathcal{H}_\infty$ norm on the control objective function $J(P_o, C)$ in Problem 2.5-1 cannot be computed.

As indicated in the previous sections, measurements from the plant P_o can be used to acquire knowledge of the plant. The measurements are used to estimating a set of models $\mathcal{P}$ that represents the knowledge available on the plant. In this way, Problem 2.5-1 is recast into a model-based procedure. By combining system identification and robust control design techniques, in the next section an approach is presented to monitor and improve the performance cost $\|J(P_o, C)\|_\infty$ progressively via a model-based approach.

2.5.2   A model-based procedure

In formulating a model-based procedure to tackle Problem 2.5-1, basically three different items can be distinguished. The first item is associated to determining the performance level of the controller C_i currently implemented on the plant P_o,

1. Firstly a procedure must be found to analyze the upper bound γ_i for $\|J(P_o, C_i)\|_\infty$, where C_i denotes the controller currently implemented on the plant P_o.

In case $\|J(P_o, C)\|$ can be expressed in terms of data coming from the feedback connection $\mathcal{T}(P_o, C)$, measurements can be used to estimate the performance level $\|J(P_o, C)\|$. Unfortunately, only a (possibly disturbed) estimate of $\|J(P_o, C)\|$ is obtained, as the performance cost $\|J(P_o, C)\|_\infty$ will be subjected to noise and finite time approximation if noisy observations are used to evaluate the performance specification. Additionally, the choice of the control objective function $J(P_o, C)$ is restricted to those performance criteria that can be measured directly by data.

---

[3]Convergence of the sequence $\{\gamma_i\}$ to a value $\gamma_{opt} \in \mathbb{R}$ does not have to imply $C_i \to C_{opt}$ for some controller C_{opt} that is optimal.

Using a model-based approach, where the knowledge of the plant P_o is represented by a set of models $\mathcal{P}_i$, evaluation of $\|J(P_o, C_i)\|_\infty$ can be done by evaluating $\|J(P, C_i)\|_\infty$ for all $P \in \mathcal{P}_i$. In this way, evaluation of the performance level is done on a model-based level, instead of data-based level and enables the specification of a control objective function $J(P_o, C)$ that cannot be measured directly on the basis of data. The set of models $\mathcal{P}_i$ can also be used to design and possibly improve the feedback controller.

2. Secondly, the synthesis of a controller C_{i+1} that satisfies (2.17) must be formulated.

Clearly, without the guarantee that the newly designed controller is able to improve the upper bound γ_i of the performance cost $\|J(P_o, C_i)\|_\infty$, the design of the controller C_{i+1} becomes superfluous. However, the set of models $\mathcal{P}_i$ being identified previously can be used to design a robust controller C_{i+1} that is able to satisfy

$$\|J(P, C_{i+1})\|_\infty \leq \gamma_{i+1} < \gamma_i \tag{2.18}$$

for all $P \in \mathcal{P}$. With $P \in \mathcal{P}_i$, the set of models $\mathcal{P}_i$ is not only used to design a new and improved feedback controller C_{i+1}. The set of models $\mathcal{P}_i$ is also used to guarantee (2.17) before implementing C_{i+1} in a feedback connection with the plant P_o.

3. Finally, once a new controller C_{i+1} is designed and implemented on the plant P_o, knowledge of an upper bound on the performance cost $\|J(P_o, C_{i+1})\|_\infty$ can be updated.

As indicated above, the design of C_{i+1} on the basis of the set of models $\mathcal{P}_i$ guarantees the bound mentioned in (2.18), even before implementing the controller C_{i+1} in a feedback connection with the plant P_o. Now that the controller C_{i+1} is implemented, the so-called *a priori* information of (2.18) can be updated with so-called *a posteriori* information by taking new data from the feedback connection $\mathcal{T}(P_o, C_{i+1})$.

This new data coming from the feedback connection $\mathcal{T}(P_o, C_{i+1})$ can be used to update or renew the set of models $\mathcal{P}_i$ to $\mathcal{P}_{i+1}$. Basically, the new feedback connection $\mathcal{T}(P_o, C_{i+1})$ is used to update the information available on the plant P_o and this information is represented in a new set of models $\mathcal{P}_{i+1}$. Furthermore, as indicated in Section 2.4.5, a set of models is suitable for control design if the set takes into account the closed-loop operation of the models within the set and the control objective function $J(P, C)$ used to evaluate the performance of the feedback connection $\mathcal{T}(P, C)$. As the controller C_i applied to the plant P_o has been updated to C_{i+1}, the set of models $\mathcal{P}_i$ needs to be updated to $\mathcal{P}_{i+1}$, using the knowledge of the new feedback connection $\mathcal{T}(P_o, C_{i+1})$ and control objective function $J(P, C_{i+1})$.

In line of the two items mentioned above, this can be labelled as a third item which highly resembles the identification of a set of models $\mathcal{P}_i$ as mentioned for the

first item. Once a new set of models $\mathcal{P}_{i+1}$ is estimated, again a new controller can be synthesized (second item) in order to improve the upper bound on the performance cost progressively. In this perspective, the identification of the set of models in the first item can be considered as an initialization of the model-based procedure.

Clearly, a model-based approach using the estimation of a set of models can be exploited to accomplish the requirements mentioned above. In accordance with the three items mentioned above, the following (model-based) procedure can be formulated.

Procedure 2.5-4 *Let a plant P_o and a controller C_i form a stable feedback connection. To evaluate $\|J(P_o, C_i)\|$, consider the following step.*

   1. Use experimental data and prior information on the data or the plant P_o to estimate a set of models $\mathcal{P}_i$ such that γ_i in

$$\|J(P, C_i)\|_\infty \leq \gamma_i \ \forall P \in \mathcal{P}_i \tag{2.19}$$

   is minimized, while $P_o \in \mathcal{P}_i$.

Subsequently, consider the following steps of control design and system identification

   2. Design a controller C_{i+1} subjected to the condition

$$\|J(P, C_{i+1})\|_\infty \leq \gamma_{i+1} < \gamma_i \ \forall P \in \mathcal{P}_i \tag{2.20}$$

   3. Use (new) experimental data and prior information on both the data and the plant P_o to estimate a set of models $\mathcal{P}_{i+1}$ such that $P_o \in \mathcal{P}_{i+1}$ and subjected to the condition

$$\|J(P, C_{i+1})\|_\infty \leq \gamma_{i+1} \ \forall P \in \mathcal{P}_{i+1} \tag{2.21}$$

The formulation of Procedure 2.5-4 points to a rather general procedure to generate a sequence of model-based controllers that will satisfy (2.17). Within this procedure, step 2 reflects the design of a robust controller. Both step 1 and step 3 are concerned with the identification problem of estimating a set of models $\mathcal{P}$. Referring to the problem formulation mentioned in Section 1.4, it can be observed that step 3 is not required. However, step 3 provides the opportunity to continue the subsequent steps of identification and control design, which is explained below.

Again the subscripts i and $i + 1$ are used to denote the repetitive nature of Procedure 2.5-4. Similar to Remark 2.5-2, the subscripts are used likewise to distinguish the set of models being estimated.

Remark 2.5-5 *In Procedure 2.5-4, the set $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ correspond to a set of models estimated on the basis of prior information on the plant P_o and data obtained respectively from the feedback connections $\mathcal{T}(P_o, C_i)$ and $\mathcal{T}(P_o, C_{i+1})$.*

Hence, both step 1 and step 3 constitute a similar identification problem to construct a set of models $\mathcal{P}$. The corresponding identification problems differ only in (the knowledge of) the feedback controller C being implemented on the plant P_o, which is also used to gather data from the feedback connection $\mathcal{T}(P_o, C)$. However, there is an other subtle difference which motivates the distinction being made between step 1 and step 3 in Procedure 2.5-4.

Remark 2.5-6 *In Procedure 2.5-4, the estimation of the set $\mathcal{P}_i$ and the usage of (2.19) in step 1 denotes a performance assessment test to evaluate $\|J(P_o, C_i)\|_\infty$ for initialization purposes. Subsequently, (2.20) and (2.21) constitute respectively a controller and modelling validation test in order to enforce (2.17) in Problem 2.5-1.*

The identification problem present in step 1 of Procedure 2.5-4 can be viewed as an initialization[4] to the design of an improved controller C_{i+1} in step 2. Subsequently, the control design is followed by a renewed identification in step 3 to update the knowledge of the plant P_o using the newly created feedback connection $\mathcal{T}(P_o, C_{i+1})$. After the initialization of step 1, repeatedly executing step 2 and step 3 will provide a design procedure in which the upper bound γ_i on a predetermined performance cost $\|J(P_o, C_i)\|_\infty$ can be reduced progressively. A possible evaluation of the subsequent steps in Procedure 2.5-4 has been depicted in Figure 2.6.
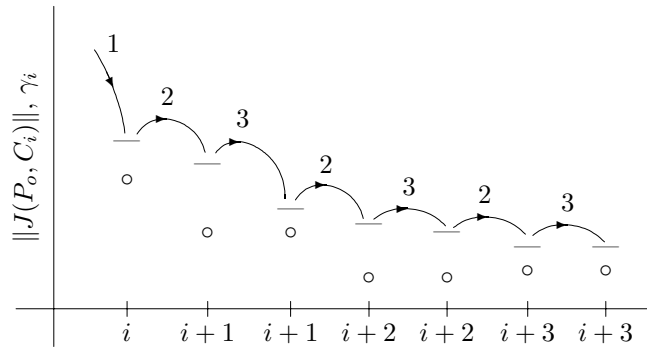
Fig. 2.6: Possible progress of subsequent steps in the model-based procedure mentioned in Procedure 2.5-4, where γ_i and $\|J(P_o, C_i)\|_\infty$ respectively are denoted by — and ∘.

It can be observed from Figure 2.6 that step 1 is used as an initialization. The consecutive design of a controller in step 2 is able to lower the upper bound γ_i and modifies $\|J(P_o, C_i)\|_\infty$. The renewed identification of a set of models $\mathcal{P}_{i+1}$ in step 3 is used to lower the upper bound γ by updating the information on the plant P_o, while

---

[4]This may also include the estimation of set $\mathcal{P}_i$ based on open-loop consideration, where $C_i = 0$.

keeping the controller implemented on the plant P_o unaltered. Although the identification in step 1 and step 3 are basically the same, they serve a different purpose. Still, both the identification problems take into account a control objective function $J(P, C)$ and will constitute a control relevant identification of a set of models $\mathcal{P}$.

Both the identification problems in step 1 and step 3 are subjected to the condition that the plant P_o should be an element of the set of models being estimated. Verifying whether or not $P_o \in \mathcal{P}$ can be viewed as a model (in)validation problem (Smith and Doyle 1992). In such a model invalidation one tries to invalidate $P_o \in \mathcal{P}$ for a given set of models $\mathcal{P}$ by searching for an experiment obtained from the plant P_o, that could not have been produced by any of the models within the set $\mathcal{P}$, see e.g. (Smith 1995, Smith *et al.* 1997). Model invalidation techniques can be applied to verify whether or not $P_o \in \mathcal{P}$, once a set of models $\mathcal{P}$ has been estimated.

Similar to (Hakvoort 1994), in this thesis prior information on the plant P_o is introduced during the estimation of a set of models $\mathcal{P}$. To accomplish $P_o \in \mathcal{P}$ in Definition 2.2-16, additional information on the data and the plant P_o is required. This is due to the fact that it is not possible to verify or guarantee that $P_o \in \mathcal{P}$ solely on the basis of finitely many input and output observations of the plant P_o (Ljung 1992, Hjalmarsson 1993). One is referred to chapter 5 for a detailed discussion on the prior information being used to estimate a set of models. If indeed the prior information is consistent with the plant P_o, the set of models $\mathcal{P}$ will contain the actual plant P_o (Hjalmarsson 1993, Hakvoort 1994). Therefore, model invalidation tools to verify whether or not $P_o \in \mathcal{P}$ will not be considered in this thesis.

Finally it can be mentioned that a monotonically decreasing upper bound γ_i on the performance cost $\|J(P_o, C_i)\|_\infty$ as depicted in Figure 2.6 is due to the conditions (2.20) and (2.21) mentioned in Procedure 2.5-4. Although (2.20) and (2.21) guarantee a monotonically decreasing upper bound γ_i, the performance cost $\|J(P_o, C_i)\|_\infty$ may still be fluctuating below the upper bound, as can be observed from in Figure 2.6. However, gradually decreasing the upper bound γ_i will ameliorate the performance eventually.

The conditions (2.20) and (2.21) reflect respectively a controller and a modelling (closed-loop) validation test in order to enforce (2.17). A controller must be designed that should satisfy the closed-loop validation test (2.20) before implementing it on the plant. After that, a set of models should be estimated that should satisfy the validation test (2.21). A schematic overview of these validation tests in Procedure 2.5-4 is depicted in Figure 2.7.

As can be seen from Figure 2.7, a distinction can be made between the situations in which a controller has not and has been implemented on the unknown plant P_o. Accordingly, the notation $\mathcal{T}(P_o, C_i)$ and $\mathcal{T}(P_o, C_{i+1})$ on the right hand side of Figure 2.7 is used to denote the data and additional information coming from the feedback connection of the plant P_o and the corresponding controllers C_i and C_{i+1}. From Figure 2.7, the initialization of step 1, the control design of step 2 and the sub-
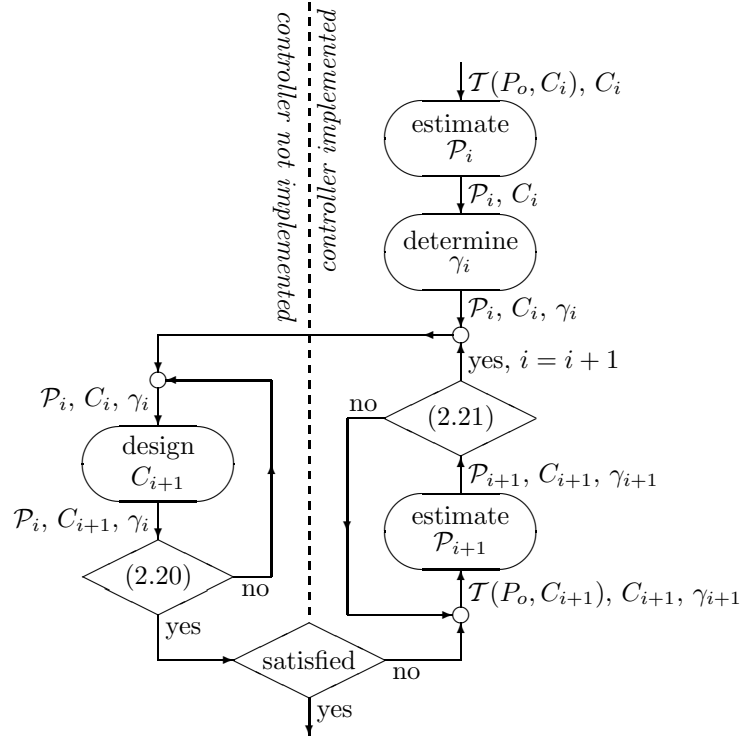
Fig. 2.7: Schematic flow chart of the three steps and the closed-loop controller (2.20) and modelling (2.21) validation test in Procedure 2.5-4.

sequent identification in step 3 can be distinguished. In all steps, the predetermined control objective function plays an important role. The circular in the lower part of the figure indicates the repetitive nature of Procedure 2.5-4.

2.5.3 Main ingredients

Although the approach sketched in Procedure 2.5-4 is fairly general and somehow trivial, it does provide a framework for the design of a sequence of controllers that yield a monotonic and non-decreasing sequence of the upper bound on $\|J(P_o, C)\|_\infty$. A similar idea was proposed also in Bayard *et al.* (1992), but the results were limited to a set of models $\mathcal{P}$ described by weighted open-loop additive perturbations on the nominal model as mentioned in (2.12) and a control objective function based on a (weighted) sensitivity function of the closed-loop system. However, the motivation to use a set of models $\mathcal{P}$ that is described by open-loop perturbations on a nominal model remains unclear. Obviously, to provide a feasible procedure for handling

Procedure 2.5-4, the choice for the structure of the set of models $\mathcal{P}$ should be addressed (Van den Hof *et al.* 1994). Summing up, the following main ingredients need additional clarification in order to use the model-based approach of Procedure 2.5-4.

- Choice for the control objective function $J(P, C) \in R\mathcal{H}_\infty$. It plays a crucial role in the closed-loop validation tests and the way the controller is to be designed. Therefore a specification of the control objective function is needed.

- Choice for the structure of the set of models $\mathcal{P}$. By characterizing the mismatch between a nominal model $\hat{P}$ and the plant P_o within a set $\mathcal{P}$, one should be able to evaluate the closed-loop performance assessment test (2.19) and the closed-loop validation test of (2.20) and (2.21) in a non-conservative way.

- Identification procedure to estimate a set of models $\mathcal{P}$. It should take into account the control objective function that is used to evaluate the set $\mathcal{P}$ in (2.19) and (2.21).

- Robust control design method. The design of a controller on the basis of a set of models $\mathcal{P}$ in step 2 is needed to ensure (2.17).

In light of Procedure 2.5-4, a discussion of the above mentioned items will form the basis for the remaining part of the thesis.

2.5.4   Contributions of this thesis

A completion of the items mentioned in Section 2.5.3 is the main contribution of this thesis. Summarizing, this thesis contributes in merging the results available in the fields of system identification and robust controller design to constitute a framework for a model-based approach to the design of robust and enhanced performing controllers for a unknown plant, on the basis of observations coming from the plant. The emphasize of this thesis lies on the field of system identification, for which new results on closed-loop identification and model uncertainty set estimation for robust control design are presented.

In this thesis, a motivation for the structure of set of models or model uncertainty set $\mathcal{P}$ is formulated that takes into account the closed-loop operation of the models within the set and the control objective function $J(P, C)$ used to evaluate the performance of the feedback connection $\mathcal{T}(P, C)$. Additionally, for such a model uncertainty set the attention is focused on finding a low complexity representation of the coefficient matrix Q in (2.11) allowing the design of possibly low order controllers.

Existing results on model uncertainty estimation and robust controller design techniques will be used. However, to improve the suitability of a model uncertainty set obtained by system identification techniques, the issue of model complexity of both the nominal model and the model uncertainty bound is addressed in this thesis. New

results are presented on the estimation of reduced order complexity models to model a possibly unstable plant. The structure of the model uncertainty set is based on coprime factor perturbations. This structure has shown to be particularly useful for identification purposes and is applicable to data obtained under closed-loop or controlled conditions.

2.5.5   Outline of the remainder

In Section 2.5.2 a solution to the problem formulation in Section 1.4 was presented. This solution is summarized in Procedure 2.5-4 and in order to complete the main ingredients of this procedure, the following 6 chapters have been given the following contents.

chapter 3

In this chapter the identification on the basis of closed-loop data coming from the feedback connection of the plant P_o and a controller C will be discussed. Especially the problem of estimating approximate models on the basis of closed-loop observations receives extra attention. Some of the existing techniques are summarized and evaluated in this chapter.

chapter 4

An identification based on fractional model representations (Van den Hof *et al.* 1995, de Callafon and Van den Hof 1995*b*) is presented in this chapter. The fractional approach will play an important role in the identification on the basis of closed-loop data and the construction of the set of models $\mathcal{P}$. It will be shown that the fractional approach is able to deal with the estimating approximate models on the basis of closed-loop observations, as discussed in the previous chapter.

chapter 5

The choice and motivation for the control objective function $J(P,C) \in R\mathcal{H}_\infty$ and the structure of the set of models $\mathcal{P}$ used in Procedure 2.5-4 are explained in more detail in this chapter. Given the control objective function $J(P,C)$ and the structure of the set of models $\mathcal{P}$, the robust control design problem of step 2 in Procedure 2.5-4 is also discussed in this chapter. The results are in line with the $\mathcal{H}_\infty$ norm-based control design as presented in Zhou *et al.* (1996).

chapter 6

This chapter is completely devoted to the identification of a set of models $\mathcal{P}$ as defined in chapter 4. The estimation of a set of models $\mathcal{P}$ is used either in step 1 or step 3 in Procedure 2.5-4. It is shown that the set of models $\mathcal{P}$ is estimated by the separate estimation of a nominal model $\hat{P}$, followed by a estimation of the allowable model perturbation Δ. Both the estimation of $\hat{P}$ and Δ are done within the framework of fractional model representations. In accordance with the results presented in chapter 3, this provides a framework to handle the approximate identification of possibly

unstable plants on the basis of closed-loop observations (de Callafon and Van den Hof 1997).

chapter 7

The application of the model-based approach discussed in Procedure 2.5-4 is illustrated in this chapter on a multivariable, three degree of freedom positioning mechanism in a wafer stepper. The results in this chapter include the estimation of a set of models $\mathcal{P}$ on the basis of closed-loop observations as discussed in chapter 3 and chapter 6. Furthermore, the control design of presented in chapter 5 is applied and it will be shown that an improved feedback controlled positioning mechanism is attained by subsequent identification and robust control design.

chapter 8

Finally, the last chapter of this thesis is used to present concluding remarks. Additional recommendations for further research with respect to topic of control relevant identification can also be found in this chapter.

Part II

Closed-Loop Identification

3

Identification in the Presence of a Feedback Controller

3.1 Synopsis

The system identification in this thesis is primarily directed towards the purpose of designing a model-based feedback controller. Estimating models for subsequent control design typically involves the identification of a system that is currently operating or aims at operating under controlled or feedback conditions. Therefore, the usage of so-called closed-loop experiments will be an important constituent for the system identification being used here.

The reasoning for the presence of a controlling device while performing experiments on the system is twofold. Firstly, many systems exhibit a poorly damped or even unstable behaviour. In order to obtain reliable data of the system in a limited amount of time, performing experiments for identification purposes can only be done in a controlled or closed-loop setting. Even for stable systems the presence of a controller is frequently required due to unremitting safety and production requirements. Without the presence of a controlling device, safe and reliable experiments cannot be performed.

The second mainspring for performing experiments on a system in the presence of a controlling device is the intended usage of the model. As indicated in Section 2.4.3, iterative schemes of subsequent identification and model-based control design are frequently used to address the inseparability of identification and control. In such an iterative scheme, closed-loop experiments are crucial to obtain an estimate of an approximate model that should model the dynamical behaviour of the system relevant for control design. As such, the usage of closed-loop experiments is motivated by the fact that the dynamics of the unknown system that exhibits in the presence of a controller is more relevant than the dynamics of the system operating in an open-

loop fashion (Schrama 1992a).

The problems associated with estimating models on the basis of closed-loop experiments is the main focus of this chapter. First, the general framework of Prediction Error identification is outlined in Section 3.2. Subsequently, it is shown under which conditions the closed-loop identification problem arises and special attention will be given to the so-called closed-loop approximate identification problem. Some of the well-known and recently introduced procedures to deal with the (approximate) closed-loop identification problem are summarized and evaluated in Section 3.3. It is shown in this chapter that these procedures can be categorized in three main stream approaches. The chapter is ended by a short summary in Section 3.4.

3.2   Identification and Closed-Loop Data

3.2.1   Closed-loop observations

The closed-loop identification problem refers to a problem of estimating models on the basis of closed-loop experiments that have been obtained in the presence of a feedback controller (Ljung 1987, Söderström and Stoica 1989). Although this is a very general description of the closed-loop identification problem, it covers the main bottleneck: the presence of feedback, while gathering data from the unknown plant. As pointed out before, the presence of feedback is favourable in many practical situations where an (unstable) plant has to be kept within a specified operating range to obtain observations for identificational purposes. However, such a closed-loop setting might introduce some additional unfavourable features that have to be dealt with.

In addition to the identifiability considerations of a feedback controlled plant (Söderström and Stoica 1989), the presence of feedback may highly influence the application and outcome of a system identification procedure being used (Ljung 1987). Especially in those situations where system identification is used to find a model that approximates the plant P_o, the presence of feedback deserves special attention. As this situation is of interest here, a more refined definition of the closed-loop identification problem will be given later in this section.

In order to discuss the problems that are associated to an identification based on closed-loop experiments, the feedback connection $\mathcal{T}(P_o, C)$ of the plant P_o and a feedback controller C will be considered here. In this feedback connection $\mathcal{T}(P_o, C)$, the variable C is used to denote any stabilizing feedback controller that is (currently) being implemented on the unknown plant P_o. On the basis of the feedback connection $\mathcal{T}(P, C)$ given in Definition 2.2-6, the feedback connection $\mathcal{T}(P_o, C)$ of the plant P_o and a controller C can be visualized by the block diagram depicted in Figure 3.1. The feedback connection $\mathcal{T}(P_o, C)$ of Figure 3.1 indicates the so-called "actual" or "achieved" feedback connection (Gevers 1993). The actual feedback connection $\mathcal{T}(P_o, C)$ indicates that the controller C is actually being implemented on the

(unknown) plant P_o in a feedback connection.
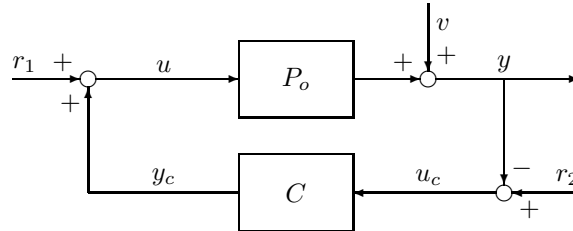
Fig. 3.1: Feedback connection $\mathcal{T}(P_o, C)$ of plant P_o and controller C.

As already indicated in Section 2.2.1, the signal v denotes an additive disturbance acting on the output of the plant. The role of the external reference signals in Figure 3.1 is twofold. From a control point of view, the reference signal r_1 can be used to denote a so-called feedforward signal (van de Vegte 1990). Due to the negative feedback connection where $u_c = r_2 - y$, the signal u_c can be interpreted as an error signal entering a servo compensator (Lauer *et al.* 1960). In this perspective, r_2 can be viewed as a set point signal present in a servo controlled system. In most control applications the presence of both signals is unavoidable in order to have the feedback connection $\mathcal{T}(P_o, C)$ work properly.

From an identification point of view, the external reference signals are introduced to provide sufficient excitation of the feedback connection $\mathcal{T}(P_o, C)$. Without the possibility to inject additional signals in the feedback connection, identification of the feedback controlled plant P_o may suffer from lack of excitation. Insufficient excitation may lead to problems associated with the identifiability of the plant P_o (Ljung 1987). For a more detailed discussion on the issues of excitation and identifiability of feedback systems, one is also referred to the book by Söderström and Stoica (1989). Additionally, the PhD-thesis by Aling (1989) provides additional insight in the conditions of excitation for identification of feedback controlled plants.

Basically, the conditions on excitation for the identification of a feedback controlled plant boil down to two basic requirements. Either the feedback controller C being used in $\mathcal{T}(P_o, C)$ should be sufficiently complex or an external reference signal should be available to provide sufficient excitation (Söderström *et al.* 1976, Gustavson *et al.* 1977, Anderson and Gevers 1982). The aim of this thesis is not directed towards the discussion of the problems that are associated with identifiability of feedback controlled systems. Therefore, it is assumed that at least one of the two reference signals depicted in Figure 3.1 is available and can be used to provide sufficient excitation for the identification of the feedback controlled plant P_o.

In this perspective, the signals r_1 and r_2 are used to represent the presence of (known) external signals that act on the feedback connection $\mathcal{T}(P_o, C)$. As such, the

signal r_2 is used to model an external signal on the input u_c of the controller C, while r_1 represents an external signal on the output y_c of the controller. It should be noted that r_1 and r_2 do not have to be known exactly.

Remark 3.2-1 *The signals $col(r_2, r_1)$ may be built up from a deliberately applied (and possibly known) pair of reference signals $col(\bar{r}_2, \bar{r}_1)$ and an unknown pair $col(\bar{v}_2, \bar{v}_1)$ such that*

$$\begin{aligned} r_2 &= \bar{r}_2 + \bar{v}_2 \\ r_1 &= \bar{r}_1 + \bar{v}_1 \end{aligned} \tag{3.1}$$

Such external reference signals may occur in case the implementation of a (digital) controller C on the plant P_o may introduce noise on the input and output of the controller due to quantization or finite word precision (Hanselmann 1987). To avoid the cumbersome notation involved in tracking all the possible noise contributions that may occur in the feedback connection $\mathcal{T}(P_o, C)$, the noise signals $\bar{v}_1$ and $\bar{v}_2$ are modelled in the single noise contribution v and $col(r_2, r_1)$ in (3.1) are assumed to be equal to the deliberately applied pair of reference signals $col(\bar{r}_1, \bar{r}_2)$.

Henceforth, the signals r_1 and r_2 are used to represent the deliberately applied pair of external signals on the output and input of the controller C. In light of the identification techniques being used, the noise v is assumed to be uncorrelated with the reference signals r_1 and r_2. Furthermore, only a finite number of samples N of the signals is assumed to be available. Hence the assumptions being made on the signals present in the feedback connection $\mathcal{T}(P_o, C)$ can be summarized as follows.

Assumption 3.2-2 *For identification purposes, N samples of the discrete-time domain signals $u(t)$ and $y(t)$ are assumed to be measurable. The external reference signals $r_1(t)$ and $r_2(t)$ in Figure 3.1 provide sufficient excitation of the feedback connection $\mathcal{T}(P_o, C)$. The additive noise $v(t)$ is assumed to be a stochastic process that is uncorrelated with the external reference signals and can be modelled as the output of a monic stable and stably invertible noise filter H_o having a white noise input $e(t)$.*

With Remark 3.2-1, Assumption 3.2-2 and the feedback connection $\mathcal{T}(P_o, C)$, the data coming from the plant P_o operating under closed-loop conditions can be described as follows.

$$\begin{bmatrix} y \\ u \end{bmatrix} = T(P_o, C) \begin{bmatrix} r_2 \\ r_1 \end{bmatrix} + \begin{bmatrix} I \\ -C \end{bmatrix} (I + P_o C)^{-1} H_o \, e \tag{3.2}$$

In (3.2), the mapping $T(P_o, C)$ is defined in (2.5), where C can be any (known) controller that is currently being implemented on the actual but unknown plant P_o. The assumption on the stochastic nature of the noise v in Assumption 3.2-2 is due to Ljung (1987) and can be used to cover a wide class of noise generating mechanisms that might be present on the data.

For notational convenience the shorthand notation

$$r := r_1 + Cr_2 \tag{3.3}$$

is introduced in order to be able to rewrite (3.2) into a convenient and reduced form. It should be stressed that (3.3) is used only as a shorthand notation and does not reflect a veritable signal that is actually being reconstructed on the basis of r_1, r_2 and the possible knowledge of the controller C.

Remark 3.2-3 *The signal r in (3.3) is used solely to denote mutual influence of the reference signals r_1 and/or r_2 that provide sufficient excitation according to Assumption 3.2-2 of the feedback connection $\mathcal{T}(P_o, C)$ in Figure 3.1. The signal r is not being constructed from (3.3) as r can be unbounded in case that the controller C is unstable and $r_2 \not\equiv 0$.*

On the basis of the shorthand notation (3.3), a reduced form of (3.2) can be formulated as follows.

$$\begin{aligned}
y &= P_o S_{in} r + S_{out} H_o e \\
u &= S_{in} r - C S_{out} H_o e
\end{aligned} \tag{3.4}$$

In (3.4) the variables S_{in} and S_{out} denote the input and output sensitivity function of the feedback connection (Maciejowski 1989) with

$$\begin{aligned}
S_{in} &:= (I + CP_o)^{-1} \\
S_{out} &:= (I + P_o C)^{-1}
\end{aligned} \tag{3.5}$$

In the case that both P_o and C are single-input–single-output (SISO) transfer functions, the multiplication of C and P_o is commutative and $S_{in} = S_{out}$. In the multivariable situation a distinction must be made between the input and output sensitivity given in (3.5). As the plant P_o is allowed to be a multivariable transfer function, a distinction is being made between the input and output sensitivity function. Finally the following algebraic relation between the reference signals r_1, r_2 and the input u and output y signals will be used frequently.

Corollary 3.2-4 *Consider the signal r defined in (3.3) and the input u and output y signal given in (3.4). Then the signal r satisfies*

$$r = r_1 + Cr_2 = u + Cy \tag{3.6}$$

and is uncorrelated with the additive noise $v = H_o e$.

Proof: From (3.4) with $S_{in} + CP_o S_{in} = (I + CP_o)S_{in} = I$, it can be verified that $u + Cy = r_1 + Cr_2$. As both r_1 and r_2 are assumed to be uncorrelated with v due to Assumption 3.2-2, r is uncorrelated with v. $\qquad\square$

With (3.6) it can be seen that r in (3.3) also satisfies $r = u + Cy$ where C indicates the controller being implemented on the plant P_o in the feedback connection $\mathcal{T}(P_o, C)$. Although both u and y in (3.4) depend on the additive noise v acting on the output of the plant P_o, the linear combination $u + Cy$ is free from the noise contribution v. This observation is also made in (Schrama 1992b). It should be mentioned that the above result is valid only for a negative feedback connection as defined in Definition 2.2-6. In case of a positive feedback with $u_c = r_2 + y$, (3.6) modifies into $r_1 + Cr_2 = u - Cy$ and the appropriate minus signs must be applied.

3.2.2   Main concepts in identification

Experiments obtained in the presence of a feedback controller constitute the basis for many closed-loop identification problems. Referring to the data generating system as mentioned in (3.2) or the short hand notation of the data coming from $\mathcal{T}(P_o, C)$ depicted in (3.4), the important and unknown quantities to be modelled are the plant P_o and the noise filter H_o. In case the controller C is unknown, it can also be added to the list of unknown quantities to be estimated. However, for analysis purposes, the attention is focused on the problem of estimating P_o and H_o first.

Estimation of parametric models

To insinuate the fact that a model is to be found by system identification techniques, a variable θ is used to denote the unknown quantity present in a linear model that needs to be estimated. For notational convenience, the plant P_o and the noise filter H_o are stacked into

$$T_o := [P_o \; H_o] \tag{3.7}$$

and the parameter dependency of a model of the unknown transfer function T_o is indicated by

$$T(\theta) = [P(\theta) \; H(\theta)] \tag{3.8}$$

where the variable θ denotes the (finite dimensional and real-valued) parameter of a model $T(\theta)$ to be estimated. The model $T(\theta)$ is built up from a noise model $H(\theta)$ and an input-output model $P(\theta)$ respectively to model H_o and P_o. In general, θ may be used to denote the numeric values present in a state space realization or a transfer function description of $T(\theta)$ stacked into a vector. As such, θ denotes a real valued parameter vector that needs to be estimated in order to complete a model.

On the basis of the notation given above, a parameter vector θ being estimated will be denoted by $\hat{\theta}$. As a result, the notation $\hat{T}$ will be used to denote $T(\hat{\theta})$ or equivalently, $\hat{P} = P(\hat{\theta})$ and $\hat{H} = H(\hat{\theta})$. As indicated in Assumption 3.2-2, a finite number N of (time domain) samples are available for estimation purposes. To indicate specifically that the parameter vector is estimated on the basis of a finite number N of samples, the notation $\hat{\theta}_N$ will be adopted. Similar to the argumentation mentioned

in Section 2.2.2, additional arguments such as the shift operator q or the complex valued argument z for a discrete-time transfer function are being omitted to simplify the notations.

As a free parameter vector θ is used to represent and estimate models $T(\theta)$, this approach is often paraphrased by parametric identification or identification of parametric models (Ljung 1987). This notion is used in order to distinguish between so-called non-parametric identification methods such as spectral analysis or Fourier analysis (Priestley 1981). In these methods, a model for T_o is represented by a finite number of frequency points instead of a real valued parameter vector θ. In both parametric and non-parametric identification a model is estimated and represented by a finite number of real or complex valued parameters. However, in parametric identification typically a model is found by estimating a limited number of parameters.

Remark 3.2-5 *Although the difference between parametric and non-parametric becomes questionable once the dimension of the parameter vector θ increases, the distinction between parametric and non-parametric will be used here for reasons of clarity.*

The way in which the parameter vector θ enters into the model $T(\theta)$ is determined by the parametrization. Formally, a parametrization is a map Π that maps the parameter vector θ onto a specific model $T(\theta)$ as follows

$$\Pi : \theta \mapsto T(\theta), \ \theta \in \Theta \subset \mathbb{R}^d \tag{3.9}$$

In the above formulation, $\mathbb{R}^d$ is used to denote the dimension d of the real valued parameter vector θ. The additional condition that θ should lie within a prespecified parameter space Θ being a subset of $\mathbb{R}^d$ allows the possibility to put additional constraints on the parameter vector θ being estimated. Such constraint may include stability considerations of the model being estimated. In case of the Prediction Error (PE) framework, an (open) parameter space Θ is used to restrict the predictor to be stable, see e.g. Ljung (1987), chapter 4.

It should be noted that the map Π in (3.9) does not specify the way in which model $T(\theta)$ is being constructed from a parameter vector θ. As such, the map Π may include various ways to parametrize a multivariable input-output model $P(\theta)$ and noise model $H(\theta)$. Parametrizations of state space matrices as introduced in McKelvey (1996) or a fully parametrized state space matrices employed in a subspace method (Viberg 1994) can also be represented by a map Π similar to (3.9). However, a parametrization that uses a parameter vector θ that has the smallest possible dimension d to represent a model $T(\theta)$ is said to be a minimal parametrization. An important property that a minimal parametrization may exhibit, is reflected by the notion of identifiability. A formal definition of identifiability is adopted from Ljung (1987) and is listed below.

Definition 3.2-6 *Consider the map* Π *given in (3.9). The parametrization* Π *is called locally identifiable at* $\theta^\star$ *if*

$$\Pi(\theta) = \Pi(\theta^\star) \Rightarrow \theta = \theta^\star \ \forall \theta \in \Theta$$

Subsequently, the parametrization Π *is globally identifiable if it is locally identifiable at almost all* $\theta^\star \in \Theta$.

It can be observed from Definition 3.2-6 that (global) identifiability of a parametrization Π is concerned with the bijectivity of the map Π. Therefore, identifiability of the parametrization can be considered as a property of finding a unique value of a parameter vector θ when applying an estimation procedure to find an estimate $\hat{\theta}$.

A classical result for multivariable models is the impossibility to construct one bijective continuous map Π that is able to cover all multivariable models up to a pre-specified McMillan degree (Luenberger 1967). To circumvent this problem, various alternative parametrizations have been proposed that satisfy the identifiability property mentioned in Definition 3.2-6. Well known examples are canonical (overlapping) parametrizations of state space realizations (Glover and Willems 1974, van Overbeek and Ljung 1982, Corrêa and Glover 1984, Corrêa and Glover 1986, Janssen 1988) or the closely related matrix polynomial parametrizations, see e.g. (Guidorzi 1975, Gevers and Wertz 1984, Van den Hof 1989). Alternative canonical parametrization based on balanced state space realizations (Ober 1987, Ober 1991) can also be used to define an identifiable parametrization.

Parametrization of transfer function models

In the parametrization Π given in (3.9), both the transfer function of the input-output model $P(\theta)$ and the noise model $H(\theta)$ have been parametrized by the same parameter vector θ. As a consequence, estimation of θ will influence both $P(\theta)$ and $H(\theta)$. A frequently used model parametrization that exhibits such mutual influence of both $P(\theta)$ and $H(\theta)$ is an Auto Regressive, Moving Average with eXogenous input (ARMAX) model (Åström and Bohlin 1965).

Adopting to the notation of parametrized transfer functions as used in the framework of the prediction error identification (Ljung 1987), the following prediction error model is considered

$$y(t) = P(q, \theta)u(t) + H(q, \theta)\varepsilon(t, \theta) \tag{3.10}$$

where $\varepsilon(t, \theta)$ denotes the (one step ahead) prediction error and $H(q, \theta)$ a monic stable and stably invertible noise filter. Following (3.10), the parametrization of a discrete-time linear time invariant ARMAX-model structure can be characterized as follows.

$$A(q, \theta_a)y(t) = B(q, \theta_b)u(t) + C(q, \theta_c)\varepsilon(t, \theta) \tag{3.11}$$

In (3.11), the parameter vector $\theta = [\theta_a^T \; \theta_b^T \; \theta_c^T]^T$ while $\varepsilon(t, \theta)$ is the prediction error. As mentioned Section 3.2.1, the signals $y(t)$ and $u(t)$ are considered to be obtained from the plant. The parameter vector θ is used as an argument of the prediction error $\varepsilon(t, \theta)$ and denotes the unknown coefficients to be estimated that appear linearly in the (matrix) polynomials $A(q, \theta_a)$, $B(q, \theta_b)$ and $C(q, \theta_c)$ (Ljung 1987).

Closely related to the ARMAX-model of (3.11) is an ARX-model or equation error model structure

$$A(q, \theta_a)y(t) = B(q, \theta_b)u(t) + \varepsilon(t, \theta) \tag{3.12}$$

for which the moving average part $C(q, \theta_c)$ in (3.11) is set to identity and $\theta = [\theta_a^T \; \theta_b^T]^T$ consequently. As the unknown coefficients appear linearly in the matrix polynomials, (3.12) yields a prediction error $\varepsilon(t, \theta)$ that is linear in the parameters θ_a and θ_b to be estimated. Linear appearance of the parameter vector θ in the prediction error $\varepsilon(t, \theta)$ greatly facilitates the estimation of θ and for that reason ARX-models are popular in the field of system identification (Rao 1973, Ljung 1987).

As can be seen from (3.11) and (3.12), $P(q, \theta) = A(q, \theta_a)^{-1}B(q, \theta_b)$ while $H(q, \theta) = A(q, \theta_a)^{-1}C(q, \theta_c)$ and $H(q, \theta) = A(q, \theta_a)^{-1}$ respectively for the ARMAX-model (3.11) and the ARX-model (3.12). Clearly, the polynomial $A(q, \theta_a)$ enforces a mutual influence between the input-output model and the noise model that may be undesirable. To decouple this mutual influence an independent parametrization of the input-output model and the noise model can be used.

$$\Pi : \begin{bmatrix} \vartheta \\ \eta \end{bmatrix} \mapsto [P(\vartheta) \; H(\eta)], \; \vartheta \in \Theta_\vartheta, \; \eta \in \Theta_\eta, \; \Theta_\vartheta \times \Theta_\eta \subset \mathbb{R}^d \tag{3.13}$$

As indicated in (3.13), the parameter vector θ is split up in $\theta = [\vartheta^T \; \eta^T]^T$ to accomplish an independent parametrization. Such an independent parametrization of the transfer functions $P(q, \vartheta)$ and $H(q, \eta)$ is fulfilled in a Box-Jenkins (BJ) model structure (Box and Jenkins 1970). Again adopting the notation of Ljung (1987), the parametrization of a Box-Jenkins model structure can be characterized by

$$y(t) = F(q, \vartheta_f)^{-1}B(q, \vartheta_b)u(t) + D(q, \eta_d)^{-1}C(q, \eta_c)\varepsilon(t, \theta) \tag{3.14}$$

where $P(q, \vartheta) = F(q, \vartheta_f)^{-1}B(q, \vartheta_b)$, $H(q, \eta) = D(q, \eta_d)^{-1}C(q, \eta_c)$ and the parameter vector $\theta = [\vartheta_f^T \; \vartheta_b^T \; \eta_d^T \; \eta_c^T]^T$.

A special case of an independent parametrization Π is obtained by choosing a fixed value $\eta = \eta^\star$ for the noise filter $H(\eta)$. As the noise filter is fixed to $H(\eta^\star)$, $\theta = \vartheta$ is the only free parameter in the parametrization of (3.13). Such a parametrization is useful if one is interested to find a model for P_o, whereas the estimation of a noise model is considered to be of less importance. A commonly used parametrization that uses such a fixed noise filter is an Output Error (OE) model structure (Ljung 1987)

$$y(t) = F(q, \theta_f)^{-1}B(q, \theta_b)u(t) + \varepsilon(t, \theta) \tag{3.15}$$

having a noise model $H(\eta)$ that has been set to identity. In this way, the estimation of a noise model is completely decoupled from the estimation of an input-output model.

Additionally, setting $F(q, \theta) = 1$ in (3.15) can be used to facilitate the estimation of the parameter vector θ. Similar to (3.12), a Finite Impulse Response (FIR) model with

$$y(t) = B(q, \theta_b)u(t) + \varepsilon(t, \theta), \ B(q, \theta_b) := \sum_{k=0}^{n_b} B_k q^{-k} \tag{3.16}$$

has the property of having both a prediction error $\varepsilon(t, \theta)$ expressed linearly in the parameter $\theta = \theta_b$ and an independent estimation of the input-output model. Although the estimation of the parameter vector θ benefits from the linear regression structure given in (3.16), typically the number of coefficients n_b to be estimated in (3.16) expands dramatically in case a lightly damped plant must be modelled (Heuberger *et al.* 1995). This is due to the fact that the poles of the discrete-time input-output model $P(q, \theta) = B(q, \theta)$ are all fixed to zero.

The number of parameters to be estimated in the linear regression structure of (3.16) can be reduced by using a linearly parametrized model structure that is able to incorporate poles that are unequal to zero. Such a linear regression structure has been proposed in Heuberger (1991) and is based on an expansion using orthonormal basis functions $V_k(q)$ that generalize the orthonormal function q^{-k} as used in (3.16). Following Heuberger (1991), a linear regression model can be formulated as

$$y(t) = \sum_{k=0}^{n_l} L_k V_k(q)u(t) + \varepsilon(t, \theta) \tag{3.17}$$

where the orthonormal functions $V_k(q)$ may contain information on the location of the (stable) poles of the plant P_o, see also Heuberger *et al.* (1995) or Ninness and Gómez (1995). The parameter vector θ now represents the parameters L_k to be estimated in the model structure labelled as ORTFIR (Heuberger *et al.* 1995). Due to the additional knowledge incorporated in $V_k(q)$, the number of coefficients n_l to be estimated can be reduced significantly compared to the number of coefficient n_b in the FIR-model structure given in (3.16) (Heuberger *et al.* 1995).

Consistent and approximate estimation

Evaluating the map $\Pi : \theta \mapsto T(\theta)$ given in (3.9) over the allowable parameter space Θ yields all possible models that are captured by the parametrization Π. In this respect, a model set $\mathcal{M}^1$ can be characterized as follows.

$$\mathcal{M} := \{T(\theta) \mid \theta \in \Theta\} \tag{3.18}$$

A similar model set $\mathcal{M}$ can be found for the independent parametrization mentioned in (3.13). In that case, the model set $\mathcal{M}$ is built up from models $[P(\vartheta) \ H(\eta)]$.

---

[1]Not to be confused with the notion of a set of models $\mathcal{P}$ as given in Definition 2.2-16.

Clearly, the model set $\mathcal{M}$ of (3.18) contains both input-output models $P(\theta)$ and noise models $H(\theta)$. For notational convenience, a model set that contains the input-output models $P(\theta)$ only is denoted by $\mathcal{G}$. Formally, this model set $\mathcal{G}$ is defined by

$$\mathcal{G} := \{P(\theta) \mid \theta \in \Theta\} \tag{3.19}$$

Perhaps superfluously, it is mentioned that the following two relations hold

$$\begin{aligned} T_o \in \mathcal{M} &\Rightarrow P_o \in \mathcal{G} \\ P_o \notin \mathcal{G} &\Rightarrow T_o \notin \mathcal{M} \end{aligned} \tag{3.20}$$

where $T_o \in \mathcal{M}$ is equivalent to the existence of a parameter $\theta_o \in \Theta$ such that $T(\theta_o) = T_o$. Similarly, $P_o \in \mathcal{G}$ is equivalent to the existence of a parameter $\theta_o \in \Theta$ such that $P(\theta_o) = P_o$.

The relations mentioned in (3.20) play an important role in characterizing the problem of estimating models on the basis of (closed-loop) observations. With the use of the previously defined model sets $\mathcal{M}$ and $\mathcal{G}$, a distinction can be made between consistent and approximate identification. This distinction is based on the fact whether or not the models to be identified are an element of either the model set $\mathcal{M}$ or $\mathcal{G}$.

Within the PE framework, estimation of the parameter θ is done by minimizing the variance of the prediction error $\varepsilon(t, \theta)$. Given N time domain samples, an estimate of θ is denoted by $\hat{\theta}_N$ and is obtained by performing the following minimization

$$\hat{\theta}_N = \min_{\theta \in \Theta} \frac{1}{N} \sum_{t=1}^{N} \mathrm{tr}\{\varepsilon(t, \theta) W \varepsilon(t, \theta)^T\} \tag{3.21}$$

where $\mathrm{tr}\{\cdot\}$ denotes the usual trace operator, W is an optional scalar weighting matrix and $\varepsilon(t, \theta)$ is the prediction error as found in (3.10). In case the prediction error $\varepsilon(t, \theta)$ is a scalar signal, the weighting matrix W can be omitted. In the sequel, the weighting matrix W will be omitted without loss of generality. The prediction error $\varepsilon(t, \theta)$ in (3.21) can also be replaced by a filtered prediction error $\varepsilon_f(t, \theta)$ given by

$$\varepsilon_f(t, \theta) = L(q)\varepsilon(t, \theta)$$

where the filter $L(q)$ can be used to emphasize, reduce or modify the frequency contents of the prediction error signal, if needed. With notation of the estimated parameter $\hat{\theta}_N$ given in (3.21), the following definition can be given.

Definition 3.2-7 *Consider T_o given in (3.7) and let the model sets $\mathcal{M}$ and $\mathcal{G}$ respectively be defined in (3.18) and (3.19).*

- *The consistent identification problem is defined as the estimation of a parameter $\hat{\theta}_N \in \Theta$ for the situation $T_o \in \mathcal{M}$ in which there exists a $\theta_o \in \Theta$ such that $T(\theta_o) = T_o$. The estimation of $\hat{\theta}_N$ is consistent if*

$$\lim_{N \to \infty} T(\hat{\theta}_N) = T(\theta_o) \ w.p. \ 1. \tag{3.22}$$

- *The partial consistent identification problem is defined as the estimation of a parameter $\hat{\theta}_N \in \Theta$ for the situation $P_o \in \mathcal{G}$ in which there exists a $\theta_o \in \Theta$ such that $P(\theta_o) = P_o$ while $T(\theta_o) \neq T_o$. The estimation of $\hat{\theta}_N$ is partial consistent if*

$$\lim_{N \to \infty} P(\hat{\theta}_N) = P(\theta_o) \ w.p. \ 1. \tag{3.23}$$

- *The approximate identification problem is defined as the estimation of a parameter $\hat{\theta}_N \in \Theta$ for the situation $P_o \notin \mathcal{G}$. Approximate identification specifically considers the problem of estimating approximate models with a limited or pre-specified McMillan degree.*

In case a model set $\mathcal{M}$ is chosen such that both the plant P_o and the noise model H_o can be modelled exactly, a consistent estimate of the parameter θ_o is preferable. The problem of consistent estimation of both plant and noise model has been studied extensively in the literature.

Results on either open- or closed-loop data (direct identification) using a PE-framework can be found in Ljung (1987) and Söderström and Stoica (1989). Consistent estimation of both the plant P_o and the noise model H_o (and possibly the controller used for the closed-loop experiments) based on a so-called joint-input joint-output method can be found in e.g. Caines and Chan (1975), Caines and Chan (1976) or Gevers and Anderson (1981). Surveys of several identification techniques for the consistent identification problem can also be found in Gustavson *et al.* (1977) and Gustavson *et al.* (1981).

Clearly, requiring $T_o \in \mathcal{M}$ implies that both the plant P_o and the noise filter H_o have to be an element of the model set $\mathcal{M}$. In most practical situations this requirement enforces the model set $\mathcal{M}$ to be highly complex in order to be able to guarantee $T_o \in \mathcal{M}$. This is due to the fact that either the plant P_o or the noise filter H_o can have a high order or McMillan degree. In most control applications modelling of the noise filter H_o is inferior to the estimation of a model for the plant P_o. Clearly, it is the plant P_o who is operating in a feedback connection. In the case that the stability of a feedback connection $\mathcal{T}(P_o, C)$ is analyzed on the basis of a model, an inaccurate modelling of a stable and stably invertible noise model will not affect the outcome. On the other hand, an inaccurate model of the plant P_o does give the opportunity to draw an incorrect conclusion (Schrama 1992*b*).

In this perspective, the partial consistent estimation given in Definition 3.2-7 is a reasonable alternative. Approaches to the (partial) consistent identification prob-

lem using instrumental variable methods, that were originally developed for open-loop based identification, have been analyzed in Söderström and Stoica (1981) and Söderström and Stoica (1983) in case the reference signals are used as instruments. Indirect methods that first estimate the complete closed-loop system in order to re-compute a consistent estimate of the plant P_o can be found in Ljung *et al.* (1974) or Gevers (1978). More recent approaches using a subspace based estimation technique can be found in Verhaegen (1993). Another indirect method that is able to address the partial consistent identification problem is the two-stage method presented in Van den Hof and Schrama (1993). An alternative indirect method, that uses a correction of the parameters of the closed-loop system being estimated via an ARX-model structure, is reported in Zheng and Feng (1995).

Approximate identification

Despite of the results available on the (partial) consistent identification problems mentioned in Definition 3.2-7, the possibility to handle the approximate identification problem would be more valuable. In that case, a (deliberate) approximate modelling of a complex and unknown plant P_o can be used to find low complexity models. This opens the possibility to estimate models with a predescribed and fixed complexity, so as to keep track of the order of the model being estimated. Requiring $P_o \in \mathcal{G}$, might imply the estimation of unnecessarily high order models contained in the set $\mathcal{G}$ in order to be able to capture the dynamics of the plant completely.

However, the approximate identification will yield a model $P(\hat{\theta}_N) \neq P_o$. Even in the case of having an infinite number of data points N, as mentioned in (3.23), the model being estimated will only yield an approximation of the actual plant P_o. For the limiting case of $N \rightarrow \infty$, such a model is characterized by $P(\theta^\star)$ with

$$\lim_{N \to \infty} P(\hat{\theta}_N) = P(\theta^\star), \ w.p. \ 1 \tag{3.24}$$

provided that this limit exists. It has been shown in (Ljung 1987) that for the general framework of PE-methods, the limit (3.24) is well defined and $\theta^\star$ denotes a parameter vector that satisfies $\theta^\star \in \Theta$.

Due to the approximate identification, $P(\hat{\theta}_N) \neq P_o$ in (3.24) and it is desirable to have an expression of the misfit between the plant P_o and the model $P(\hat{\theta}_N)$. Such a tunable (bias) expression can be exploited to shape the approximation being made and may serve as a design tool while estimating an approximate model of the plant P_o.

Within the PE-framework, the availability of such a tunable and explicit expression of the misfit between the plant P_o and the limiting model $P(\theta^\star)$ can be characterized by writing down an equivalent frequency domain expression of (3.21). In conformance

with (Ljung 1987), the variance of $\varepsilon(t, \theta)$ as used in (3.21) for $N \to \infty$, is denoted by

$$\overline{\mathrm{E}}\{\varepsilon(t,\theta)\varepsilon(t,\theta)\} := \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \mathrm{E}\{\varepsilon(t,\theta)\varepsilon(t,\theta)^T\} \tag{3.25}$$

where $\mathrm{E}\{\cdot\}$ denotes the usual expectation operator. By Parseval's relation (Ljung 1987, Jackson 1991) the following frequency domain expression holds

$$\overline{\mathrm{E}}\{\varepsilon(t,\theta)\varepsilon(t,\theta)^T\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_\varepsilon(\omega,\theta) \, d\omega$$

where Φ_ε denotes the (auto) spectrum of the prediction error ε. As a consequence, for the limiting case $N \to \infty$ the estimate $\hat{\theta}_N$ is given by the following frequency domain expression.

$$\lim_{N\to\infty} \hat{\theta}_N := \theta^\star = \min_{\theta\in\Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{tr}\{\Phi_\varepsilon(\omega,\theta)\} \, d\omega \tag{3.26}$$

Equation (3.26) is a fundamental expression in deriving frequency domain expressions for the misfit between the plant P_o and the model being estimated, in case of an approximate identification.

3.2.3 Approximate identification without feedback

In case of open-loop identification, (3.6) reduces to $u = r_1$, while the data coming from the plant P_o given in (3.2) can be reduced to

$$y(t) = P_o(q)u(t) + H_o(q)e(t) \tag{3.27}$$

describing the open-loop operational conditions of the plant P_o. In (3.27), u is uncorrelated with the noise $v(t) = H_o(q)e(t)$ acting on the output of the plant.

Considering the variance of an unfiltered[2] prediction error $\varepsilon(t,\theta)$ mentioned in (3.25), in Ljung (1987) a frequency domain expression for the misfit between the plant P_o and the limiting model $P(\theta^\star)$ of (3.24) is formulated. For the case in which the input-output model $P(\theta)$ and the noise model $H(\theta)$ may have parameters in common, an expression for the misfit between the plant and the model being estimated can be obtained by observing that

$$\begin{aligned}
\varepsilon(t,\theta) &= H(q,\theta)^{-1}((P_o(q) - P(q,\theta))u(t) + H_o(q)e(t)) \\
&= H(q,\theta)^{-1}((P_o(q) - P(q,\theta))u(t) + (H_o(q) - H(q,\theta))e(t)) + e(t).
\end{aligned} \tag{3.28}$$

As both $H_o(q)$ and $H(q,\theta)$ are monic noise filters and $e(t)$ is a white noise, it can be verified that $\overline{\mathrm{E}}\{e(t)\tilde{e}(t)\} = 0$, where $\tilde{e}(t) := (H_o(q) - H(q,\theta))e(t)$. Furthermore,

---

[2]Similar expression can also be obtained in case a filtered version of the prediction error is being minimized.

$\overline{\mathrm{E}}\{e(t)u(t-\tau)^T\} = 0 \ \forall \tau$ as $e(t)$ and $u(t)$ are assumed to be uncorrelated. Restricting to the single-input single-output case to simplify notations, the following (frequency domain) expression for the auto spectrum $\Phi_\varepsilon(\omega,\theta)$ of the prediction error $\varepsilon(t,\theta)$ can be obtained

$$
\begin{aligned}
\Phi_\varepsilon(\omega,\theta) &= \frac{|P_o(\mathrm{e}^{i\omega}) - P(\mathrm{e}^{i\omega},\theta)|^2 \Phi_u(\omega) + |H_o(\mathrm{e}^{i\omega})|^2 \lambda}{|H(\mathrm{e}^{i\omega},\theta)|^2} \\
&= \frac{|P_o(\mathrm{e}^{i\omega}) - P(\mathrm{e}^{i\omega},\theta)|^2 \Phi_u(\omega) + |H_o(\mathrm{e}^{i\omega}) - H(\mathrm{e}^{i\omega},\theta))|^2 \lambda}{|H(\mathrm{e}^{i\omega},\theta)|^2} + \lambda
\end{aligned}
\tag{3.29}
$$

by taking expectation and applying Fourier transform subsequently to (3.28). In (3.29), $\Phi_u(\omega)$ denotes the auto spectrum of the input signal u, while λ denotes the level of the (constant) auto spectrum of the white noise $e(t)$, making $\Phi_\varepsilon(\omega) \geq \lambda$.

For the limiting case $N \to \infty$, the estimation of a parameter $\theta^\star$ found by minimizing the variance of the prediction error $\varepsilon(t,\theta)$ as a function of θ, can be written as

$$
\min_{\theta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|P_o(\mathrm{e}^{i\omega}) - P(\mathrm{e}^{i\omega},\theta)|^2 \Phi_u(\omega) + |H_o(\mathrm{e}^{i\omega}) - H(\mathrm{e}^{i\omega},\theta)|^2 \lambda}{|H(\mathrm{e}^{i\omega},\theta)|^2} \, d\omega
\tag{3.30}
$$

by using (3.26). The above expression is slightly different from the frequency domain expression mentioned in Ljung (1987, pp. 224), as the level λ in (3.29) of the (constant) auto spectrum Φ_e of the white noise e has been eliminated, so that the integral in (3.30) is larger than or equal to zero.

It can be observed from (3.30) that a trade off is made in the approximate identification on the basis of open loop data using an input-output model and noise model with common parameters (ARX- or ARMAX-model structure). This trade off consists of fitting an input-output model $P(\theta)$ to P_o and fitting a noise model $H(\theta)$ to H_o. In this trade off, the difference $|P_o - P(\theta)|^2$ is weighted by the (frequency dependent) input spectrum Φ_u and the difference $|H_o - H(\theta)|^2$ is weighted by a constant λ, while both are weighted by $1/|H(\theta)|^2$.

Although (3.30) yields an expression for the resulting estimate obtained by approximate identification and illuminates the trade off being made, it is not an explicit expression for the misfit between plant P_o and a model $P(\theta)$ being estimated. This is due to the fact that the noise filter $H(\theta^\star)$ to be estimated influences the approximate estimate $P(\theta^\star)$ being obtained, while $H(\theta^\star)$ is not known beforehand.

As indicated in Section 3.2.2, the mutual influence of the estimation of a noise filter on the estimation of the input-output model can be decoupled by employing an independently parametrized model structure (3.13) such as a BJ-model structure (3.14). Specifically in the case of choosing a fixed noise filter $H(\eta^\star)$, not necessarily equal to H_o, the difference between $|H_o - H(\eta^\star)|$ will not contribute to the minimization of (3.30). Hence, an explicit and tunable expression for the bias or misfit between the

plant P_o and the model $P(\theta^\star)$ can be formulated. In case of the OE-model structure (3.15) for which $H(\eta^\star)$ is set to identity, (3.30) modifies into

$$\min_{\theta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_o(\mathrm{e}^{\,i\omega}) - P(\mathrm{e}^{\,i\omega}, \theta)|^2 \Phi_u(\omega)\, d\omega \qquad (3.31)$$

which is 'a clear-cut best mean-square approximation' (Ljung 1987) of P_o obtained via a minimization of the additive difference between the plant P_o and the model $P(\theta)$, weighted by the frequency dependent input spectrum Φ_u.

As mentioned above, the integral in (3.31) is larger than or equal to zero and it can be observed that (3.31) is zero or minimized when $P(\theta) = P_o$ in case $P_o \in \mathcal{G}$. This addresses the partial consistent identification problem of Definition 3.2-7 and is a known result (Ljung 1987). In case $P_o \notin \mathcal{G}$, an explicit and tunable expression for the bias of a model found by approximate identification on the basis of open loop experiments is obtained.

3.2.4   Ignoring the feedback: direct identification

Approximate identification on the basis of data obtained under feedback controlled conditions has attained attention only recently in the literature. Results on the estimation of models on the basis of closed-loop data, that address the approximate identification problem according to Definition 3.2-7, are less plentiful. This is due to the fact that finding an explicit and tunable (bias) expression for the misfit between the plant P_o and a model $P(\theta)$ being estimated is more involved in case closed-loop data is used for the approximate identification.

To illustrate the effects associated to approximate identification on the basis of closed-loop data, consider the data coming from the plant P_o given in (3.2). The prediction error (3.10) with

$$\varepsilon(t, \theta) = H(q, \theta)^{-1}(y(t) - P(q, \theta)u(t))$$

can be proposed to (directly) use the input u and output y signal of the plant P_o in an open-loop way for identification purposes. In this way, the feedback is ignored and such an identification is labelled as a direct closed-loop identification.

As the signals u and y are obtained under feedback, substitution of (3.4) yields the following prediction error

$$\varepsilon(\theta) = H(\theta)^{-1}((P_o - P(\theta))S_{in}r + (I + P(\theta)C)S_{out}H_o e) \qquad (3.32)$$

In (3.32) the arguments t and q are omitted for clarity. Furthermore, r denotes the signal given in the shorthand notation (3.3), whereas S_{in} and S_{out} are given in (3.5). As mentioned in Corollary 3.2-4, r and e are uncorrelated. Again restricting to the

single-input single-output case to simplify notations, the following (frequency domain) expression for the auto spectrum $\Phi_\varepsilon(\omega, \theta)$ can be obtained

$$\Phi_\varepsilon(\theta) = \frac{|P_o - P(\theta)|^2|S_{in}|^2\Phi_r + |1 + P(\theta)C|^2|S_{out}|^2|H_o|^2\lambda}{|H(\theta)|^2} \qquad (3.33)$$

where the arguments $e^{i\omega}$ and ω are omitted to further simplify notations and Φ_r denotes the (frequency dependent) auto spectrum of the reference signal r given in (3.3).

Again it can be verified from (3.33) that a trade off is made between fitting a input-output model to P_o and a noise model to H_o by minimizing

$$\min_{\theta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_\varepsilon(\omega, \theta)\, d\omega \qquad (3.34)$$

where $\Phi_\varepsilon(\omega, \theta)$ is given in (3.33). However, compared to (3.29), the following two observations can be made.

- The input S_{in} and output sensitivity function S_{out} act as additional frequency dependent weightings. This is due to the fact that the signals used for identification are obtained under feedback or closed-loop conditions. As a result, the reference signal r, used to excite the input u of the plant, will be filtered by the input sensitivity S_{in}. In a similar way, the additive noise v will appear filtered by the output sensitivity S_{out} on the output of the plant.

- The noise filter is weighted by an additional term $(1 + P(\theta)C)$ that depends on the model $P(\theta)$ to be estimated. This is due to the fact that the input signal u is correlated with the noise v due to feedback.

The latter causes (3.33) to be untunable expression for the misfit between the plant P_o and a model $P(\theta)$ being estimated. Even in the case when a fixed noise filter $H(\eta^\star) \neq H_o$ is chosen, the term

$$\frac{|1 + P(\theta)C|^2|S_{out}|^2|H_o|^2\lambda}{|H(\eta^\star)|^2} \qquad (3.35)$$

still contributes to the minimization (3.34) as $P(\theta)$ appears in (3.35), whereas this is not the case in an open-loop identification. Hence, the estimation of an input-output model that uses the input u and output y of the plant directly, is implicitly influenced by the noise contribution present on the data. The misfit between a model and the plant cannot be tuned explicitly, making the closed-loop more complicated than an open-loop approximate identification.

To illustrate the problem associated to a closed-loop identification, consider an OE-model structure with $H(\eta^\star) = 1$. In that case, $\Phi_\varepsilon(\omega, \theta)$ given in (3.33), reduces

to

$$|P_o(\mathrm{e}^{\,i\omega}) - P(\mathrm{e}^{\,i\omega}, \theta)|^2 |S_{in}(\mathrm{e}^{\,i\omega})|^2 \Phi_r(\omega) +$$

$$+ |1 + P(\mathrm{e}^{\,i\omega}, \theta) C(\mathrm{e}^{\,i\omega})|^2 |S_{out}(\mathrm{e}^{\,i\omega})|^2 \|H_o(\mathrm{e}^{\,i\omega})\|^2 \lambda \tag{3.36}$$

As can be seen from (3.36), no explicit and tunable expression for the misfit between P_o and $P(\theta)$ is obtained, as in the open-loop case. The noise (filter H_o) still influences the estimate of the input-output model $P(\theta)$ in the second term of (3.36).

Furthermore, the minimization (3.34) of (3.36) with $P_o \in \mathcal{G}$ does not imply $P(\theta^\star) = P_o$, as minimization of the second term in (3.36) does not imply $P(\theta^\star) = P_o$. This result on partial consistency was found in the open-loop case, but does not hold in case closed-loop data is used. In the extreme situation of $\Phi_r = 0 \ \forall \omega$ (no reference signal) a model $P(\theta^\star)$ is found that minimizes

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |1 + P(\mathrm{e}^{\,i\omega}, \theta) C(\mathrm{e}^{\,i\omega})|^2 |S_{out}(\mathrm{e}^{\,i\omega})|^2 |H_o(\mathrm{e}^{\,i\omega})|^2 \lambda \, d\omega \tag{3.37}$$

In the case that $-C^{-1} \in \mathcal{G}$, the minimum of (3.37) is attained at $P(\theta^\star) = -C^{-1}$, even if $P_o \in \mathcal{G}$. Clearly, this is an undesired biased estimate of the plant P_o and is caused by the fact that without a reference signal, no distinction can be made between P_o and C present in the feedback connection $\mathcal{T}(P_o, C)$. As indicated in Figure 3.2, a natural consequence is the estimation of a model that approximates the inverse of the controller, no matter what the plant P_o might be.
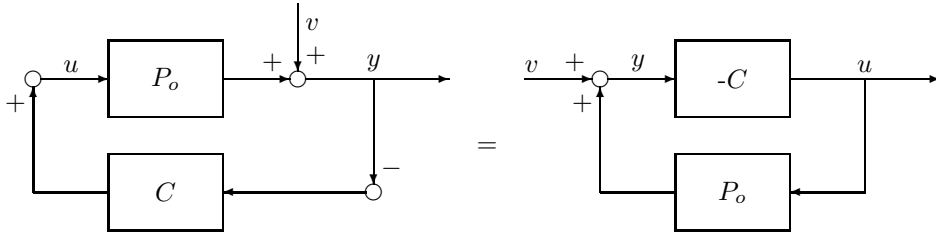
Fig. 3.2: Closed-loop identification without reference signal.

Assumption 3.2-2 mentions the fact that the external reference signals are introduced to provide sufficiently excitation of the feedback connection $\mathcal{T}(P_o, C)$. Insufficient excitation may lead to problems associated with the consistent estimation of the plant P_o, as indicated above. Hence, in many practical situations both reference and noise are present on the feedback connection $\mathcal{T}(P_o, C)$ while gathering data for identification purposes.

Still, even in the case of choosing a fixed noise filter $H(q, \eta^\star)$, the term (3.35) will remain to influence the model $P(\theta)$ being estimated. This influence is not known

beforehand and depends on the noise present on the closed-loop data. As a result, a model $P(\theta)$ being estimated will be biased, even if no approximate identification is performed, where the bias or misfit between the model and the plant P_o depends on the noise present on the closed-loop data.

Example 3.2-8 *The noise may highly influence an identification that is based directly on the input u and output y signal of a plant P_o obtained under feedback. To illustrate this effect, consider the plant*

$$P_o(q) = \frac{q - 0.6}{(q - 0.95)(q - 0.95)} = \frac{q^{-1} - 0.6q^{-2}}{1 - 1.9q^{-1} + 0.9025q^{-2}}$$

that is controlled by unity feedback ($C = 1$) to form a feedback connection $\mathcal{T}(P_o, C)$ as depicted in Figure 3.1. The reference signal $r = r_1 + Cr_2$ is chosen as a zero mean normally distributed white noise signal with variance 1. The noise v acting on $\mathcal{T}(P_o, C)$ is given by

$$v(t) = \lambda H_o(q)e(t), \ H_o(q) = \frac{q - 0.1}{q - 0.9} \tag{3.38}$$

where $e(t)$ is a zero mean normally distributed white noise signal with variance 1 and uncorrelated with r. To illustrate the influence of the noise v acting on the system, the noise intensity variable λ in (3.38) is used to modify the variance of the noise.

To model the plant P_o, an OE-model (3.15) is used with

$$F(q, \theta) = 1 + f_1 q^{-1} + f_2 q^{-2}, \ B(q, \theta) = b_0 + b_1 q^{-1} + b_2 q^{-2}$$

hence, $P_o \in \mathcal{G}$ but $T_o \notin \mathcal{M}$. For the identification of a model, 1000 data points of the closed-loop input u and output y signals are used. To illustrate the effect of the noise, 3 different experiments are simulated where the noise intensity λ in (3.38) is set to 5, 1 and 0.2 respectively, while the variance of the reference signal r is kept the same.

On the basis of these 3 experiments, 3 OE-models are estimated. A Monte Carlo simulation is performed where the above mentioned experiments and subsequent OE-model identification are invoked 10 times. An amplitude Bode plot of the resulting models $P(\hat{\theta})$ being estimated is plotted in Figure 3.3.

It can be observed from Figure 3.3 that the intensity of the noise highly influences the bias of the input-output model, whereas no noise model is being estimated. Furthermore it can be seen that the models are biased from the plant P_o, even though an OE-model structure is chosen that is able to capture the dynamics of the plant completely. In case the noise level is large ($\lambd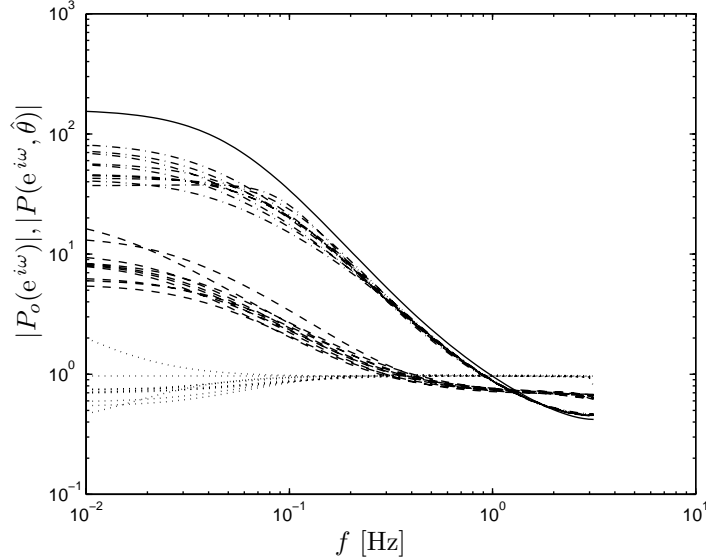a = 5$) the model being estimated approaches $-C^{-1} = -1$.   ◇

Fig. 3.3: Amplitude Bode plots of plant $P_o(\mathrm{e}^{i\omega})$ (—) and 10 realizations of OE-models $P(\mathrm{e}^{i\omega}, \hat{\theta})$ being estimated for noise intensity $\lambda = 5$ ($\cdots$), $\lambda = 1$ ($--$) and $\lambda = 0.2$ ($-\cdot-$).

3.2.5   The closed-loop identification problem

In case of a closed-loop identification, noise is fed back to the input. As indicated in Example 3.2-8, a model estimated on the basis of the input u and output data y directly, depends on the noise acting on the feedback connection. No explicit and tunable expression is obtained for the misfit between the plant and a model being estimated and even the partial consistent identification problem mentioned in Definition 3.2-7 cannot be handled in case the identification uses the closed-loop input u and output y of the plant P_o directly. For the direct identification, two situations can be distinguished that are able to deal with the partial consistency identification problem.

- $\lambda = 0$

  This means that no noise is present on the closed-loop data. In that case, (3.33) modifies into

  $$\Phi_\varepsilon(\omega, \theta) = |P_o(\mathrm{e}^{i\omega}) - P(\mathrm{e}^{i\omega}, \theta)|^2 |S_{in}(\mathrm{e}^{i\omega})|^2 \Phi_r(\omega) \qquad (3.39)$$

  that is similar to the integrand of (3.31) as $\Phi_u(\omega) = |S_{in}(\mathrm{e}^{i\omega})|^2 \Phi_r(\omega)$ in case no noise is present on the input signal u. The same results, as obtained for the open-loop situation, can be applied to the closed-loop identification in the

absence of noise. A closed-loop identification without noise can be treated as an open-loop identification problem. An explicit and tunable expression for the misfit between the model $P(\theta)$ and the plant P_o is obtained, although the misfit does depend on the unknown transfer function of the input sensitivity function S_{in} in (3.39).

- $H(q, \eta^\star) = H_o(q)$

Choosing a fixed noise filter $H(\eta^\star)$ equal to the actual (but unknown) noise filter H_o modifies (3.33) into

$$
\begin{aligned}
\Phi_\varepsilon(\omega, \theta) = |P_o(\mathrm{e}^{\,i\omega}) - P(\mathrm{e}^{\,i\omega}, \theta)|^2 \frac{|S_{in}(\mathrm{e}^{\,i\omega})|^2 \Phi_r(\omega)}{|H(\mathrm{e}^{\,i\omega}, \eta^\star)|^2} + \\
|1 + P(\mathrm{e}^{\,i\omega}, \theta) C(\mathrm{e}^{\,i\omega})|^2 |S_{out}(\mathrm{e}^{\,i\omega})|^2 \lambda
\end{aligned}
\tag{3.40}
$$

Although (3.40) is slightly more complicated than (3.39), it can be observed that $P(\theta^\star) = P_o$ in case $P_o \in \mathcal{G}$. As

$$
\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_\varepsilon(\omega, \theta)\, d\omega \geq \lambda
\tag{3.41}
$$

the minimum of (3.41) is attained at $P(\theta^\star) = P_o$ as $(1 + P_o C) S_{out} = 1$. Unfortunately, in case of approximate identification (3.40) does not provide a tunable expression for the misfit between a model $P(\theta)$ and the plant P_o as the level λ of the noise still determines the trade off between the two terms in (3.40).

Actually, having both $H(q, \eta^\star) = H_o$ and $P_o \in \mathcal{G}$ is equivalent to the situation of $T_o \in \mathcal{M}$ addressing the consistent identification problem. The result on consistent identification in the presence of feedback is known to hold (Ljung 1987).

Both situations discussed above cannot be emulated in many practical situations. Noise free data is considered to be unavailable, whereas choosing a noise filter $H(\eta^\star) = H_o$ can be considered as an identification problem involving a consistent estimation of the noise filter dynamics. As such, the partial consistent identification problem for closed-loop data needs to be addressed in more detail. Additionally, an identification on the basis of closed-loop data that allows an explicit tuning of the misfit between model and plant is required as a design tool for estimating an approximate model of the plant P_o.

In the direct identification, only the input u and output y of the (possibly unstable) plant P_o are used for identification. In order to formalize the closed-loop identification problem discussed here, the information to be used for the closed-loop identification is not restricted solely to the signals u and y. Additional knowledge of either the reference signals r_1 and r_2, the controller C or any other information of the

components in (3.4) remains open and in this perspective, the following closed-loop identification problem is considered here.

Definition 3.2-9 *Consider the model set $\mathcal{G}$ given in (3.19) and the assumptions posed in Assumption 3.2-2. Based on the available knowledge of signals in (3.4) or additional information on the feedback connection $\mathcal{T}(P_o, C)$, the closed-loop identification problem is defined as the estimation of a (nominal) model $\hat{P} = P(\hat{\theta}_N)$ for which the limiting model $P(\theta^\star)$ in (3.24) should satisfy the following requirements.*

- *The closed-loop partial consistent estimation problem: in case $P_o \in \mathcal{G}$, the model $P(\theta^\star)$ should satisfy $P(\theta^\star) = P_o$.*

- *The closed-loop approximate estimation problem: in case $P_o \notin \mathcal{G}$, an expression should be obtained in which the misfit between the plant P_o and a model $P(\theta^\star)$ of specified limited complexity is tunable and can be made independent of the noise v present on the closed-loop data.*

The first item in Definition 3.2-9 addresses the partial consistent identification problem mentioned also in Definition 3.2-7. The second item is concerned with the approximate identification of the plant P_o. Specifically, an explicit and tunable expression for the misfit or bias of the model that approximates the plant P_o is formulated as a desired property.

Although the estimation of the noise filter H_o is not explicitly mentioned in Definition 3.2-9, both items (may) incorporate the approximate identification of a noise model, as $T_o \in \mathcal{M}$ is not required. As mentioned before, the plant P_o is a crucial element in the construction of a feedback connection $\mathcal{T}(P_o, C)$ and therefore the main attention is focused on modelling the plant P_o in this thesis.

3.3   Approaches to the Closed-Loop Identification Problem

3.3.1   Exploiting the knowledge of feedback

In the classical approaches to closed-loop identification such as the joint-input-joint-output method (Anderson and Gevers 1982, Söderström and Stoica 1989) or the direct identification (Ljung 1987, Ljung 1993a) as discussed in Section 3.2.4, neither the (exact) knowledge of the reference signals r_1 and r_2 nor the feedback controller C is used. Only the input u and output y signals are used for identification purposes.

Although the classical approaches provide consistency results in case $T_o \in \mathcal{M}$, these approaches cannot deal with the partial consistent or the approximate closed-loop identification problems mentioned in Definition 3.2-9. During the last decade, the approximate identification on the basis of closed-loop data has considered to be of more practical importance. As a result, alternative methods to deal with data obtained under feedback have been proposed. In light of the closed-loop identification

problem mentioned in Definition 3.2-9, some of these methods will be summarized in this section.

As mentioned in Definition 3.2-9, available knowledge of the signals in (3.4) or additional information on the feedback connection $\mathcal{T}(P_o, C)$ can be used for the identification on the basis of closed-loop data. Exploiting the freedom in specifying this knowledge enables the formulation of alternative solutions to the closed-loop identification problem. In order to delineate the various approaches to tackle the closed-loop identification problem, the following distinction is being made.

Remark 3.3-1 *The available knowledge of the signals in (3.4) and the additional information on the feedback connection $\mathcal{T}(P_o, C)$ is considered to consists of the following items.*

- *Knowledge of input signal u and/or output signal y.*

- *Knowledge of the reference signals r_1 and/or r_2.*

- *Knowledge of the controller C.*

In most of the closed-loop identification procedures discussed here, the (possibly disturbed) input u and output y signal of the plant P_o are assumed to be known and measurable in order to estimate a model of the plant. Requirements on the knowledge of the reference signals r_1 and/or r_2 can refer to either known and measurable reference signals or knowledge regarding the presence and excitation properties of reference signals. In the latter, the reference signals are considered to be unknown, as they are not used for identification purposes but merely for excitation of the closed-loop system. Finally, the knowledge of the (possibly unstable) controller that creates a stabilizing feedback connection $\mathcal{T}(P_o, C)$ can be used in the identification of models. In accordance with Assumption 2.2-4, the knowledge of the linear controller C is given by either a transfer function or a state space representation.

It should be noted that knowledge of the signals u and/or y, the reference signals r_1 and r_2 or the controller C cannot be considered to be independent. With (3.3) it can be seen that

$$r_1 + u = C(r_2 - y)$$

and implies that knowledge of C the reference signals r_1 and r_2 and the output signal y can be used to compute u. Alternatively, knowledge of the reference signals r_1 and r_2 and the input u and output y signals implies knowledge of the controller C, as $r_2 - y$ and $r_1 + u$ are respectively the noise free input and output signals of C.

A further classification of the different approaches of closed-loop (approximate) identification discussed in this chapter, is based on the following properties that involve stability.

Remark 3.3-2 *With respect to stability of the plant P_o and the controller C the following items play a role while estimating a model on the basis of closed-loop data.*

- *The possibility to consistently estimate unstable plants P_o.*

- *The ability to deal with closed-loop experiments obtained with an unstable controller C.*

- *The guarantee that a model $\hat{P}$ being estimated is stabilized by the controller C used in the feedback connection $\mathcal{T}(P_o, C)$.*

Both the plant and the controller are allowed to be unstable, as long as the feedback connection $\mathcal{T}(P_o, C)$ is stable. Consequently, a closed-loop identification technique should also be able to identify an unstable plant P_o, on the basis of time domain data obtained via a stabilizing feedback connection $\mathcal{T}(P_o, C)$. The ability to deal with an unstable controller C during the identification procedure is due to the fact that many controllers, like a PID, are equipped with an integrator to provide steady-state tracking of the reference signal r_2 (Maciejowski 1989, Boyd and Barrat 1991). As a consequence, the controller is marginally stable and the identification procedure should be able to deal with such controllers. Finally, the possibility to guarantee that a model found by a system identification technique is guaranteed to be stabilized by the controller C, used during the closed-loop experiments, is sometimes preferable. A model of a plant P_o that is not stabilized by a controller C, while the same controller is used in a stable feedback connection $\mathcal{T}(P_o, C)$, will certainly cast doubts on the quality of the model.

Basically, three main approaches to address the closed-loop identification problem are discussed in the remaining part of this section. These three approaches are split up in the following distinguishable closed-loop identification methods.

1. Indirect identification

2. Customized identification

3. Two-stage identification

The above mentioned methods provide valuable tools and insights for the items associated to the closed-loop identification problem mentioned in Definition 3.2-9. The items mentioned in Remark 3.3-1 and Remark 3.3-2 will be quoted in the discussion of the closed-loop identification procedures presented in the remaining part of this section. Each of the different methods will be evaluated for their usefulness considering the closed-loop identification problem mentioned in Definition 3.2-9 and the above mentioned items.

3.3.2 Indirect identification

Opposite to the direct identification mentioned in Section 3.2.4, an alternative to deal with closed-loop experiments is called indirect identification. This method is based on the idea of first estimating a closed-loop transfer function and then recalculating the model by using the knowledge of the controller C present in the estimated closed-loop transfer function. Summarizing, an indirect identification method consists of the following two steps.

1. Identify a closed-loop transfer function by considering the reference r_1 and/or r_2 as input signals and the plant signals u and/or y as output signals.

2. Recalculate an open-loop model $\hat{P}$ (and a noise filter model $\hat{H}$) from the identified closed-loop transfer functions (possibly by using the knowledge of the controller C)

Referring to the Remark 3.3-1, the indirect method of closed-loop identification requires the knowledge (known and measurable) of at least one of the reference signals r_1 or r_2 and one of the plant signals u or y. These signals are used to estimate a closed-loop transfer function. Additionally, the knowledge of the controller is needed to recompute an (open-loop nominal) model $\hat{P}$ on the basis of identified closed-loop transfer functions.

A model $\hat{P}$ is estimated indirectly, as first a closed-loop transfer function is being estimated from which the model $\hat{P}$ is computed. The crucial part is the reconstruction of a (nominal) model $\hat{P}$ from the identified closed-loop transfer function. Both steps are discussed in more detail.

Estimation of closed-loop transfer function

It can be observed from (3.2) that essentially four possible closed-loop input-output transfer functions and two possible closed-loop noise models can be accessed on the basis of closed-loop data. The four different transfer functions are given by the entries of the $T(P_o, C)$ matrix given in (2.5), while the two different noise filters consists of $S_{out}H_o$ and $-CS_{out}H_o$ as given in (3.4).

The first step in indirect identification is the estimation of a closed-loop transfer function. For notational convenience, the four available closed-loop transfer functions in $T(P_o, C)$ are denoted by the short hand notation

$$G_o = \begin{bmatrix} G_{o,11} & G_{o,12} \\ G_{o,21} & G_{o,22} \end{bmatrix} := \begin{bmatrix} P_o S_{in} C & P_o S_{in} \\ S_{in} C & S_{in} \end{bmatrix} \qquad (3.42)$$

while the two different noise filters are denoted by

$$L_o = \begin{bmatrix} L_{o,1} \\ L_{o,2} \end{bmatrix} := \begin{bmatrix} I \\ -C \end{bmatrix} (I + P_o C)^{-1} H_o \qquad (3.43)$$

Similar to (3.10) a prediction error model can be defined that uses the reference signals $col(r_2, r_1)$ as inputs and $col(y, u)$ as outputs. Consequently, a prediction error model can be expressed as follows

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = G(q, \theta) \begin{bmatrix} r_2(t) \\ r_1(t) \end{bmatrix} + L(q, \theta)\varepsilon(t, \theta) \tag{3.44}$$

where $G(q, \theta)$ and $L(q, \theta)$ are used to model respectively the closed-loop input-output transfer function G_o given in (3.42) and the closed-loop noise filter L_o given in (3.43). Clearly, depending on the available knowledge of the reference signals r_1, r_2 and the plant signals u, y, either all or parts of G_o or L_o can be estimated.

The reasoning to estimate a closed-loop transfer function first, is induced by the fact that the identification does not suffer from undesirable bias effects. The reference signals that are used for the identification are not correlated with the noise acting on the feedback connection $\mathcal{T}(P_o, C)$ as in a direct identification. As a result, the estimation of a closed-loop transfer function reduces to a standard open-loop identification problem, as discussed in Section 3.2.3. The only difference is the estimation of a (closed-loop) input-output model $\hat{G}$ and noise filter $\hat{L}$, instead of estimating an (open-loop) model $\hat{P}$ and noise filter $\hat{H}$ directly.

Under the same conditions as in the open-loop case, a consistent estimation can be obtained of a closed loop transfer function. Furthermore it should be noted that not all the closed loop transfer functions in (3.42) need to be estimated. Similar to the open-loop case, an OE-model structure in (3.44) allows a separate consistent estimation, or partial consistent estimation see Definition 3.2-7, of the closed-loop input-output transfer function G_o given in (3.42).

Computation of a model

From the estimated closed-loop transfer functions, either an estimate of the plant P_o, noise filter H_o or the controller C can be computed. This constitutes the second step in an indirect closed-loop identification. The reconstruction of the different elements within the feedback connection $\mathcal{T}(P_o, C)$ can be done in different ways.

It can be observed from (3.42) that the controller C can be reconstructed via $C = G_{o,22}^{-1} G_{o,21}$. Hence, estimating models $\hat{G}_{22}$ and $\hat{G}_{21}$ to model respectively $G_{o,22}$ and $G_{o,21}$ can be used to come up with an estimate of the controller given by $\hat{G}_{22}^{-1}\hat{G}_{12}$. In a similar way, a model $\hat{P}$ of P_o and a noise model $\hat{H}$ of H_o can be reconstructed from the estimates $\hat{G}_{11}$, $\hat{G}_{12}$, $\hat{G}_{22}$ and $\hat{L}_1$ via

$$\begin{aligned} \hat{P} &= \hat{G}_{12}\hat{G}_{22}^{-1} \\ \hat{H} &= (I - \hat{G}_{11})^{-1}\hat{L}_1 \end{aligned} \tag{3.45}$$

These estimates can be computed, provided that the inverse of $\hat{G}_{22}$ and $(I - \hat{G}_{11})$ are well-defined. As $\hat{G}_{22}$ and $(I - \hat{G}_{11})$ are estimates of respectively the input sensitivity

$(I + CP_o)^{-1}$ and the output sensitivity $(I + P_oC)^{-1}$ invertibility can be guaranteed, since $\mathcal{T}(P_o, C)$ is assumed to be a well-posed feedback connection. Finally, some remarks can be given with respect to the monicity of the constructed noise filter $\hat{H}$. Clearly, the estimate of $\hat{H}$ is monic if $(I - \hat{G}_{11})^{-1}$ is monic, as $\hat{L}_1$ is estimated as a monic stable noise filter. As $(I - \hat{G}_{11})$ is an estimate of the input sensitivity function $(I + CP_o)^{-1}$, the inverse $(I - \hat{G}_{11})^{-1}$ is likely to be monic. In fact, if the model $\hat{G}_{11}$ does not have a feedthrough term, the inverse $(I - \hat{G}_{11})^{-1}$ is monic.

Procedures that estimate all the closed loop transfer functions in (3.42) in order to reconstruct an estimate of P_o and possibly C and H_o, have been presented in Bretthauer and Heckert (1991) and Verhaegen (1993). In Verhaegen (1993) a subspace identification method is used that aims at estimating the closed-loop transfer function G_o given in (3.42) and/or the closed-loop noise filter L_o in (3.43) consistently. Subsequently, models for P_o, C (and H_o) can be reconstructed as mentioned above.

In case the controller C is known, this knowledge can be used to recompute models of P_o (or the noise filter H_o). Instead of estimating multiple input-output transfer functions, it suffices to estimate only one input-output closed-loop transfer function (Van den Hof and de Callafon 1996). The result has been summarized in the following proposition.

Proposition 3.3-3 *Let $\hat{G}_{11}$, $\hat{G}_{12}$, $\hat{G}_{21}$ and $\hat{G}_{22}$ denote estimates of the input-output closed-loop transfer functions given in (3.42) while $\hat{L}_1$ and $\hat{L}_2$ denote estimates of the closed-loop noise filters given in (3.43). Under the assumption that $\mathcal{T}(\hat{P}, C)$ is a well-posed feedback connection, the following reconstructions of an (open-loop) model $\hat{P}$ and an (open-loop) noise filter $\hat{H}$ can be considered by using the knowledge of the controller C.*

(a) $\hat{G}_{11} = \hat{P}(I + C\hat{P})^{-1}C$ and $\hat{L}_1 = (I + \hat{P}C)^{-1}\hat{H}$ imply that $\hat{Q} := (I - \hat{G}_{11})^{-1}$ is well defined and

$$\begin{aligned} \hat{P} &= (\hat{Q} - I)C^\dagger \\ \hat{H} &= \hat{Q}\hat{L}_1 \end{aligned} \tag{3.46}$$

provided that C has a right inverse $C^\dagger$

(b) $\hat{G}_{12} = \hat{P}(I + C\hat{P})^{-1}$ and $\hat{L}_1 = (I + \hat{P}C)^{-1}\hat{H}$ imply that $\hat{Q} := (I - \hat{G}_{12}C)^{-1}$ is well defined and

$$\begin{aligned} \hat{P} &= \hat{Q}\hat{G}_{12} \\ \hat{H} &= \hat{Q}\hat{L}_1 \end{aligned} \tag{3.47}$$

(c) $\hat{G}_{21} = (I + C\hat{P})^{-1}C$ and $\hat{L}_2 = -C(I + \hat{P}C)^{-1}\hat{H}$ imply that $\hat{Q} := (\hat{G}_{21}C^{-1})^{-1}$ is well defined and

$$\begin{aligned} \hat{P} &= C^{-1}(\hat{Q} - I) \\ \hat{H} &= -C^{-1}\hat{Q}\hat{L}_2 \end{aligned} \tag{3.48}$$

provided that C has an inverse C^{-1}

(d) $\hat{G}_{22} = (I + C\hat{P})^{-1}$ *and* $\hat{L}_2 = -C(I + \hat{P}C)^{-1}\hat{H}$ *imply that* $\hat{Q} := (\hat{G}_{22})^{-1}$ *is well defined and*

$$
\begin{aligned}
\hat{P} &= C^{\dagger}(\hat{Q} - I) \\
\hat{H} &= -C^{\dagger}\hat{Q}\hat{L}_2
\end{aligned}
\tag{3.49}
$$

provided that C has a left inverse $C^{\dagger}$

Proof: It can be verified that the transfer function $\hat{Q}$ denotes either the input sensitivity $(I+C\hat{P})^{-1}$ or output sensitivity $(I+C\hat{P})^{-1}$ for which the inverse is well defined. The subsequent expressions for $\hat{P}$ and $\hat{H}$ are found by algebraic manipulations. $\square$

The reconstructions of a (nominal) model $\hat{P}$ and noise filter $\hat{H}$ in Proposition 3.3-3 point to the different situations that may occur in the application of indirect closed-loop identification. The different situations depend on the availability of the reference signals r_1, r_2 and plant signals u, y used during the first step of an the indirect identification.

The reconstruction in (a) occurs when the reference signal r_2 is used as an input signal and y is used as an output signal. In that case $G_{11} = P_o S_{in} C$ and $L_1 = S_{out} H_o$ need to be estimated. In a similar way, the other situations can be characterized in terms of the use or availability of the reference signals r_1 or r_2 and the plant signal u or y. In the following, the attention is focused on the estimation of a (nominal) model $\hat{P}$ as this is of main interest for the closed-loop identification problems mentioned in Definition 3.2-9.

Special attention deserves situation (b) mentioned in Proposition 3.3-3. It can be observed that for (3.47) no conditions on the controller C are posed. In all other situations, either a left inverse, right inverse or both (an inverse) of the controller must be computable in order to reconstruct the model. Situation (b) reflects the situation in which r_1 is used as an input and y as an output signal in the first step of the indirect identification. With (3.3), this is also the situation where the signal r can be considered to be an input signal while identifying the closed-loop transfer functions G_{22} and L_1. As mentioned in Remark 3.2-3, using the signal r directly requires the controller C to be stable in order reconstruct r from the available r_1 and r_2.

As the reconstruction in (3.47) does not pose any conditions on the controller C, most of the indirect closed-loop identification schemes are based on the estimation of the closed-loop transfer function $G_{o,12} = P_o S_{in}$ (Ljung *et al.* 1974, Söderström and Stoica 1975, Zhu *et al.* 1988, Söderström and Stoica 1989, Zheng and Feng 1995). It should be noted that not every indirect method uses the reconstruction of a model $\hat{P}$ accordingly to Proposition 3.3-3. In Zheng and Feng (1995) the relation between the parameters of an open-loop model and an estimated closed-loop transfer function is

used to recompute an open-loop model $\hat{P}$. In some of the closed-loop indirect identification schemes, such as Zhu *et al.* (1988), the reconstruction of a nominal (open-loop) model $\hat{P}$ is omitted and the estimated closed-loop transfer function is considered to be a model of the plant P_o operating in feedback. In that case, knowledge of the controller C is not required as no open-loop model $\hat{P}$ is being computed. Although the other reconstructions may be used in case the appropriate conditions on the controller C are satisfied, an indirect identification that uses either (3.47) of (3.45) will be considered in the sequel.

Evaluation

As the first step of the indirect identification is just an open-loop identification problem, a consistent estimate $\hat{G}$ of a closed-loop transfer function G_o can be obtained, provided that the reference signals are persistently exciting (Assumption 3.2-2). Subsequently, either from (3.45) or (3.47) an estimate $\hat{P}$ of P_o can be obtained. It can be verified that $G_{o,12} = \hat{G}_{12}$ implies $P_o = \hat{P}$, hence the indirect identification is able to deal with the closed-loop partial consistent identification problem.

In order to get a consistent estimation of a closed-loop transfer function, a model set $\mathcal{M}$ must be used that is able to capture both the dynamics of the plant P_o and the controller C. Provided that no pole/zero cancellations occur between the plant P_o and the controller C, the McMillan degree n_g of the closed-loop transfer functions given in (3.42) is the sum of the McMillan degree n_o of P_o and the McMillan degree n_c of C. Consequently, a consistent identification of a closed-loop transfer function in the first step of an indirect method requires a model set $\mathcal{M}$ that parametrizes models having a McMillan degree $n_g = n_o + n_c$ that is higher than needed in order to model the plant P_o.

Additionally, a (nominal) model $\hat{P}$ computed via (3.45) or (3.47), in general will have a McMillan degree that is larger than or equal to the McMillan degree n_g of the estimated closed-loop transfer function. Possible pole/zero cancellations (unobservable or uncontrollable modes) must be eliminated to find a low order model $\hat{P}$. This also effects the closed-loop approximate identification problem mentioned in Definition 3.2-9. Referring to the construction of $\hat{P}$ described in (3.47), an approximate identification of $G_{o,12}$, using an OE-model $G_{12}(\theta)$ similar to (3.15), leads to the following frequency domain expression for the variance of the prediction error

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |G_{o,12}(e^{\,i\omega}) - G_{12}(e^{\,i\omega}, \theta)|^2 \Phi_r(\omega)\, d\omega \qquad (3.50)$$

Clearly, the integral expression in (3.50) is similar to (3.31) and yields a tunable expression for the misfit between $G_{o,12}$ and the closed-loop transfer function model $\hat{G}_{12}$ being estimated. Furthermore, the approximate identification is independent of the noise v acting on the feedback connection $\mathcal{T}(P_o, C)$.

There is however an essential difference between the integral expressions of (3.31) and (3.50). Although $\hat{G}_{12} = \hat{P}(I + C\hat{P})^{-1}$ implies $\hat{P} = (I - \hat{G}_{12}C)^{-1}\hat{G}_{12}$, the transfer function $G_{12}(\theta)$ in (3.50) is not parametrized according to $P(\theta)(I + CP(\theta))^{-1}$. In other words, the tuning of the approximate identification of $\hat{P}$ is clear from (3.50), but the set of models $\mathcal{M}$ over which the optimization takes place is different. This is due to the fact that the model $P(\theta)$ in the direct identification is parametrized as $P(\theta) = (I - \hat{G}_{12}(\theta)C)^{-1}\hat{G}_{12}(\theta)$. As a result, the estimation of a low order (nominal) model becomes difficult as the McMillan degree of the model $P(\theta)$ is influenced by the order of $G(\theta)$ and he controller C. Consequently, in general the indirect identification cannot deal with the closed-loop approximate identification problem mentioned in Definition 3.2-9.

Referring to Remark 3.3-2, it can be observed that the indirect method is able to estimate possibly unstable models. In fact, there is no restriction on the stability of the nominal model $\hat{P}$ being constructed. Furthermore, it is possible to deal with closed-loop data obtained from a feedback connection $\mathcal{T}(P_o, C)$ in which both the plant P_o and the controller C may be unstable as the estimation is done via stable closed-loop transfer functions. However, the different possibilities of Proposition 3.3-3 or the construction of a signal r in (3.3) as mentioned above, come with additional restrictions on the controller C. These conditions might include the existence of a left and/or right inverse or stability of C.

Although stability of the nominal model $\hat{P}$ is not enforced, stability of $\mathcal{T}(\hat{P}, C)$ can be guaranteed in specific cases (Van den Hof and de Callafon 1996). These specific cases can be found by posing stability conditions on the controller C that is used in the feedback connection $\mathcal{T}(P_o, C)$ when gathering data for closed-loop observations. One of these special cases, that can occur in many practical situations, is found when the controller C is stable.

Corollary 3.3-4 *Consider situation (b) of Proposition 3.3-3 where the controller C satisfies $C \in R\mathcal{H}_\infty$ and let the model $\hat{P}$ being constructed according to (3.47). Then $T(\hat{P}, C) \in R\mathcal{H}_\infty$ if and only if $\hat{G}_{12} \in R\mathcal{H}_\infty$.*

Proof: Investigation of $T(\hat{P}, C) \in R\mathcal{H}_\infty$ in case $C \in R\mathcal{H}_\infty$ was discussed in Corollary 2.2-11, from which the result immediately follows.   □

In case the controller C is not stable, but the inverse of C exist and is stable, a similar corollary can be given by considering situation (c) in Proposition 3.3-3.

Corollary 3.3-5 *Consider situation (c) of Proposition 3.3-3 and let the model $\hat{P}$ being constructed according to (3.48) where the inverse C^{-1} satisfies $C^{-1} \in R\mathcal{H}_\infty$. Then $T(\hat{P}, C) \in R\mathcal{H}_\infty$ if and only if $\hat{G}_{21} \in R\mathcal{H}_\infty$.*

Proof: Necessity is obvious. Sufficiency is found by considering $\hat{G}_{12} = (I + C\hat{P})^{-1}C \in R\mathcal{H}_\infty$ and $C^{-1} \in R\mathcal{H}_\infty$. Then $\hat{G}_{12}C^{-1} = (I + C\hat{P})^{-1} \in R\mathcal{H}_\infty$, $I - \hat{G}_{12}C^{-1} = \hat{P}(I+C\hat{P})^{-1}C \in R\mathcal{H}_\infty$ and $(I - \hat{G}_{12}C^{-1})C^{-1} = \hat{P}(I+C\hat{P})^{-1} \in R\mathcal{H}_\infty$.
$\square$

The above mentioned corollaries provide stability results for the feedback connection $\mathcal{T}(\hat{P}, C)$ in case a specific indirect closed-loop identification is being used. The indirect method to closed-loop identification gives a possibility to deal with the closed-loop partial consistent estimation problem as mentioned in Definition 3.2-9. However, the approximate closed-loop identification is not fully covered. Although it is clear how the nominal model is tuned in terms of a frequency domain expression, an explicit tuning of a fixed order approximate nominal model $\hat{P}$ becomes intractable.

3.3.3   Customized identification

Mainly due to the fact that a closed-loop transfer function $G(\theta)$ is not parametrized in terms of the model $P(\theta)$ to be estimated, the McMillan degree of the model $\hat{P}$ found by a recomputation will generally be larger than the McMillan degree of the estimated closed-loop transfer function. As mentioned before, this parametrization issue makes an explicit tuning of a fixed order approximate model $\hat{P}$ intractable.

Customized parametrization

A natural extension of the indirect method to accommodate the approximate closed-loop identification problem is to perform an estimation of a closed-loop transfer function $G(\theta)$ in which a model $P(\theta)$ has been parametrized explicitly. In this way a customized or so-called tailor-made parametrization (Landau and Boumaïza 1996, Donkelaar and Van den Hof 1996) is used while estimating a closed-loop transfer function. In the light of the PE-framework, such a customized identification involves the minimization of a prediction error

$$\varepsilon(t, \theta) = \begin{bmatrix} y(t) \\ u(t) \end{bmatrix} - G(q, \theta) \begin{bmatrix} r_2(t) \\ r_1(t) \end{bmatrix} \tag{3.51}$$

where $G(q, \theta)$ is parametrized in conformance with the map (2.5) and given by

$$G(q, \theta) = \begin{bmatrix} P(q, \theta) \\ I \end{bmatrix} (I + C(q)P(q, \theta))^{-1} \begin{bmatrix} I & C(q) \end{bmatrix}, \ \theta \in \Theta \tag{3.52}$$

or

$$G(q, \theta) = \begin{bmatrix} P(q, \theta) \\ I \end{bmatrix} (I + C(q)P(q, \theta))^{-1}, \ \theta \in \Theta \tag{3.53}$$

in case r from (3.3) is considered to be an input signal to the OE-model (3.51). Alternatively, only specific parts of the closed-loop transfer function $G(q, \theta)$ in (3.52) need to be parametrized, in case only one of the reference signals $col(r_2, r_1)$ or one of the plant signals $col(y, u)$ is available. In Donkelaar and Van den Hof (1996) such a (SISO) customized identification is performed where the closed-loop transfer function to be estimated is restricted to $G_{o,12}$. Consequently, the tailor-made parametrization reads as follows

$$G_{12}(q, \theta) = \frac{P(q, \theta)}{1 + C(q)P(q, \theta)}, \; \theta \in \Theta \qquad (3.54)$$

and involves the estimation of only one of the transfer function given in (3.52).

The estimation of $G_{o,12}$ corresponds to the scenario reflected by option (b) in Proposition 3.3-3 but in case of a customized identification, $G_{12}(q, \theta)$ is equipped with the customized parametrization (3.54). Similar to the indirect method of closed-loop identification, knowledge (known and measurable) of the reference signals r_1 and/or $r2$, the plant signals u and/or y and knowledge of the controller C is required in case a customized identification needs to be performed.

Clearly, a least-squares minimization of the prediction error $\varepsilon(t, \theta)$ in the OE-model (3.51) requires a dedicated non-linear optimization. Even in the case when a model $P(\theta)$ is restricted to be a FIR-model (3.16) or the poles of the model $P(\theta)$ are fixed in case of an ORTFIR-model (3.17), the least-squares minimization of $\varepsilon(t, \theta)$ in (3.51) requires a non-linear minimization, whereas in an open-loop based situation a linearly parametrized model structure can be used to facilitate the least-squares estimation.

Additionally, it can be observed from (3.51) that $G(q, \theta)$ is required to be stable for all $\theta \in \Theta$ in order to have a well-defined and bounded prediction error $\varepsilon(t, \theta)$. As such, it must be verified whether or not restricting $G(q, \theta)$ to be stable is going to effect the non-linear optimization. Requiring stability of $G(q, \theta)$ may result in an allowable parameter space Θ that is not (path-wise) connected (Ljung 1987), which will greatly complicate the optimization. Restricting $G_{12}(q, \theta)$ to be stable, in Donkelaar and Van den Hof (1996) a sufficient condition on the path-wise connectivity of Θ is formulated in terms of the order or McMillan degree of the model $P(q, \theta)$ and the controller $C(q)$. As long as the order n_p of the parametrized model $P(q, \theta)$ is larger than or equal to the order n_c of the controller, the parameter set Θ is (path-wise) connected. As such, the model to be estimated must have a higher complexity than the controller used for the closed-loop experiments in order to ensure connectivity of the parameter set Θ.

Evaluation

Provided that the reference signals $col(r_2, r_1)$ are persistently exciting, which is mentioned in Assumption 3.2-2, and the parametrization of $G(q, \theta)$ is rich enough to capture G_o, minimizing (a norm of) the difference between $G_o(q)$ and the customized

parametrization $G(q,\theta)$ given in (3.52) must yield an estimate $\hat{G} = G(q,\theta^\star) = G_o$ for the limiting case $N \to \infty$. Clearly, $\hat{G} = G_o$ implies $\hat{P} = P_o$ and the customized identification is able to deal with the partial consistent identification problem as mentioned in Definition 3.2-9.

To discuss the approximate closed-loop identification problem, consider the tailor-made parametrization given in (3.54). For the limiting case mentioned in (3.25) and (3.26), minimizing the least-squares prediction error using the tailor-made parametrization of (3.54) gives rise to the following frequency domain expression

$$\min_{\theta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |\frac{P_o(\mathrm{e}^{\,i\omega})}{1 + C(\mathrm{e}^{\,i\omega})P_o(\mathrm{e}^{\,i\omega})} - \frac{P(\mathrm{e}^{\,i\omega},\theta)}{1 + C(\mathrm{e}^{\,i\omega})P(\mathrm{e}^{\,i\omega},\theta)}|^2 \Phi_r(\omega)\, d\omega \qquad (3.55)$$

It can be observed from (3.55) that an expression for the misfit between P_o and the model $P(\theta)$ to be estimated is obtained in which the difference between two closed-loop transfer functions is being minimized for a model $P(\theta)$ of specified limited complexity.

Although, (3.55) does not characterize explicitly an additive difference between the plant P_o and the model $\hat{P}$ being estimated, it is clear that the approximate identification deals with finding a model $\hat{P}$ by matching two closed-loop transfer functions. In case of (3.55), $G_{12}(\theta)$ is used to approximate the closed-loop transfer function $G_{o,12}$, but this can be generalized to the case in which all transfer function $G(\theta)$ in (3.52) are used to approximate the map $col(r_2, r_1)$ to $col(y, u)$ given in (2.5). Furthermore, the approximate identification is independent of the noise v acting on the feedback connection $\mathcal{T}(P_o, C)$. As such, the customized identification is able to deal with the approximate closed-loop identification problem mentioned in Definition 3.2-9 at the sake of a (non-standard) non-linear optimization.

With respect to Remark 3.3-2, the same remarks as given for the indirect identification can be given. Clearly, an estimated stable closed-loop transfer function $G(\hat{\theta})$ as given in (3.52) is equivalent to $T(\hat{P}, C) \in R\mathcal{H}_\infty$. In case only a stable model $G_{12}(\hat{\theta})$ as given in (3.54) is found, similar to Corollary 2.2-11 and Corollary 3.3-4, stability of $T(\hat{P}, C)$ is guaranteed provided that $C \in R\mathcal{H}_\infty$.

3.3.4 Two-stage identification

Both the indirect and the customized identification use the knowledge of the controller respectively to recompute an open-loop plant model or to set up a specific customized parametrization. As such, both methods heavily rely on the knowledge of the controller C being used. An alternative method that does not use the knowledge of the controller C and is still able to perform an approximate identification of a model with a specified limited complexity is known under the name of two-stage method (Van den Hof and Schrama 1993).

In order to estimate a model on the basis of observations obtained from a feedback connection $\mathcal{T}(P_o, C)$, the two-stage method requires, as expected, two steps. Based on

the reduced form (3.4) that describes the data coming from the feedback connection $\mathcal{T}(P_o, C)$ the following two steps are considered.

1. Identify a model $\hat{S}_{in}$ of the input sensitivity S_{in} by considering the map from r_1 or r in (3.3) to the plant input u and simulate a so-called intermediate or input signal $u_r(t)$ via

$$u_r(t) = \hat{S}_{in}(q)r(t) \qquad (3.56)$$

2. Estimate a model $\hat{P}$ for P_o by considering the map from the noise free intermediate signal $u_r(t)$ to the output signal y of the plant P_o.

Compared to the indirect method, the open-loop model $\hat{P}$ of the plant P_o is being estimated on the basis of closed-loop data, instead of reconstructing it from the estimated closed-loop (sensitivity) transfer function. As the reference signals are uncorrelated with the noise v acting on the feedback connection $\mathcal{T}(P_o, C)$, the intermediate signal $u_r(t)$ is not perturbed by the noise v directly. Consequently the correlation of the input u with the noise v as mentioned in (3.4) is eliminated. As a result, the approximate identification of a model $\hat{P}$ can be decoupled from the noise v present on the closed-loop data.

Clearly, the two-stage method does not require the knowledge of the controller C. Only knowledge of the reference signals and the output signal y of the plant P_o are needed to perform the two-stage method. The knowledge of the reference signals might be limited to r_1 or r_2 in case either CS_{in} is being estimated in the first step or the reference signal r (3.3) is being used to estimate S_{in}. The two steps will be summarized in the following.

Estimation of sensitivity

The estimation of the (input) sensitivity function in the first step of the two stage method is used solely to reconstruct the (noise free) intermediate signal u_r mentioned in (3.56). For the simulation of u_r in (3.56) no noise model is required and the estimation of S_{in} can be limited to the estimation of an input-output model $\hat{S}_{in}$ only.

For this purpose, an OE-model structure

$$u(t) = S(q, \beta)r(t) + \varepsilon(t, \beta) \qquad (3.57)$$

can be used, where $r(t)$ can be either r_1 or r in (3.3). As mentioned above, instead of estimating a model for the input sensitivity, a model for CS_{in} can be identified in case the reference signal r_2 is available only.

Due to the OE-model structure and the open-loop character of (3.57), a consistent estimate $\hat{S}_{in} = S(q, \hat{\beta})$ can be obtained, irrespective of the noise present on the input signal u. In case a consistent estimate of the input sensitivity S_{in} is obtained, the

intermediate signal u_r in (3.56) will match the noise free part of the input signal u in (3.4), as

$$u(t) = S_{in}(q)r(t) - C(q)S_{out}(q)v(t) = u_r(t) - C(q)S_{out}(q)v(t).$$

where it is assumed that $\hat{S}_{in} = S_{in}$.

It should be noted that the estimate $\hat{S}_{in}$ is used solely to simulate the intermediate signal u_r and therefore the model $\hat{S}_{in}$ only acts as a filter. The actual estimation of a limited complexity (nominal) model that is able to deal with the closed-loop identification problem mentioned in Definition 3.2-9 is performed in the second step of the two stage method.

To accommodate the optimization involved in estimating the parameter β in (3.57), a linear parametrized model structure that maintains the OE-model structure of (3.57) can be used. As mentioned in Section 3.2.2, such a model structure can be either a FIR- or ORTFIR-model structure, respectively given in (3.16) and (3.17). Although the complexity (McMillan degree) of a FIR or ORTFIR-model can be reasonably high, the model $\hat{S}_{in}$ is used only for filtering. Successful application of such an ORFIR-model structure in the first step of the two-stage procedure has been reported for example in de Callafon *et al.* (1993).

Estimation of open-loop plant model

In the second step of the two-stage procedure, the intermediate signal u_r is used as an input signal to estimate a model $\hat{P}$ for the plant P_o. The use of u_r as an input signal opens the possibility to estimate a (nominal) model $\hat{P}$ in an open-loop way. This is due to the fact that u_r in (3.56) is uncorrelated with the noise v acting on the feedback connection $\mathcal{T}(P_o, C)$ and the output y in (3.4) can be written as

$$y(t) = P_o(q)S_{in}(q)u(t) + S_{out}(q)v(t) = P_o(q)u_r(t) + S_{out}(q)v(t)$$

in case $\hat{S}_{in}$ in (3.56) satisfies $\hat{S}_{in} = S_{in}$. Consequently, estimating a (nominal) model $\hat{P}$ of the plant P_o can be achieved by employing an OE-model structure

$$y(t) = P(q, \theta)u_r(t) + \varepsilon(t, \theta) \tag{3.58}$$

where the intermediate signal $u_r(t)$ is considered to be an input signal.

Note that the estimation of a model $\hat{P}$ by minimizing a least squares criterion on the prediction error $\varepsilon(t, \theta)$ in (3.58) again reduces to an open-loop identification problem, as u_r in (3.56) is uncorrelated with the noise v in (3.58). Consequently, the same results on both consistent and approximate identification as mentioned in Section 3.2.3 will apply to the second step of the two stage method.

Even in the case where the estimate $\hat{S}_{in}$ from the first step does not satisfy $\hat{S}_{in} = S_{in}$, the simulated intermediate signal u_r in (3.56) remains uncorrelated with the

noise v in (3.58). However, the situation in which $\hat{S}_{in} \neq S_{in}$ will effect the result on approximate identification, as u_r will not match the noise free part of the input signal u. This effect will be discussed in the following.

Evaluation

As the second step of the two stage procedure is just an open-loop identification problem, a consistent estimate $\hat{P}$ of the plant P_o can be obtained, provided that the reference signals are persistently exciting and a consistent estimate $\hat{S}_{in}$ of the input sensitivity is obtained in the fist step. In case $\hat{S}_{in}$ is an inaccurate, approximate or inconsistent estimate of the input sensitivity S_{in}, a consistent estimation in the second step of the two stage procedure requires the modelling of a more complicated closed-loop transfer function. This is due to the fact that

$$y(t) = P_o S_{in} r(t) + S_{out} v(t) = P_o S_{in} \hat{S}_{in}^{-1} u_r(t) + S_{out} v(t)$$

requiring a consistent estimation of $P_o S_{in} \hat{S}_{in}^{-1}$. Consequently, performing an accurate identification on the basis of the input signal u_r and output signal $y(t)$ might lead to the conclusion that the identified model $\hat{P}$ is a consistent estimation of the plant P_o, while actually a consistent estimation of $P_o S_{in} \hat{S}_{in}^{-1}$ is obtained. Henceforth, an accurate estimate, or even better a consistent estimate, of the input sensitivity S_{in} is required to attain a consistent estimation of the plant P_o(Van den Hof and Schrama 1993). Concluding, the two-stage procedure is able to deal with the partial consistent identification problem as mentioned in Definition 3.2-9.

Due to the similarity of the second step in the two-stage procedure with an ordinary open-loop identification problem, the approximate closed-loop identification problem of Definition 3.2-9 can be addressed quite easily. For the limiting case $N \to \infty$ mentioned in (3.24), an approximate identification of a model $\hat{P} = P(\theta^\star)$ is obtained that satisfies

$$\theta^\star = \arg\min_{\theta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_o(e^{i\omega}) S_{in}(e^{i\omega}) - P(e^{i\omega}, \theta) S(e^{i\omega}, \beta^\star)|^2 \Phi_r(\omega)\, d\omega \qquad (3.59)$$

where

$$\beta^\star = \arg\min_{\beta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_{in}(e^{i\omega}) - S(e^{i\omega}, \beta)|^2 \Phi_r(\omega)\, d\omega \qquad (3.60)$$

is the estimate $\hat{S}_{in} = S(\beta^\star)$ obtained in the first step of the two-stage procedure. In (3.59) and (3.60), the possibility that $\hat{S}_{in}$ is not a consistent estimate of the input sensitivity S_{in} has been incorporated. Clearly, if $\hat{S}_{in} = S_{in}$, the integrand of (3.59) reduces to

$$|[P_o(e^{i\omega}) - P(e^{i\omega}, \theta)] S_{in}(e^{i\omega})|^2 \Phi_r(\omega) \qquad (3.61)$$

which is again 'a clear-cut best mean-square approximation', similar as in the open-loop situation discussed in Section 3.2.3. Clearly, a tunable expression for the misfit

between the plant P_o and the model $\hat{P}$ of limited is obtained and as such, the two-stage procedure is able to deal with the approximate closed-loop identification problem mention in Definition 3.2-9. In case $\hat{S}_{in} \neq S_{in}$, the integrand of (3.59) can be rewritten as follows.

$$|[P_o(\mathrm{e}^{\,i\omega}) - P(\mathrm{e}^{\,i\omega}, \theta)]S_{in}(\mathrm{e}^{\,i\omega}) + P(\mathrm{e}^{\,i\omega}, \theta)[S_{in}(\mathrm{e}^{\,i\omega}) - S(\mathrm{e}^{\,i\omega}, \beta^\star)]|^2 \Phi_r(\omega) \qquad (3.62)$$

From (3.62) it can be observed that a term $P(\theta)[S_{in} - S(\beta^\star)]$ is introduced that will effect the fitting of $P(\theta)$ to P_o, weighted by S_{in}. Clearly, no explicit tunable expression for the misfit between P_o and $\hat{P}$ is obtained. However, this term can be made small by fitting $S(\beta)$ to S_{in} where $P(\theta)$ is relatively large.

Although the open-loop character of the second step in the two-stage procedure allows an explicit and tunable expression for the misfit between the plant P_o and the (nominal) model $\hat{P}$, it brings along one disadvantage. As the plant P_o is identified in an open-loop way, the estimation of an unstable plant P_o using PE-based methods is not always possible. This problem can be avoided if a model parametrization is used where both the input-output model $P(\theta)$ and the noise model $H(\theta)$ are parametrized dependently, see e.g. example 4.4 in Ljung (1987). Unfortunately, such a model parametrization is unfavourable in case an explicit tuning of the misfit between model $P(\theta)$ to be estimated and the plant is desired. As such, it can be mentioned that the two-stage method in general is not able to handle the (approximate) estimation of unstable plants.

In case the reference signal r_1 is available, the first step should estimate S_{in}, being the map from r_1 to u. Alternatively, in case r_2 is available, CS_{in} should be estimated in the first step. In case both r_1 and r_2 are available, the reconstruction of r in (3.3) is not possible in case the controller C is unstable, whereas the two-stage method aims at not using the knowledge of the controller explicitly. As the knowledge of the controller is not used to construct and/or parametrize the model $\hat{P}$, no statements with respect to the stability of the feedback connection $\mathcal{T}(\hat{P}, C)$ can be given beforehand.

3.4 Precis

The three main stream approaches described in the previous section address the problem of approximate identification on the basis of closed-loop data. Clearly, every approach has its favourable properties and its specific disadvantages. In case a specific disadvantage is taken for granted, the previously discussed methods can be well suited for approximate identification on the basis of closed-loop data.

Furthermore, in some special cases a disadvantage can be evaded. As an example, a memory-less (constant gain) feedback controller is not going to enlarge the McMillan degree of the model $\hat{P}$ being computed via e.g. (3.47). In that case, the indirect identification is well suited for the purpose of estimating (approximate) models on the basis of closed-loop data. Although the two-stage method does not require the

knowledge of the controller C to estimate an approximate (nominal) model $\hat{P}$ of the plant P_o, the customized identification is able to exploit this knowledge entirely at the sake of a slightly more complicated parametrization and associated optimization.

Although the above mentioned procedures provide tools to deal with the closed-loop identification problem mentioned in Definition 3.2-9, a more general approach that can be used without any restrictions is preferable. Such a general approach should cover special situations such as employing the possible knowledge of an unstable controller C and the estimation of a possibly unstable plant P_o. The approach should be able to estimate a (nominal) model $\hat{P}$ of limited complexity via an approximate identification consistent with Definition 3.2-9. Additionally, such an approach should allow a possible solution to Procedure 2.5-4, in which the identification and control are intertwined for the purpose of designing a robust and high performing controller. Not only a nominal model $\hat{P}$, but a set of models $\mathcal{P}$ needs to be estimated on the basis of closed-loop data. In the next chapter it will be clarified that a fractional approach is able to provide such a general solution to the closed-loop identification problem.

4

Identification Using Closed-Loop Data: a Fractional Approach

4.1 Fractional Representations

4.1.1 Motivations and background

One of the primary motivations to use fractional model representations is due to the fact that both stable or unstable systems can be represented by (the ratio of two) stable factors (Desoer *et al.* 1980, Vidyasagar 1985, Antsaklis 1986, Bakri 1988). To anticipate on the definitions given below, it can be said here that a (possibly unstable) system P is represented by $P = ND^{-1}$ where (N, D) constitutes a right fractional representation of P in which both N and D are BIBO stable mappings. As such, a unified approach to handle both stable and unstable systems can be formulated. The opportunity to deal with stable factorizations, can be exploited in the identification of an unstable plant P_o, as only its stable factorizations have to be estimated.

Additionally, the algebraic approach to fractional model representations of Desoer *et al.* (1980) or Vidyasagar (1985) has opened alternatives to study stability of interconnected systems, such as feedback connections $\mathcal{T}(P_o, C)$ (Nett 1986, Smith 1989). On the basis of this algebraic approach, also a set of all stabilizing controllers for a given system P can be characterized and is known as the Youla parametrization (Youla *et al.* 1976*b*, Youla *et al.* 1976*a*). This set of controllers is expressed in terms of possible factorizations of the controllers that are based on a factorization of the system P and a factor that is allowed to vary over all stable transfer functions. The results on the possibility to parametrize a set of stabilizing controllers have for example been used in the characterization of $\mathcal{H}_\infty$ optimal controllers (Doyle 1984, Francis 1987, McFarlane and Glover 1990). Alternatively, a factorizational approach is used frequently to discuss the problems associated to robust or simultaneous stabilization (Saeks and Murray 1982, Anantharam 1985, Obinata and Moore 1988, Sefton *et al.* 1990).

The possibility to represent an unstable system P by a stable factorization (N, D) is used in Vidyasagar *et al.* (1982) to define a graph of a system P and to set up a graph topology. One of the main results obtained with the graph topology shows that small perturbations of the graph of a system P do not violate the stability of a feedback connection $\mathcal{T}(P, C)$. As a result, approximation of an (unknown) plant P_o by a model $\hat{P}$ in the graph topology, either by considering the graph-metric (Vidyasagar 1984) or the gap-metric (El-Sakkary 1985, Packard and Helwig 1989), is considered to be a promising approximation in view of a feedback in which both P_o and $\hat{P}$ are operating.

Not surprisingly, the application of the algebraic framework of fractional model representation for the purpose of system identification has been recognized and used by several researchers. The possibility to be able to estimate an unstable plant P_o via the estimation of (only) a stable factor has been recognized in Hansen and Franklin (1988) and Hansen (1989) as a favourable property. Additionally, the fractional approach to system identification provides a method to deal with closed-loop experiments and to perform an (approximate) identification on the basis of closed-loop data (Hansen *et al.* 1989, Schrama 1991, de Bruyne 1996). Using this property, several approaches to closed-loop identification using fractional model representations have been developed (Zhu and Stoorvogel 1992, Mäkilä and Partington 1992, Lee *et al.* 1993*a*).

4.1.2 The use of fractional model representations in this thesis

It is shown in this chapter that the approximate identification of stable factorizations is readily applicable to data coming from a feedback controlled plant P_o. As a result the closed-loop identification problem, mentioned in the previous chapter, can be solved when the approximate identification is able to estimate a stable factorization of P_o. To anticipate on the results mentioned in this chapter it can be mentioned here that a stable factorization is not unique but it is shown in this chapter that a simple filtering of the closed-loop data can provide access to any stable factorization of P_o. With the opportunity of choosing this filter operation, access to stable factorizations of the plant P_o can be obtained that have favourable properties, such as a minimal McMillan degree or normalized comprime factorization.

Furthermore, for the characterization and estimation of a set of models $\mathcal{P}$, as mentioned in Chapter 2, a set $\mathcal{P}$ is used that is tuned towards the intended robust control design of Procedure 2.5-4. The characterization (2.11) of this set of models $\mathcal{P}$ is postponed until the chapter 5 but it can be mentioned here that the characterization of $\mathcal{P}$ is also based upon a fractional model representation. Consequently, the characterization and estimation of the coefficient matrix Q in Definition 2.2-16 is entirely based on a fractional model approach in this thesis. As the coefficient matrix Q will contain an estimate of a nominal model (or nominal factorization), in this chapter special attention is given to the approximate identification of a nominal model on the

basis of closed-loop data.

Closely related to the topics discussed in this thesis it is worth mentioning that the approximation identification in view of the feedback design using fractional model representations has been studied previously in Schrama (1992b). However, the identification in Schrama (1992b) primarily focuses on the control relevant approximate estimation of nominal models $\hat{P}$ or nominal factorizations $(\hat{N}, \hat{D})$. This thesis aims at estimating sets of models $\mathcal{P}$ tuned towards a robust control design application, as mentioned in Procedure 2.5-4, instead of estimating a nominal model only.

4.2 Factorizations, Coprimeness and Stability

4.2.1 Coprime factorizations

Following Vidyasagar (1985), coprime factorizations are defined using the algebraic theory based on rings. Using this ring theory, a principal ring or principal ideal domain is used to built up the algebraic structure and opens the possibility to study various (multivariable) dynamical systems in a rigorous mathematical way. This study might include topological aspects of (feedback) systems (Vidyasagar et $al.$ 1982), fractional approaches to feedback system design (Desoer et $al.$ 1980) or stability analysis of feedback systems described by fractional representations (Desoer and Gündes 1988).

The primary motivation to introduce this algebraic framework in this thesis, is to study the properties of FDLTI dynamical systems. For that purpose it satisfies to associate the principal ring with $R\mathcal{H}_\infty$, being the set of all stable proper real-rational systems with bounded $\mathcal{H}_\infty$ norm. In the sequel, the association of $R\mathcal{H}_\infty$ with the principal ring is used without mentioning. Accordingly, the following definition of a coprime factorization (over $R\mathcal{H}_\infty$) will be used.

Definition 4.2-1 *Let $N, D \in R\mathcal{H}_\infty$ then the pair (N, D) is a right coprime factorization (rcf) if there exist $X, Y \in R\mathcal{H}_\infty$ such that*

$$XN + YD = I \tag{4.1}$$

Let $\tilde{N}, \tilde{D} \in R\mathcal{H}_\infty$ then the pair $(\tilde{D}, \tilde{N})$ is a left coprime factorization (lcf) if there exist $\tilde{X}, \tilde{Y} \in R\mathcal{H}_\infty$ such that

$$\tilde{N}\tilde{X} + \tilde{D}\tilde{Y} = I \tag{4.2}$$

The difference between a *rcf* and a *lcf* is due to the non-commutative property of multivariable systems. To create a notational distinguishability between a *rcf* and a *lcf*, a *lcf* is equipped with a ˜, while the ordering of the pair $(\tilde{D}, \tilde{N})$ is reversed, compared to the pair (N, D).

Both (4.1) and (4.2) are called Bezout identities. According to Definition 4.2-1, existence of the Bezout identities (4.1) and (4.2) implies the coprimeness of respectively the pair (N, D) and $(\tilde{D}, \tilde{N})$. The Bezout identities also imply the fact that a

coprime factorization cannot have a common unstable zero. This can be verified as follows. In case of a *rcf*, such a common unstable zero must be cancelled by a common unstable pole in X and Y in order to be able to make the right hand side of (4.1) equal to identity. However, both X and Y are restricted to be an element of $R\mathcal{H}_\infty$, hence the Bezout identity cannot be satisfied.

With the notion of coprimeness given in Definition 4.2-1, a *rcf* or *lcf* of a system P can be defined as follows.

Definition 4.2-2 *Let (N, D) be rcf and $(\tilde{D}, \tilde{N})$ be a lcf. Then the pair (N, D) is a rcfof a system P if*

- $\det\{D\} \not\equiv 0$
- $P = ND^{-1}$

Similarly, the pair $(\tilde{D}, \tilde{N})$ is a lcfof a system P if

- $\det\{\tilde{D}\} \not\equiv 0$
- $P = \tilde{D}^{-1}\tilde{N}$

The requirement on the determinant of D and $\tilde{D}$ is needed to ensure that D^{-1} and $\tilde{D}^{-1}$ are well-defined, real rational transfer functions. Due to the multiplication of N and D^{-1} in case of a *rcf* and $\tilde{N}$ and $\tilde{D}^{-1}$ in case of a *lcf*, a coprime factorization of a system P is not unique.

Corollary 4.2-3 *Let (N, D) and $(\tilde{D}, \tilde{N})$ be respectively a rcf and a lcf of a system P. Consider a Q and a $\tilde{Q}$ for which $\det\{Q\} \not\equiv 0$ and $\det\{\tilde{Q}\} \not\equiv 0$. Then (NQ, DQ) is a rcf of P if and only if $Q, Q^{-1} \in R\mathcal{H}_\infty$. Similarly, $(\tilde{Q}\tilde{D}, \tilde{Q}\tilde{N})$ is a lcf of P, if and only if $\tilde{Q}, \tilde{Q}^{-1} \in R\mathcal{H}_\infty$.*

Proof: A proof can be found in Vidyasagar (1985). $\square$

As Corollary 4.2-3 indicates, a coprime factorization is not unique and any non-proper coprime factorization can be modified into a (strictly) proper coprime factorization. Furthermore, for any (strictly) proper transfer function P there always exists a *rcf*, or a *lcf*, that consists of (strictly) proper N, $\tilde{N}$ and proper D, $\tilde{D}$. This result is summarized in the following corollary.

Corollary 4.2-4 *Consider a (strictly) proper real-rational system P, then there exist rcf's (N, D) and lcf's $(\tilde{D}, \tilde{N})$ of P that satisfy*

$$\begin{bmatrix} Y & X \\ -\tilde{N} & \tilde{D} \end{bmatrix} \begin{bmatrix} D & -\tilde{X} \\ N & \tilde{Y} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.3}$$

and for which N, $\tilde{N}$ are (strictly) proper, D, $\tilde{D}$ are proper and D^{-1}, $\tilde{D}^{-1}$ are well defined proper transfer functions.

Proof: A proof can be found in Nett *et al.* (1984) where the explicit state space formulae of the double coprime factorization (4.3) can be found. $\qquad\square$

Uniqueness of a coprime factorization can be addressed by considering a special kind of coprime factorizations that are said to be normalized. In the sequel, the notion of a normalized coprime factorization will be used frequently.

Definition 4.2-5 *A rcf (N_n, D_n) is called a normalized right coprime factorization (nrcf) if the rcf satisfies*

$$N_n^* N_n + D_n^* D_n = I \tag{4.4}$$

Similarly, a lcf $(\tilde{D}_n, \tilde{N}_n)$ is called a normalized left coprime factorization (nlcf) if the lcf satisfies

$$\tilde{D}_n \tilde{D}_n^* + \tilde{N}_n \tilde{N}_n^* = I \tag{4.5}$$

Restricting a coprime factorization to be normalized, restricts the freedom in constructing the coprime factorization. According to Corollary 4.2-3, the freedom in a *rcf* is given by a stable and stably invertible parameter Q. In case of a *nrcf*, the freedom is narrowed. This can be seen by substituting $N_n = NQ$ and $D_n = DQ$ in (4.4), where (N, D) is also a *nrcf*. This implies $Q^*Q = I$ and hence Q is restricted to be a unimodular matrix.

As mentioned before, state space formulae for coprime factors can be found in Nett *et al.* (1984). State space solutions to the computation of normalized coprime factors can for example be found in Meyer and Franklin (1987), Vidyasagar (1988) or Bongers and Bosgra (1993) for continuous-time and in Meyer (1990) or Bongers and Heuberger (1990) for discrete-time systems.

4.2.2 Youla parametrization

With the algebraic framework based on coprime factorizations, an alternative formulation of internal stability of a feedback connection can be given. In Lemma 2.2-10 the internal stability of a feedback connection $\mathcal{T}(P, C)$ is stated in terms of the transfer function matrix $T(P, C)$ given in (2.5). Formally, to check internal stability, all transfer functions in $T(P, C)$ must be evaluated. As mentioned in Corollary 2.2-11 and Corollary 2.2-12, special situations in which either the controller C or the system P is stable require only the evaluation of a single transfer function.

On the basis of the algebraic framework, testing internal stability of $\mathcal{T}(P, C)$ can be brought down to an evaluation of a single transfer function, regardless of P or C being stable. The result is summarized in the following lemma.

Lemma 4.2-6 *Let $P = ND^{-1} = \tilde{D}^{-1}\tilde{N}$ where (N, D) is a rcf and $(\tilde{D}, \tilde{N})$ a lcf of P. Let $C = N_c D_c^{-1} = \tilde{D}_c^{-1}\tilde{N}_c$ where (N_c, D_c) is a rcf and $(\tilde{D}_c, \tilde{N}_c)$ a lcf of C. The following statements are equivalent*

    i. the feedback connection $\mathcal{T}(P,C)$ given in Figure 2.2 is internally stable

    ii. $T(P,C) \in R\mathcal{H}_\infty$

    iii. $\Lambda^{-1} \in R\mathcal{H}_\infty$, with $\Lambda := \begin{bmatrix} \tilde{D}_c & \tilde{N}_c \end{bmatrix} \begin{bmatrix} D \\ N \end{bmatrix}$

    iv. $\tilde{\Lambda}^{-1} \in R\mathcal{H}_\infty$, with $\tilde{\Lambda} := \begin{bmatrix} \tilde{D} & \tilde{N} \end{bmatrix} \begin{bmatrix} D_c \\ N_c \end{bmatrix}$

Proof: The equivalence of item *i* and *ii* was already proven for Lemma 2.2-10. A proof of the equivalence of *iii* or *iv* can be found in Schrama (1992*b*) or Bongers (1994) and is based on the fact that $T(P,C)$ in (2.5) can be written as

$$T(P,C) = \begin{bmatrix} N \\ D \end{bmatrix} \Lambda^{-1} \begin{bmatrix} \tilde{N}_c & \tilde{D}_c \end{bmatrix} \tag{4.6}$$

where $\Lambda = \tilde{D}_c D + \tilde{N}_c N$. □

    From (4.6) it can be observed that the map from $col(r_2, r_1)$ to $col(y, u)$ given by $T(P,C)$, is a series connection of the *lcf* $(\tilde{D}_c, \tilde{N}_c) \in R\mathcal{H}_\infty$ of the controller, the (possibly unstable) map Λ^{-1} and the *rcf* $(N,D) \in R\mathcal{H}_\infty$ of the system P. Provided that $\Lambda^{-1} \in R\mathcal{H}_\infty$, the matrix $T(P,C)$ remains stable. A similar argumentation holds in case $\tilde{\Lambda}$ is being considered.

    For a given *rcf* (N,D) of a system P, a *lcf* $(\tilde{D}_c, \tilde{N}_c)$ of a controller C that satisfies $\Lambda^{-1} \in R\mathcal{H}_\infty$ can readily be found from (4.1). Setting $(\tilde{D}_c, \tilde{N}_c) = (Y, X)$ makes $(\tilde{D}_c, \tilde{N}_c)$ a *lcf*. As in this case $\Lambda = I$, $\Lambda^{-1} \in R\mathcal{H}_\infty$, making $C = Y^{-1}X$ an internally stabilizing controller. Clearly, $C = Y^{-1}X$ is not the only possible stabilizing controller for the feedback connection $\mathcal{T}(P,C)$. Varying the *lcf* $(\tilde{D}_c, \tilde{N}_c)$ such that Λ^{-1} remains stable, characterizes all controllers $C = \tilde{D}_c^{-1}\tilde{N}_c$ that yield a stable feedback connection $\mathcal{T}(P,C)$. Such a characterization of all stabilizing controllers is known as the Youla parametrization (Youla *et al.* 1976*a*, Vidyasagar 1985).

    Clearly, by interchanging the role of P and C, a characterization of all systems $P = ND^{-1}$ that yield an internally stable feedback connection $\mathcal{T}(P,C)$ can also be given. Since such a characterization is dual to the well-known Youla parametrization, it is labelled as a dual-Youla parametrization and reads as follows.

Lemma 4.2-7 *Let (N_x, D_x) be a rcf of an arbitrary auxiliary model $P_x = N_x D_x^{-1}$ and (D_c, N_c) be a rcf of a controller $C = N_c D_c^{-1}$ such that $T(P_x, C) \in R\mathcal{H}_\infty$, then a system P with a rcf (N,D) satisfies $T(P,C) \in R\mathcal{H}_\infty$ if and only if $\exists\, R \in R\mathcal{H}_\infty$ such that*

$$\begin{aligned} N &= N_x + D_c R \\ D &= D_x - N_c R \end{aligned} \tag{4.7}$$

Proof: For a proof one is referred to Schrama (1992*b*). □

Consequently, allowing R in (4.7) to vary freely over all possible transfer functions in $R\mathcal{H}_\infty$ such that $\det\{D_x - N_c R\} \not\equiv 0$, characterizes a set of LTI systems P, written in terms of a *rcf*, that are internally stabilized by the controller C. Although Lemma 4.2-7 is based on a *rcf* of the system, a similar result can be written down for a *lcf* of P, see e.g. Lee *et al.* (1992), Lee *et al.* (1993*a*) or de Bruyne (1996).

4.3 Fractional Approach to Closed-Loop Identification

4.3.1 Algebraic framework in system identification

The algebraic framework of coprime factorizations opens the possibility to characterize a possibly unstable plant P_o by a stable factorization. From an identification point of view, this means that an unstable plant P_o, operating in a stabilized feedback connection $\mathcal{T}(P_o, C)$, can be estimated by simply estimating the stable transfer functions associated to a coprime factorization of the plant.

Two different approaches can be distinguished. Firstly, the identification can be directed towards the estimation of a *rcf* of the (unknown) plant P_o (Schrama and Bosgra 1993, Zhu and Stoorvogel 1992, de Callafon and Van den Hof 1997). Secondly, with explicit knowledge of the controller C, the dual-Youla parametrization can be used to parametrize a *rcf* of the plant and to estimate the stable transfer function R in (4.7) (Hansen 1989, Lee *et al.* 1993*a*, de Bruyne 1996, Lee *et al.* 1995). In both cases, the identification involves the estimation of stable transfer functions to find models of the (possibly unstable) plant P_o.

Both approaches are discussed below. As mentioned in Corollary 4.2-3, a *rcf* of the pant P_o is not unique and the question arises how to access different *rcf*'s of the plant. Furthermore, the problem of identifying a factorization of the plant P_o on the basis of closed-loop data is illuminated. As such, the question whether or not a fractional approach to closed-loop identification is able to address the closed-loop identification problem mentioned in Definition 3.2-9 is elaborated in more detail.

4.3.2 Estimation using dual-Youla parametrization

In the feedback connection $\mathcal{T}(P_o, C)$ of (unknown) plant P_o and (possibly known) controller C, the knowledge or information of the feedback controller C is useful in analyzing the feedback connection $\mathcal{T}(P_o, C)$. This knowledge might include the fact that the controller C forms a stable feedback connection $\mathcal{T}(P_o, C)$. If, in addition, full knowledge of the controller C is available, the fact that $\mathcal{T}(P_o, C)$ is a stable feedback connection, opens the possibility to characterize a *rcf* (N_o, D_o) of the plant P_o with a dual-Youla parametrization.

The use of the dual-Youla parametrization

As the feedback connection $\mathcal{T}(P_o, C)$ is internally stable, Lemma 4.2-7 states that there exists a $R_o \in R\mathcal{H}_\infty$ that can characterize the *rcf* (N_o, D_o) of the plant P_o as follows

$$N_o = N_x + D_c R_o \tag{4.8}$$
$$D_o = D_x - N_c R_o \tag{4.9}$$

In the above equations, (N_x, D_x) is a *rcf* of any auxiliary model that satisfies $T(P_x, C) \in R\mathcal{H}_\infty$ and (N_c, D_c) is a *rcf* of the controller C. The transfer function R_o in (4.8) and (4.9) is an unknown, but stable, transfer function.

Henceforth, estimation of a model $\hat{R}$ of the stable transfer function R_o would yield an estimate $(\hat{N}, \hat{D})$ of a *rcf* of the plant P_o. According to (4.8) and (4.9), a *rcf* $(\hat{N}, \hat{D})$ of a model $\hat{P} = \hat{N}\hat{D}^{-1}$ to be estimated can be represented by

$$\begin{aligned} \hat{N} &= N_x + D_c \hat{R} \\ \hat{D} &= D_x - N_c \hat{R} \end{aligned} \tag{4.10}$$

where $\hat{R} \in R\mathcal{H}_\infty$ indicates a model of the stable transfer function R_o. Representing a *rcf* of a model $\hat{P}$ according to (4.10) has a favourable property.

Remark 4.3-1 *If the estimate $\hat{R}$ is a stable mapping, the resulting model $\hat{P} = \hat{N}\hat{D}^{-1}$ computed from the rcf's given in (4.10) is guaranteed to be stabilized by the controller C used in the feedback connection $\mathcal{T}(P_o, C)$. This is a direct consequence of the dual-Youla parametrization, as mentioned in Lemma 4.2-7.*

The guarantee to find a model $\hat{P}$ that will satisfy $T(\hat{P}, C) \in R\mathcal{H}_\infty$ provided that the estimated model $\hat{R} \in R\mathcal{H}_\infty$ is advantageous. Since the controller C used in the feedback connection $\mathcal{T}(P_o, C)$ is know to stabilize the unknown plant P_o, it is worthwhile to find a model $\hat{P}$ of the plant P_o that is also stabilized by the same controller C. A model $\hat{P}$ that is not stabilized by the controller C will certainly cast doubt on the ability of the model to describe the dynamical aspects of the plant P_o. Therefore it is worthwhile to parametrize a model $\hat{P}$ with the dual-Youla parametrization to guarantee $T(\hat{P}, C) \in R\mathcal{H}_\infty$.

Access to dual-Youla parameter

Access to and estimation of the stable dual-Youla parameter R_o has been studied for example in Hansen (1989), Lee *et al.* (1993*a*) and Schrama (1991). The main conclusion in these references is the fact that the estimation of the dual-Youla parameter, either in a parametrization using a *lcf* or a *rcf*, is an open-loop identification problem that can be tackled by standard system identification techniques (Ljung 1987).

To clarify the equivalent open-loop identification problem of the dual-Youla parameter R_o in the characterization of the rcf (N_o, D_o) given in (4.8) and (4.9), consider the following result.

Lemma 4.3-2 *Consider the data coming from a plant P_o operating in an internally stable feedback connection $\mathcal{T}(P_o, C)$ be described by (3.2). Let the controller C have a rcf (N_c, D_c) and consider an auxiliary model P_x with a rcf (N_x, D_x) that satisfies $\mathcal{T}(P_x, C) \in R\mathcal{H}_\infty$. If the so-called intermediate signal x is defined by the filter operation*

$$x := (D_x + CN_x)^{-1} \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \tag{4.11}$$

and the so-called dual-Youla signal z is defined by the filter operation

$$z := (D_c + P_x N_c)^{-1} \begin{bmatrix} I & -P_x \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \tag{4.12}$$

then (3.2) can be rewritten into

$$z = R_o x + S_o e \tag{4.13}$$

where x is uncorrelated with e and the transfer functions of R_o and S_o respectively are given by

$$R_o = D_c^{-1}(I + P_o C)^{-1}(P_o - P_x)D_x \tag{4.14}$$
$$S_o = D_c^{-1}(I + P_o C)^{-1}H_o \tag{4.15}$$

Proof: The conditions on the internal stability of $\mathcal{T}(P_o, C)$ and $\mathcal{T}(P_x, C)$ allow the use of the the dual-Youla parametrization given in (4.8) and (4.9). Post-multiplication of (4.9) with $-P_x$ and adding (4.8) yields

$$(D_c + P_o N_c)R_o = (P_o D_x - N_x)$$

and with $P_x = N_x D_x^{-1}$, $C = N_c D_c^{-1}$ this can be rewritten into (4.14). Furthermore, post-multiplication of (4.8) with C and adding (4.9) yields $D_o + CN_o = D_x + CN_x$, making

$$(I + CP_o)^{-1} = D_o(D_x + CN_x)^{-1} \tag{4.16}$$

With $P_o = N_o D_o^{-1}$ and $C = N_c D_c^{-1}$, (4.16) can be used to rewrite (3.4) into

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} N_o \\ D_o \end{bmatrix} (D_x + CN_x)^{-1}r + \begin{bmatrix} I \\ -C \end{bmatrix} S_{out} H_o e \tag{4.17}$$

Defining the intermediate signal x as in (4.11) and using the dual-Youla parametriza-
tion given in (4.8) and (4.9), (4.17) can be rewritten into

$$
\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} N_x \\ D_x \end{bmatrix} x + \begin{bmatrix} D_c \\ -N_c \end{bmatrix} R_o x + \begin{bmatrix} D_c \\ -N_c \end{bmatrix} S_o e \tag{4.18}
$$

where S_o is given in (4.15). With Corollary 3.2-4 it can be verified that the interme-
diate signal $x = (D_x + CN_x)^{-1} r$, and hence x is uncorrelated with e. Subsequently,
computing $y - P_x u$ with $P_x = N_x D_x^{-1}$ yields

$$
y - P_x u = (D_c + P_x N_c) R_o x + (D_c + P_x N_c) S_o e
$$

which reduces to (4.13), in case z is given as in (4.12).   □

In the representation of the closed-loop data in (3.2), the plant P_o and the noise
filter H_o are unknown and need to be identified. However, rewriting (3.2) into the
short hand notation of (4.13) indicates that R_o and S_o replace the missing knowledge
of plant P_o and the noise filter H_o by an alternative characterization of P_o and H_o.
This alternative characterization of P_o and H_o in terms of R_o and S_o is given by

$$
\begin{aligned}
P_o &= (N_x + D_c R_o)(D_x - N_c R_o)^{-1} \\
H_o &= D_c (I + P_o C) S_o
\end{aligned} \tag{4.19}
$$

and is found by using (4.8), (4.9) and (4.15). However, opposite to the identification
of P_o and H_o on the basis of closed-loop data, the identification of R_o and S_o is
equivalent to a standard open-loop identification problem and will be discussed in the
following.

Estimation of models in a dual-Youla parametrization

It can be seen from (4.18) that R_o and S_o are the only unknown transfer functions
needed to describe the data $col(y, u)$ coming from the feedback connection $\mathcal{T}(P_o, C)$.
The $rcf\,(N_x, D_x)$ of the auxiliary model P_x and the $rcf\,(N_c, D_c)$ of the controller C in
(4.18) are assumed to be known and given.

It should be pointed out that (4.18) is just an alternative representation of the
closed-loop data as given in (3.2). This representation involves an alternative charac-
terization of the unknown plant P_o and the unknown noise filter H_o with respectively
an unknown dual-Youla parameter R_o and an unknown noise filter S_o.

However, this alternative representation serves an important purpose, as the es-
timation of R_o in (4.14) and S_o in (4.15) can be done by an equivalent open-loop
identification. The open-loop identification is due to the fact that the so-called inter-
mediate signal x mentioned in Lemma 4.3-2, is uncorrelated with the noise e. Hence,
the intermediate signal x of (4.11) and the output signal z in (4.12) can be considered

as respectively an input and (possibly) disturbed output signal, for which the input x is uncorrelated with the disturbance acting on the output signal z. The representation (4.18) can be visualized in the block-diagram given in Figure 4.1, where the equivalent open-loop identification problem has also been indicated.

Fig. 4.1: Block diagram of the reparametrization of plant P_o and noise filter H_o given in (4.18).

Equivalent to (4.19), models $P(\theta)$ and $H(\theta)$ of the plant P_o and the noise filter H_o are represented in a dual-Youla parametrization as

$$
\begin{aligned}
P(\theta) &= (N_x + D_c R(\theta))(D_x - N_c R(\theta))^{-1} \\
H(\theta) &= D_c(I + P(\theta)C)S(\theta)
\end{aligned}
\tag{4.20}
$$

where $R(\theta)$ and $S(\theta)$ are parametrized models of respectively R_o and S_o. As mentioned in Remark 4.3-1, the parametrization (4.20) has a favourable property. However, it can be observed from (4.20) that the McMillan degree of the model $\hat{P} = P(\hat{\theta})$ is determined by the order of the controller C with the rcf (N_c, D_c), the order of the model P_x with the rcf (N_x, D_x) and the McMillan degree of the estimate $\hat{R} = R(\hat{\theta})$ of R_o given in (4.14).

In Schrama (1992b) the alternative characterization of P_o and H_o in (4.19) or the parametrization of the model $\hat{P}$ and $\hat{H}$ in (4.20) is labelled as an R, S-parametrization. The estimation of a model $\hat{R}$ and a noise model $\hat{S}$ can be solved by the standard open-loop identification techniques mentioned in Section 3.2.3.

Relation with the indirect identification method

The reparametrization and the equivalent open-loop identification problem seems to be innovative, but it can be viewed as a generalization of the indirect identification method discussed in Section 3.3.2 as observed in Van den Hof and de Callafon (1996). It can be observed from (4.14) or (4.15) that the transfer functions to be estimated in the equivalent open-loop identification problem are closed-loop transfer functions. Although the closed-loop transfer functions involve additional weightings such as D_c, D_x or P_x, the R, S-parametrization and the equivalent open-loop identification can be made equivalent to the indirect identification problem, in case either the plant P_o or the controller C is known to be stable.

Example 4.3-3 *Consider a feedback connection $\mathcal{T}(P_o, C)$ for which the controller C satisfies $C \in R\mathcal{H}_\infty$. In that case, a rcf (N_c, D_c) of C can be given by $(N_c, D_c) = (C, I)$. An auxiliary model P_x that satisfies $T(P_x, C) \in R\mathcal{H}_\infty$ can be chosen as $P_x = 0$, since the controller C is known to be stable already, and the rcf (N_x, D_x) can be set to $(N_x, D_x) = (0, I)$. As a result, the transfer function for R_o in (4.14) and S_o in (4.15) to be identified are given by*

$$
\begin{aligned}
R_o &= (I + P_oC)^{-1}P_o = P_o(I + CP_o)^{-1} \\
S_o &= (I + P_oC)^{-1}H_o
\end{aligned}
\tag{4.21}
$$

It can be verified that the transfer functions R_o and S_o to be estimated are equivalent to the closed-loop transfer functions of the indirect identification approach discussed in Proposition 3.3-3(b). ◇

In case the controller C is unstable, the dual-Youla parametrization or the R, S-parametrization generalizes the indirect identification method. As mentioned in Remark 4.3-1, in the case where it must be guaranteed beforehand that the model $\hat{P}$ is stabilized by the controller C used in the feedback connection $\mathcal{T}(P_o, C)$, the dual-Youla parametrization can serve this purpose.

4.3.3   Access to factorizations

An alternative to the approach mentioned in the previous section, is the identification of a rcf (N_o, D_o) of the plant P_o directly, without using the dual-Youla parametrization. As mentioned in the previous section, the order of the model $\hat{P}$ is highly influenced by the controller C, the auxiliary model P_x and the model $\hat{R}$ used in the

parametrization (4.20). To eliminate this effect, a direct identification of the *rcf* of the plant is a reasonable alternative. It is shown below that in this way the order of the resulting model can be controlled more efficiently, as no controller or auxiliary model dependent parametrization is used to characterize the *rcf* being estimated.

Access to a stable factorization

To illustrate the idea of accessing and estimating a coprime factorization of an unknown plant P_o, consider the reduced form of the closed-loop data generating system (3.2) as given in (3.4).

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} P_o S_{in} \\ S_{in} \end{bmatrix} r + \begin{bmatrix} S_{out} H_o \\ -C S_{out} H_o \end{bmatrix} e \tag{4.22}$$

Since the controller C is used for the closed-loop experiments, the feedback connection $\mathcal{T}(P_o, C)$ is again assumed to be internally stable and Lemma 2.2-10 yields $T(P_o, C) \in R\mathcal{H}_\infty$ making both $P_o S_{in}, S_{in} \in R\mathcal{H}_\infty$.

Hence, $P_o S_{in}, S_{in}$ can be considered to be a stable, right, but not necessarily coprime, factorization (N_o, D_o) of the plant P_o, with $N_o := P_o S_{in}$ and $D_o := S_{in}$. This stable factorization can accessed easily by considering the map from r in (3.3) or (3.6) to $col(y, u)$. On the coprimeness of the pair $(P_o S_{in}, S_{in})$ the following result can be obtained.

Corollary 4.3-4 *Let a system P and a controller C create an internally stable feedback connection $\mathcal{T}(P, C)$ then (PS_{in}, S_{in}) is a rcf of P if and only if $C \in R\mathcal{H}_\infty$.*

Proof: Since $\mathcal{T}(P, C)$ is internally stable, $T(P, C) \in R\mathcal{H}_\infty$ and implies $(PS_{in}, S_{in}) \in R\mathcal{H}_\infty$ and it remains to show that $\exists\, X, Y \in R\mathcal{H}_\infty$ such that $XPS_{in} + YS_{in} = I$ for coprimeness of (PS_{in}, S_{in}).
$\Leftarrow$ If $C \in R\mathcal{H}_\infty$, taking $X = C \in R\mathcal{H}_\infty$ and $Y = I \in R\mathcal{H}_\infty$ immediately shows coprimeness of (PS_{in}, S_{in}).
$\Rightarrow$ Prove by contradiction: consider C to be unstable and assume (PS_{in}, S_{in}) to be a *rcf* of P. Now let (N, D) be any *rcf* of the system P then $XPS_{in} + YS_{in} = I$ can be rewritten into $XN[D + CN]^{-1} + YD[D + CN]^{-1} = I$ which equals

$$XN + YD = [D + CN] \tag{4.23}$$

Using the fact that $N, D \in R\mathcal{H}_\infty$ and $X, Y \in R\mathcal{H}_\infty$, the left hand side in the equality of (4.23) is stable, while the right hand side is unstable, since C is unstable and N and D are coprime. This contradicts the assumption on the existence of $X, Y \in R\mathcal{H}_\infty$ in $XPS_{in} + YS_{in} = I$. Hence (PS_{in}, S_{in}) is not coprime in case the controller C is unstable. $\square$

Hence, a factorization $(P_o S_{in}, S_{in})$ of the plant P_o is not a *rcf* in case the controller C is unstable. Limiting to the case of stable controllers is restrictive, as many controllers are equipped with an integral action for tracking purposes. Additionally to the lack of coprimeness of the factorization, the signal r given in (3.3) or (3.6) with $r_2 \neq 0$ yields an unbounded signal for an unstable controller C. Furthermore, a factorization in general is not unique and access to factorizations different from $(P_o S_{in}, S_{in})$ is desirable to exploit the freedom in choosing a factorization.

Access to coprime factorizations

In order to deal with the problems mentioned above an additional filtering of the signal r can be proposed. Defining this filtered signal as $x := Fr$, similar as in Van den Hof *et al.* (1995) or de Callafon and Van den Hof (1995*b*), with (3.6) the following relations are obtained

$$x = F \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} r_2 \\ r_1 \end{bmatrix} = F \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \tag{4.24}$$

and (4.22) now rewrites into

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} P_o S_{in} F^{-1} \\ S_{in} F^{-1} \end{bmatrix} x + \begin{bmatrix} S_{out} H_o \\ -C S_{out} H_o \end{bmatrix} e \tag{4.25}$$

where $(P_o S_{in} F^{-1}, S_{in} F^{-1})$ can be considered to be a (right) factorization of the plant P_o. The choice for the filter F determines whether or not $(P_o S_{in} F^{-1}, S_{in} F^{-1})$ is a *rcf*. Furthermore, the filter F can be used to access all possible *rcf*'s of the plant P_o. This result has been summarized in the following lemma.

Lemma 4.3-5 *Let a system P and a controller C with a lcf $(\tilde{D}_c, \tilde{N}_c)$ form an internally stable feedback connection $\mathcal{T}(P, C)$. Then the following statements are equivalent and moreover, each of these statements imply $F[C \ I] \in R\mathcal{H}_\infty$.*

    i. $(P S_{in} F^{-1}, S_{in} F^{-1})$ is a rcf.

    ii. $F = W \tilde{D}_c$ with W, $W^{-1} \in R\mathcal{H}_\infty$

Proof:
$i \Rightarrow ii$ $(P S_{in} F^{-1}, S_{in} F^{-1})$ is a *rcf* and hence $\exists \ \bar{X}, \bar{Y} \in R\mathcal{H}_\infty$ such that

$$\bar{X} P S_{in} F^{-1} + \bar{Y} S_{in} F^{-1} = I. \tag{4.26}$$

Taking any *rcf* (N, D) of P and the *lcf* $(\tilde{D}_c, \tilde{N}_c)$ of C, (4.26) can be rewritten into

$$[\bar{X} N \Lambda^{-1} + \bar{Y} D \Lambda^{-1}] \tilde{D}_c F^{-1} = I \tag{4.27}$$

where $\Lambda = [\tilde{D}_c D + \tilde{N}_c N] \in R\mathcal{H}_\infty$ according to Lemma 4.2-6. Postmultiplication of (4.27) with $F\tilde{D}_c^{-1}$ yields

$$W := F\tilde{D}_c^{-1} = [\bar{X}N\Lambda^{-1} + \bar{Y}D\Lambda^{-1}] \in R\mathcal{H}_\infty \tag{4.28}$$

Since $(PS_{in}F^{-1}, S_{in}F^{-1})$ is a *rcf*, also

$$\begin{bmatrix} PS_{in} \\ S_{in} \end{bmatrix} F^{-1} \in R\mathcal{H}_\infty \tag{4.29}$$

holds. With the definition of $F = W\tilde{D}_c$ from (4.28) and $X, Y \in R\mathcal{H}_\infty$ from $XN + YD = I$, (4.29) can be rewritten into

$$\begin{bmatrix} N \\ D \end{bmatrix} \Lambda^{-1}W^{-1} \in R\mathcal{H}_\infty \tag{4.30}$$

and premultiplication of (4.30) with $\Lambda[X\ Y] \in R\mathcal{H}_\infty$ yields $W^{-1} \in R\mathcal{H}_\infty$.

$ii \Rightarrow i$ With any *rcf* (N, D) of P, the *lcf* $(\tilde{D}_c, \tilde{N}_c)$ of C and $F = W\tilde{D}_c^{-1}$, where W, $W^{-1} \in R\mathcal{H}_\infty$, the factorization $(PS_{in}F^{-1}, S_{in}F^{-1})$ can be rewritten into

$$\begin{bmatrix} PS_{in} \\ S_{in} \end{bmatrix} F^{-1} = \begin{bmatrix} N \\ D \end{bmatrix} \Lambda^{-1}W^{-1} \in R\mathcal{H}_\infty$$

and hence both factors are stable. Consider $X, Y \in R\mathcal{H}_\infty$ with

$$XN + YD = I \tag{4.31}$$

where (N, D) is any *rcf* of P. Since Λ, Λ^{-1}, W, $W^{-1} \in R\mathcal{H}_\infty$, premultiplication of (4.31) with $W\Lambda$ and postmultiplication with $\Lambda^{-1}W^{-1}$ yields

$$\bar{X}N\Lambda^{-1}W^{-1} + \bar{Y}D\Lambda^{-1}W^{-1} = I \tag{4.32}$$

with $\bar{X} := W\Lambda X \in R\mathcal{H}_\infty$, $\bar{Y} := W\Lambda Y \in R\mathcal{H}_\infty$. However, (4.32) equals

$$\bar{X}PS_{in}F^{-1} + \bar{Y}S_{in}F^{-1} = I$$

which proves coprimeness of $(PS_{in}F^{-1}, S_{in}F^{-1})$.

Both conditions are equivalent and it suffices to show that one of the conditions implies $F[C\ I] \in R\mathcal{H}_\infty$. Consider $\tilde{X}, \tilde{Y} \in R\mathcal{H}_\infty$ from $\tilde{N}_c\tilde{X} + \tilde{D}_c\tilde{Y} = I$ where $(\tilde{D}_c, \tilde{N}_c)$ is the *lcf* of C. Then

$$F \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix} = F\tilde{D}_c^{-1} \begin{bmatrix} \tilde{N}_c & \tilde{D}_c \end{bmatrix} \begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix} = F\tilde{D}_c^{-1} = W \in R\mathcal{H}_\infty$$

hence $F[C\ I] \in R\mathcal{H}_\infty$. $\qquad \square$

A similar result can be found in Van den Hof *et al.* (1995), but there the freedom in choosing the filter F is found by restricting both the factorization $(PS_{in}F^{-1}, S_{in}F^{-1})$ and the map $F[C\ I]$ in (4.24) to be stable. However, stability of the map $F[C\ I]$ is not necessary in general. In the case that $r_2(t) = 0\ \forall t$, $x = Fr_1$ and hence only stability of F is required. By restricting $(PS_{in}F^{-1}, S_{in}F^{-1})$ to be a *rcf*, stability of $F[C\ I]$ is implied directly, as mentioned in Lemma 4.3-5.

Lemma 4.3-5 is a generalisation of Corollary 4.3-4 and characterizes the freedom in choosing the filter F by the choice of any stable and stably invertible transfer function W. Moreover, with the additional freedom in W it is possible to have access to all possible right coprime factorizations of a plant P_o. To indicate that the *rcf* of the plant P_o depends on the filter F used in Lemma 4.3-5, the following remark is given.

Remark 4.3-6 *The rcf of the plant P_o that can be accessed by considering the map from x to $col(y, u)$ depends on the choice of the filter F used in (4.24). This dependency is denoted by adopting the notation $(N_{o,F}, D_{o,F})$ to indicate the particular rcf of the plant P_o given by*

$$\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} = \begin{bmatrix} P_o S_{in} F^{-1} \\ S_{in} F^{-1} \end{bmatrix}.$$

A stable and stably invertible transfer function W in Lemma 4.3-5 can be constructed easily. However, from Lemma 4.3-5 it is not clear how the choice of a stable and stably invertible W will change the *rcf* $(P_o S_{in} F^{-1}, S_{in} F^{-1})$ of the plant P_o. To illuminate this effect, the freedom in choosing W can be related to the choice of a *rcf* of a so-called auxiliary model P_x and a *lcf* of an auxiliary controller C_x. In this perspective, the auxiliary model P_x and controller C_x may represent respectively knowledge of the plant P_o and the controller C used in the feedback connection $\mathcal{T}(P_o, C)$.

Proposition 4.3-7 *Let (N_x, D_x) be any rcf of any auxiliary model P_x and $(\tilde{D}_{c,x}, \tilde{N}_c)$ be any lcf of any auxiliary controller C_x then*

$$W := [\tilde{D}_{c,x} D_x + \tilde{N}_{c,x} N_x]^{-1}$$

satisfies W, $W^{-1} \in R\mathcal{H}_\infty$ if and only if $T(P_x, C_x) \in R\mathcal{H}_\infty$.

Proof: Follows directly from Lemma 4.2-6 with $\Lambda = [\tilde{D}_{c,x} D_x + \tilde{N}_{c,x} N_x]$.   □

With Proposition 4.3-7 the freedom of choosing a stable and stably invertible filter W is replaced by the choice of a *rcf* (N_x, D_x) of an auxiliary model P_x and a *lcf* $(\tilde{D}_{c,x}, \tilde{N}_{c,x})$ of an auxiliary controller C_x. It should be noted that the specific choice of W in Proposition 4.3-7 is not a restriction on the set of all possible stable and stably invertible filters W. In fact, it can be shown that for any given controller C, the

freedom in choosing a stable and stably invertible filter W can be fully characterized by the freedom in choosing the *rcf* (N_x, D_x) of the auxiliary model P_x. This result has been stated in the following proposition.

Proposition 4.3-8 *Let W be a filter that satisfies W, $W^{-1} \in R\mathcal{H}_\infty$ and consider a controller C with a lcf $(\tilde{N}_c, \tilde{D}_c)$. Then there exists an auxiliary model P_x with a rcf (N_x, D_x) such that $W = [\tilde{D}_c D_x + \tilde{N}_c N_x]^{-1}$ and $T(P_x, C) \in R\mathcal{H}_\infty$.*

Proof: As $(\tilde{N}_c, \tilde{D}_c)$ is a *lcf* of C, there exist $\tilde{X}$, $\tilde{Y} \in R\mathcal{H}_\infty$ such that

$$\tilde{N}_c \tilde{X} + \tilde{D}_c \tilde{Y} = I. \tag{4.33}$$

Postmultiplying (4.33) with W^{-1} and taking the inverse yields

$$[\tilde{N}_c \tilde{X} W^{-1} + \tilde{D}_c \tilde{Y} W^{-1}]^{-1} = W \in R\mathcal{H}_\infty.$$

With $(\tilde{X}, \tilde{Y})$ satisfying (4.33) and W, $W^{-1} \in R\mathcal{H}_\infty$, it can be verified that $(N_x, D_x) := (\tilde{X} W^{-1}, \tilde{Y} W^{-1})$ is a *rcf* of a model P_x that satisfies $W = [\tilde{D}_c D_x + \tilde{N}_c N_x]^{-1}$. Furthermore, $T(P_x, C) \in R\mathcal{H}_\infty$ from Lemma 4.2-6. $\qquad\square$

As mentioned above C_x can be any (auxiliary) controller. In case C_x is used to represent the knowledge of the controller C used in the feedback connection $\mathcal{T}(P_o, C)$, the filter F in Lemma 4.3-5 can be characterized as follows.

$$F = W\tilde{D}_c = [D_x + CN_x]^{-1} = D_x^{-1}[I + CP_x]^{-1} \tag{4.34}$$

Hence, the role of the *lcf* of the controller C_x drops out in (4.34). This gives rise to an alternative characterization of the filter F to create the signal x in (4.24). This result is summarized in the following corollary.

Corollary 4.3-9 *Let a plant P_o and a controller C create an internally stable feedback connection $\mathcal{T}(P_o, C)$ and let (N_x, D_x) be any rcf of any auxiliary model P_x, then*

$$F = [D_x + CN_x]^{-1}$$

satisfies the conditions of Lemma 4.3-5 if and only if $T(P_x, C) \in R\mathcal{H}_\infty$.

Proof: Follows by application of Proposition 4.3-7 with $C_x = C$ and (4.34). $\qquad\square$

With F given by Corollary 4.3-9, the *rcf* $(N_{o,F}, D_{o,F})$ is related to the *rcf* (N_x, D_x) of the auxiliary model P_x used in the construction of the filter F. This result is summarized in the following corollary.

Corollary 4.3-10 *The rcf $(N_{o,F}, D_{o,F})$ of the plant P_o that can be accessed on the basis of the map from x to $col(y,u)$ in (4.25) satisfies*

$$\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} = \begin{bmatrix} P_o \\ I \end{bmatrix} S_{in} F^{-1} = \begin{bmatrix} P_o \\ I \end{bmatrix} [I + CP_o]^{-1} [I + CP_x] D_x \qquad (4.35)$$

and

$$[D_{o,F} + CN_{o,F}] = F^{-1} = [D_x + CN_x] \qquad (4.36)$$

Proof: Equation (4.35) is found directly with $N_{o,F} = P_o S_{in} F^{-1}$ and $D_{o,F} = S_{in} F^{-1}$ and substituting (4.34). Subsequently, $[D_o + CN_o] = [I + CP_o]S_{in}F^{-1} = F^{-1}$, proving equation (4.36). □

From an identification point of view $[D_{o,F} + CN_{o,F}]$ is unknown, since it contains the *rcf* $(N_{o,F}, D_{o,F})$ of the unknown plant P_o, but (4.36) indicates that this can be replaced by the filter operation F^{-1}, which is completely known. From Corollary 4.3-10 it can also be seen that $(N_{o,F}, D_{o,F})$ can be of high order and containing redundant dynamics. A sensible choice of the auxiliary model P_x may lead to cancelling of redundant dynamics (Van den Hof *et al.* 1995).

4.3.4   Identification of fractional representations

It can be noted that the construction of the signal x in (4.24) with the filter F given in Corollary 4.3-9 is similar to the intermediate signal x as used in (4.11). The filtering (4.24) is visualized in Figure 4.2 and it can be seen that the intermediate signal x is found by a simple filtering of the signal $col(y,u)$.
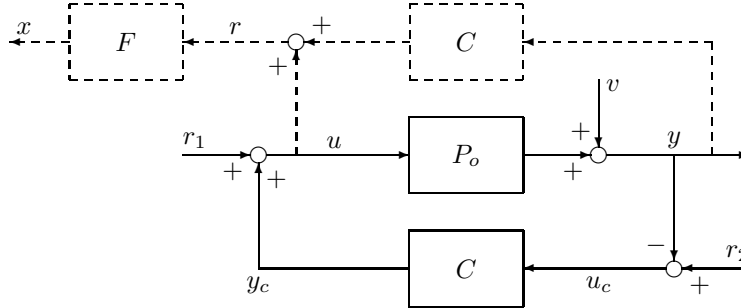
Fig. 4.2: Construction of signal x in (4.24) and admittance to *rcf* $(N_{o,F}, D_{o,F})$ of plant P_o operating in a feedback connection $\mathcal{T}(P_o, C)$.

However, in the delineation presented here, it is indicated that it is possible to access and estimate any *rcf* of the plant P_o operating in closed-loop conditions by simply filtering the signals present during the closed-loop experiments.

Equivalent open-loop identification of coprime factors

As the signal x is uncorrelated with the noise present on either the plant input signal u or the plant output signal y, estimation of the rcf $(N_{o,F}, D_{o,F})$ of the plant P_o is again on open-loop based identification problem. With the result of Lemma 4.3-5 the following proposition for the identification of a right coprime factor can be given, see also Van den Hof *et al.* (1995) or de Callafon and Van den Hof (1995b).

Proposition 4.3-11 *Let the plant P_o and a known controller C create a stable feedback system $\mathcal{T}(P_o, C)$, then the closed-loop data $col(y, u)$ in (4.22) can be rewritten into*

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} x + \begin{bmatrix} D_c \\ -N_c \end{bmatrix} S_o e$$

where S_o is given in (4.15), x given in (4.24) is uncorrelated with e, F is any filter satisfying Lemma 4.3-5 and $(N_{o,F}, D_{o,F})$ is a rcf of the plant P_o that satisfies (4.35) and (4.36).

Proof: By use of (4.25) with $N_{o,F} := P_o S_{in} F^{-1}$ and $D_{o,F} := S_{in} F^{-1}$ and direct application of Corollary 4.3-9.   □

The signal x is uncorrelated with the noise v acting on the feedback connection $\mathcal{T}(Po, C)$. Consequently the fractional approach to system identification yields an equivalent open-loop identification of the plant's coprime factorization $(N_{o,F}, D_{o,F})$ on the basis of closed-loop data. A similar observation was made in (Schrama 1992b) on the basis of the intermediate signal x in (4.11). Here it can be seen that different intermediate signals can be used to access any rcf $(N_{o,F}, D_{o,F})$ of the plant P_o by modifying the filter F in (4.24) and depicted in Figure 4.2. The (open-loop) map $(N_{o,F}, D_{o,F})$ from the intermediate signal x to the closed-loop signals $col(y, u)$ has been depiced in Figure 4.3.

Compared to the parametrization mentioned in (4.20) estimating the rcf $(N_{o,F}, D_{o,F})$ of the plant P_o yields a model $\hat{P}$ and a noise model $\hat{H}$ that are parametrized as follows.

$$\begin{aligned} P(\theta) &= N(\theta)D(\theta)^{-1} \\ H(\theta) &= D_c(I + P(\theta)C)S(\theta) \end{aligned} \tag{4.37}$$

where $(N(\theta), D(\theta))$ is a parametrized model for the rcf $(N_{o,F}, D_{o,F})$ given in (4.35) and $S(\theta)$ is a parametrized model for S_o given in (4.15). Although the parametrization of $H(\theta)$ is the same compared to (4.20), the factorization $(N(\theta), D(\theta))$ of the model $P(\theta)$ is estimated directly. In this way, the McMillan degree of the parametrized model $P(\theta)$ is influenced only by the order of the factorization $(N(\theta), D(\theta))$ being estimated. As such, the filter F will not affect the complexity of the model $\hat{P} = P(\hat{\theta})$
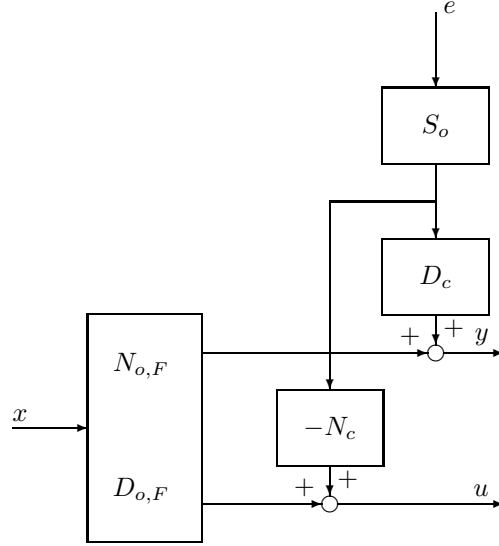
Fig. 4.3: Equivalent open-loop identification of the rcf $(N_{o,F}, D_{o,F})$ of the plant P_o using the signal x of (4.24) and the closed-loop data $col(y, u)$.

being estimated. Opposite to (4.20), the auxiliary model P_x that can be used to construct a filter F, will not be reused in the reconstruction of a model $\hat{P}$.

Remark 4.3-12 *The filter F in (4.24) is only used to construct the signal x for identification purposes. As such, the complexity of F does not affect the complexity of the model $\hat{P}$ in (4.37) being estimated.*

However, the filter F does have an effect on the rcf $(N_{o,F}, D_{o,F})$ that can be accessed on the basis of the (filtered) closed-loop data. It can be observed from (4.35) that choosing a (high order) auxiliary model P_x with a rcf (N_x, D_x) such that $P_x = P_o$ would simplify the rcf $(N_{o,F}, D_{o,F})$ to

$$
\left[\begin{array}{c} N_{o,F} \\ D_{o,F} \end{array} \right] = \left[\begin{array}{c} N_x \\ D_x \end{array} \right]
$$

In that case, the coprime factors of the plant P_o that can be accessed on the basis of closed-loop data are equivalent to the (high order) rcf of P_x. This motivation has been used in de Callafon *et al.* (1994) and Van den Hof *et al.* (1995) to gain access and to estimate a *nrcf* of a plant P_o.

To anticipate on the results mentioned in Chapter 6, it can be mentioned here the order of the model $\hat{P}$ can be made less than or equal to the order of the factorization

$(\hat{N}, \hat{D})$ being parametrized. For that purpose, $N(\theta)$ and $D(\theta)$ should be parametrized in either a state space realization with a common A and B matrix or a Matrix Fraction Description (MFD) with a common right polynomial matrix. More details can be found in Section 6.2.3.

Relation with two-stage identification method

The same approach to use a (filtered) signal in order to obtain an equivalent open-loop identification problem is also being used in the two-stage identification method (Van den Hof and Schrama 1993) as described in Section 3.3.4. In this method the filter F is given by an (accurate) estimate of the input sensitivity function $S_{in} = (I + CP_o)^{-1}$. With this filter F, the signal x has the interpretation of a noise free input signal entering the plant P_o in the closed-loop configuration. As a result, the factorization $(N_{o,F}, D_{o,F})$ to be identified becomes (P_o, I), since $N_{o,F} = P_o S_{in} F^{-1}$, $D_o = S_{in} F^{-1}$. In the two-stage method, an estimate of P_o is found by estimating the coprime factor $N_{o,F}$ only, since $D_{o,F}$ is assumed to be equal to identity.

It should be noted that $F = (I + CP_o)^{-1}$ does not satisfy the conditions mentioned in Lemma 4.3-5 and clearly, the factorization $(N_{o,F}, D_{o,F})$ is not coprime over $R\mathcal{H}_\infty$ for an unstable plant P_o. If the filter F is given by an approximation of the input sensitivity function $(I + CP_o)^{-1}$, where the plant P_o or the controller C is unstable, the situation can become even worse since both $N_{o,F}$ and $D_{o,F}$ can be unstable. This is due to the fact that F^{-1}, which is the inverse of the estimated input sensitivity function, will be unstable and the unstable modes will not be cancelled completely in the operation $P_o S_{in} F^{-1}$ or $S_{in} F^{-1}$. The consequence for the two-stage method is that in the second step an unstable transfer function $P_o S_{in} F^{-1}$ needs to be estimated.

4.4   Closed-loop Identification Problem Revisited

To summarize, the fractional approach to closed-loop identification, either based on the dual-Youla parametrization or the direct identification of a coprime factorization of the plant, basically consists of three steps.

1. Filtering: the construction of an intermediate signal, being a filtered version of the signals measured from the feedback connection $\mathcal{T}(P_o, C)$.

2. Identification: an equivalent open-loop identification of a fractional representation of the plant P_o and/or the estimation of a closed-loop noise model.

3. Reconstruction: computation of model and noise model from the estimates obtained.

The filtering step is needed to construct the signals used for the identification in the second step. In this step, knowledge of the input signal u and the output signal y or

knowledge of r_1 and r_2 is used. In case of the dual-Youla parametrization, knowledge of the controller C is also required to construct the intermediate signal in (4.11), The direct coprime factor identification approach is more robust against incomplete knowledge of the controller C. As indicated in Lemma 4.3-5, the minimum knowledge needed is the location of unstable poles of the controller that should appear in the factor $\tilde{D}_c$ to construct the filter F.

As the identification step is just an equivalent open-loop identification problem, a consistent estimate of either the transfer function R_o in (4.14) of the rcf $(N_{o,F}, D_{o,F})$ in (4.35) and/or the closed-loop noise model S_o can be obtained, provided that the reference signals are persistently exciting (Assumption 3.2-2). Depending on the method used, either from (4.20) or (4.37) a consistent estimate $\hat{P}$ of P_o can be obtained. As such, the closed-loop partial consistent estimation problem mentioned in Definition 3.2-9 can be solved. Furthermore, the equivalent open-loop identification problem also enables the estimation of tunable approximate models of either R_o or $(N_{o,F}, D_{o,F})$ having a limited complexity and independent of the noise v present on the closed-loop data $col(y, u)$.

Due to fractional approach in which only stable transfer functions are being estimated, there is no distinction between the estimation of stable or unstable plants P_o. Furthermore, the controller is also factorized in a stable factorization, so that no distinction have to be made between stable or unstable controllers. In this way the methods are applicable to situations in case both the plant and the controller, used in the feedback connection $\mathcal{T}(P_o, C)$, are unstable.

However, there are differences between the two fractional approaches discussed in this chapter. These differences are due to the different parametrizations of the model $P(\theta)$ used in (4.20) for the dual-Youla approach and (4.37) for the coprime factor approach. As mentioned in Remark 4.3-1, the dual-Youla parametrization enables the property to guarantee that a model $\hat{P}$ being estimated, is stabilized by the controller C used in the feedback connection $\mathcal{T}(P_o, C)$. This property is not shared when a model $\hat{P}$ is estimated and constructed via (4.37).

On the other hand, a serious drawback from the dual-Youla parametrization is the complexity of the model $\hat{P}$ being estimated. As mentioned before, the McMillan degree of the model $\hat{P}$ found by (4.20) is determined by the order of the controller C with the rcf (N_c, D_c), the order of the model P_x with the rcf (N_x, D_x) and the McMillan degree of the estimate $\hat{R}$ of R_o given in (4.14). Hence, performing an approximate identification of R_o to obtain a low complexity model $\hat{R}$ will still lead to a high complexity model $\hat{P}$ in case either the controller C or the auxiliary model P_x being used in (4.20) have a relatively high McMillan degree.

In case of the parametrization given in (4.37), the complexity of a controller C and a filter F[1] will not affect the complexity of the model $\hat{P}$ being estimated. Therefore,

---

[1]The filter F may be based on an auxiliary model P_x, see Corollary 4.3-9.

directly estimating a coprime factorization $(N_{o,F}, D_{o,F})$ is found to be more suitable to find low complexity models $\hat{P}$ to address the closed-loop approximate estimation problem mentioned in Definition 3.2-9.

Therefore, the estimation of a model $\hat{P}$ of low complexity will be based on the coprime factor approach in the remaining part of this thesis. This approach is generally applicable and able to handle the problem of approximate identification of stable and unstable plants P_o on the basis of closed-loop data. However, the favourable properties associated to the dual-Youla parametrization will be elaborated in the construction of a set of model $\mathcal{P}$ for which an estimated nominal factorization $(\hat{N}, \hat{D})$ will serve as a nominal model.

In this chapter the attention is focused solely on the estimation of a nominal model $\hat{P}$ on the basis of closed-loop data using the coprime factor approach. However, the nominal model is just a part of the set of models $\mathcal{P}$ needed in Procedure 2.5-4. The following chapter discusses the construction of the set of models $\mathcal{P}$ and the role of the identification of the set $\mathcal{P}$ within Procedure 2.5-4 in more detail. After that, in chapter 6, the tools presented in this chapter and chapter 5 will be combined to elucidate the identification of a set of models $\mathcal{P}$, including a nominal model $\hat{P}$, on the basis of closed-loop data.

Part III

Procedure

5

Control and Performance Enhancement using a Set of Models

5.1 Retrospect on the Problem Formulation

Looking back on the problem formulation discussed in the first chapter and the approach outlined in Section 2.5.2 of Chapter 2, it has become clear that the knowledge of the (unknown) plant P_o is represented by means of a set of models $\mathcal{P}$. Henceforth, the design of an enhanced performing feedback controller for the plant P_o must be done on the basis of this set $\mathcal{P}$. To be able to complete the design of such an enhanced performing controller C on the basis of a set of model $\mathcal{P}$, the following three main ingredients must be considered.

- Characterization of the set of models $\mathcal{P}$.

  As mentioned in Section 2.5.3, the set of models $\mathcal{P}$ should allow the evaluation of the closed loop performance assessment test (2.19) and the closed loop validation test of (2.20) and (2.21) in a non-conservative way.

- Definition of the performance.

  In Remark 2.2-15, the performance has been characterized by the norm $\|J(P,C)\|_\infty$ of a control objective function $J(P,C) \in R\mathcal{H}_\infty$. A specification of the control objective function is required to perform the steps of Procedure 2.5-4.

- Design of a robust controller on the basis of a set of models.

  Once the structure of the set of models $\mathcal{P}$ and the $\mathcal{H}_\infty$ norm-based control objective has been defined, a robust performing controller, according to step 2 in Procedure 2.5-4, must be computed.

As of now, these items have not been discussed in detail, but do constitute the main ingredients of Procedure 2.5-4. In this chapter, the above mentioned ingredients will

be discussed and the exact formulation and characterization of the set of models $\mathcal{P}$, the performance and the design of a robust controller is illuminated. As such, this chapter presents new results on the characterization of a set of models $\mathcal{P}$ using a dual-Youla perturbation. It is shown how this set of models is constructed and motivated from the perspective of a closed-loop performance evaluation. Furthermore, for this set of models, performance evaluation and robust control design will be discussed. These results are based on the existing theory of μ-analysis and synthesis.

According to Section 2.5.3, there is a fourth main ingredient to be considered: the identification procedure to actually estimate the set of models $\mathcal{P}$. The estimation of a set $\mathcal{P}$ plays an important role in step 1 and step 3 of Procedure 2.5-4. As the identification procedure is subjected to the control objective function $J(P, C)$ and the structure of the set $\mathcal{P}$ being used, a discussion of the estimation of $\mathcal{P}$ is postponed and can be found in the next chapter.

The outline of this chapter is as follows. In light of the favourable properties associated to the identification of fractional model representations mentioned in the previous chapter, the fractional approach will be continued here to define the structure of the set of models $\mathcal{P}$. In Section 5.2 this fractional approach to construct a set of models $\mathcal{P}$ is illuminated and additional benefits are mentioned. Subsequently, in Section 5.3 the form of the $\mathcal{H}_\infty$ norm-based control objective function $J(P, C)$ is given and the way in which the closed loop performance assessment test (2.19) and the closed loop validation test of (2.20) and (2.21) can be evaluated. Finally, in Section 5.4 the procedure to the design of an enhanced and robust performing controller on the basis of the set of models $\mathcal{P}$ is given.

5.2   Characterization of the Set of Models

5.2.1   Motivation for fractional approach

The fractional approach presented in the previous chapter does not distinguish between the estimation of stable or unstable systems. Similarly, the knowledge of a feedback controller C is also represented by a stable factorization and no distinction has to be made between stable and unstable controllers. Finally, it has been illustrated in the previous chapter that the estimation of a *rcf* of the plant P_o on the basis of data obtained under closed-loop conditions does not differ much from a standard open-loop identification problem.

In light of the advantages associated to this fractional model approach, it is a natural consequence to exploit the framework of fractional model representations in the characterization of a set of models $\mathcal{P}$ used to represent the knowledge of the possibly unstable plant P_o. As indicated in Section 2.2.4, the set $\mathcal{P}$ is built up from a nominal model $\hat{P}$ along with an allowable model perturbation Δ. The allowable model perturbation Δ represents the incomplete knowledge of the plant P_o with respect to

the nominal model $\hat{P}$ and allows the construction of a set $\mathcal{P}$ such that $P_o \in \mathcal{P}$.

In this section, some of the possibilities to construct a set of models $\mathcal{P}$ on the basis of perturbations within a fractional model approach are illustrated. In the next section, the construction of a set of models based on additive (or multiplicative) perturbations on coprime factorizations and distance measures based uncertainty sets are presented. In Section 5.2.4, a set of models described by a perturbation in a dual-Youla parametrization will be discussed in more detail. The latter has several favourable properties and motivates to structure the set of models $\mathcal{P}$ accordingly, in order to be used in Procedure 2.5-4 for both system identification and control design purposes.

5.2.2 Perturbations on coprime factors and distance measures

In the fractional approach discussed in the previous chapter, the plant P_o is represented by a rcf $(N_{o,F}, D_{o,F})$ as given in (4.35) and the nominal model $\hat{P}$ is represented by a rcf $(\hat{N}, \hat{D})$. Following Example 2.2-17, a typical set of models $\mathcal{P}$ can be constructed straightforwardly by considering either additive or multiplicative perturbations on the rcf $(\hat{N}, \hat{D})$ of the nominal model $\hat{P}$.

In case of additive perturbations on the rcf $(\hat{N}, \hat{D})$ the following a set of models $\mathcal{P}_A$ can be considered.

$$
\mathcal{P}_A(\hat{N}, \hat{D}, V, W) = \left\{ P \mid P = ND^{-1} \text{ where } \begin{bmatrix} N \\ D \end{bmatrix} = \begin{bmatrix} \hat{N} \\ \hat{D} \end{bmatrix} + \begin{bmatrix} \Delta_N \\ \Delta_D \end{bmatrix} \text{ with} \right.
$$
$$
\left. \begin{bmatrix} \Delta_N \\ \Delta_D \end{bmatrix} \in R\mathcal{H}_\infty \text{ and } \Delta := \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} \Delta_N \\ \Delta_D \end{bmatrix} W \text{ satisfies } \|\Delta\|_\infty < 1 \right\}
\tag{5.1}
$$

It can be observed from (5.1) that a set of models $\mathcal{P}$ is found by considering an additive stable perturbation (Δ_N, Δ_D) on the nominal rcf $(\hat{N}, \hat{D})$ where the perturbation (Δ_D, Δ_N) is assumed to be unknown but bounded. The weighting functions V and W serve as appropriate frequency dependent weighting functions to normalize the $\mathcal{H}_\infty$ norm bound on the allowable model perturbation Δ. In a similar way, a set of models $\mathcal{P}$ based on a (input or output) multiplicative perturbation can be constructed.

In general, a rcf $(\hat{N}, \hat{D})$ of a model $\hat{P}$ is not unique and in (5.1) the rcf $(\hat{N}, \hat{D})$ being chosen is not specified. Furthermore, the size of the (unweighted) perturbation (Δ_N, Δ_D), measured by

$$
\left\| \begin{bmatrix} \Delta_N \\ \Delta_D \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} N \\ D \end{bmatrix} - \begin{bmatrix} \hat{N} \\ \hat{D} \end{bmatrix} \right\|_\infty
$$

depends on the difference between the specific rcf (N, D) of any system $P \in \mathcal{P}$ and the chosen rcf $(\hat{N}, \hat{D})$ of the nominal model $\hat{P}$. Restricting the rcf $(\hat{N}, \hat{D})$ to be normalized and considering the smallest difference between the rcf $(\hat{N}, \hat{D})$ and the rcf

(N, D) can be formalized by the notion of a distance measure, known as the gap- and graph-metric (El-Sakkary 1985, Vidyasagar 1984, Vidyasagar 1985, Zhu 1989).

Definition 5.2-1 *Let the distance between two systems P_1 and P_2 in the graph and gap metric respectively be denoted by $d(P_1, P_2)$ and $\delta(P_1, P_2)$, and let $P_i = N_i D_i^{-1}$ where (N_i, D_i) be a nrcf for $i \in [1, 2]$.*
Then $d(P_1, P_2)$ is defined by

$$d(P_1, P_2) := \max\{\vec{d}(P_1, P_2), \vec{d}(P_2, P_1)\}, \text{ with}$$

$$\vec{d}(P_i, P_j) := \inf_{Q \in R\mathcal{H}_\infty, \, \|Q\|_\infty \leq 1} \left\| \begin{bmatrix} D_i \\ N_i \end{bmatrix} - \begin{bmatrix} D_j \\ N_j \end{bmatrix} Q \right\|_\infty$$

and $\delta(P_1, P_2)$ is defined by

$$\delta(P_1, P_2) := \max\{\vec{\delta}(P_1, P_2), \vec{\delta}(P_2, P_1)\}, \text{ with}$$

$$\vec{\delta}(P_i, P_j) := \inf_{Q \in R\mathcal{H}_\infty} \left\| \begin{bmatrix} D_i \\ N_i \end{bmatrix} - \begin{bmatrix} D_j \\ N_j \end{bmatrix} Q \right\|_\infty$$

The distance $\delta(P_1, P_2) \in [0, 1]$ is called 'the gap' between the systems P_1, P_2. It should be noted that, formally, the gap and graph metric are defined in terms of the graph topology (El-Sakkary 1985, Vidyasagar 1985), whereas the above mentioned definitions are obtained from the result mentioned in Georgiou (1988). Note that the computation of a *nrcf* and the minimization in Definition 5.2-1 are needed to compute the distance between the two systems P_1 and P_2, while the minimization associated to the gap-metric is much easier to perform (Georgiou 1988). However, due to the similarity between the two distance measures, the gap and the graph metric can be related by the following triangular inequality (Packard and Helwig 1989)

$$d(P_1, P_2) \leq \delta(P_1, P_2)$$

making the gap between P_1 and P_2 an upper bound for the graph metric.

On the basis of the distance measures as given in Definition 5.2-1, again a set of models $\mathcal{P}$ can be characterized. Using the gap metric as distance measure between two systems, a set of models $\mathcal{P}_{\vec{\delta}}$ can be formulated as

$$\mathcal{P}_{\vec{\delta}}(\hat{P}, \vec{\delta}_{max}) := \{P \mid \vec{\delta}(\hat{P}, P)\vec{\delta}_{max}^{-1} < 1\} \tag{5.2}$$

where $\vec{\delta}_{max} \in \mathbb{R}$. In Sefton and Ober (1993) and de Callafon *et al.* (1996*b*) the similarities between the set $\mathcal{P}_{\vec{\delta}}$ in (5.2) and the set $\mathcal{P}_A$ in (5.1) have been investigated. In these references it has been shown that both sets can be made identical for specific choices of the *rcf* of the nominal model $\hat{P}$.

5.2.3   Introducing knowledge of the controller

In de Callafon *et al.* (1996*b*) it is observed that either in the set $\mathcal{P}_A$ in (5.1) or $\mathcal{P}_{\vec{\delta}}$ in (5.2), the closed-loop operation of any system $P \in \mathcal{P}$ is not taken into account. Stated differently, knowledge of a controller C is not used in the construction of any of the sets of models as discussed above.

Clearly, in case it is known that the (unknown) plant P_o is stabilized by a certain feedback controller C, this knowledge can and should be exploited to construct a set of models $\mathcal{P}$. During the construction of $\mathcal{P}$ this information can be taken into account. As result, the set of models $\mathcal{P}$ can be tuned towards the available information that is based on the closed-loop operation of the plant P_o. An attempt to exploit the knowledge of a stabilizing controller is the use of a weighted gap or so-called Λ-gap as introduced in Bongers (1991) or Bongers (1994).

Definition 5.2-2 *Let* $P_i = N_i D_i^{-1}$, *where* (N_i, D_i) *is a nrcf for* $i \in [1,2]$. *Let* $C = \tilde{D}_c^{-1} \tilde{N}_c$ *be any controller having a nlcf* $(\tilde{D}_c, \tilde{N}_c)$, *that creates an internally stable feedback connection* $\mathcal{T}(P_1, C)$ *then* $\vec{\delta}_\Lambda(P_1, P_2)$ *is defined as*

$$\vec{\delta}_\Lambda(P_1, P_2) := \inf_{\bar{Q} \in R\mathcal{H}_\infty} \left\| \begin{bmatrix} D_1 \\ N_1 \end{bmatrix} \Lambda^{-1} - \begin{bmatrix} D_2 \\ N_2 \end{bmatrix} \bar{Q} \right\|_\infty$$

with $\Lambda = [\tilde{D}_c D_1 + \tilde{N}_c N_1]$.

The difference between $\vec{\delta}(P_1, P_2)$ and $\vec{\delta}_\Lambda(P_1, P_2)$ is the additional shaping of the nrcf (N_1, D_1) of P_1 with Λ^{-1} into a rcf $(\bar{N}, \bar{D})$ that makes $\bar{\Lambda} := \tilde{D}_c \bar{D} + \tilde{N}_c \bar{N} = I$. As a result, the distance between P_1 and P_2 is measured via a rcf $(\bar{N}, \bar{D})$ of P_1 that takes into account the closed loop operation of P_1. This makes the distance between P_1 and P_2 dependent on (the nrcf of) the controller C. However, the distance measure $\vec{\delta}_\Lambda(P_1, P_2)$ is not a metric since $\vec{\delta}_\Lambda(P_1, P_2) \neq \vec{\delta}_\Lambda(P_2, P_1)$ as to the influence of the controller C that may be different for P_2 and P_1.

Similar to the set $\mathcal{P}_{\vec{\delta}}$ as given in (5.2), a set of models described by the Λ-gap can be defined as follows

$$\mathcal{P}_{\vec{\delta}_\Lambda}(\hat{P}, \vec{\delta}_{\Lambda,max}) := \{P \mid \vec{\delta}_\Lambda(\hat{P}, P) \vec{\delta}_{\Lambda,max}^{-1} < 1\}. \tag{5.3}$$

The tuning of the set of models $\mathcal{P}_{\vec{\delta}_\Lambda}$ so that knowledge of a stabilizing feedback controller C is used, yields a favourable property not shared by other "open-loop" based uncertainty sets as given in (5.1), (5.2) or Example 2.2-17. To clarify this property, consider the following result.

Proposition 5.2-3 *Consider the set* $\mathcal{P}_{\vec{\delta}}$ *given in (5.2) and the set* $\mathcal{P}_{\vec{\delta}_\Lambda}$ *given in (5.3), where* $\vec{\delta}_{max}^{-1} = \vec{\delta}_{\Lambda,max}^{-1}$. *Let* $\mathcal{P}_{stab}$ *denote all systems* P *for which* $\mathcal{T}(P, C)$ *is an*

internally stable feedback connection and C is the feedback controller as used in the construction of the set $\mathcal{P}_{\vec{\delta}_\Lambda}$. If

$$\begin{aligned} \mathcal{P}_a &:= \mathcal{P}_{\vec{\delta}} \cap \mathcal{P}_{stab} \\ \mathcal{P}_b &:= \mathcal{P}_{\vec{\delta}_\Lambda} \cap \mathcal{P}_{stab} \end{aligned}$$

then $\mathcal{P}_a \subset \mathcal{P}_b$.

Proof: For a proof one is referred to de Callafon *et al.* (1996*b*).   □

Clearly, the above mentioned proposition addresses the robust stability problem as given in Definition 2.2-18. In words, Proposition 5.2-3 indicates that the set of models that are stabilized by the controller C and described by the structure given in (5.3) encloses the set of models that are stabilized by the controller C and described by the structure given in (5.2).

As a set of models $\mathcal{P}$ is used to capture the limited knowledge available on the plant P_o, it is obvious to use the knowledge that the plant P_o is stabilized by a controller C. This knowledge is helpful in estimating and constructing a set of models $\mathcal{P}$ for which is known beforehand that all systems $P \in \mathcal{P}$ are stabilized by the controller C. With the definition of $\mathcal{P}_{stab}$ given in Proposition 5.2-3, the knowledge of the controller C must be used in a more elaborated way to guarantee that the set of models $\mathcal{P}_i$ being estimated satisfies $\mathcal{P} = \mathcal{P}_{stab}$. The construction of such a set of models can be found in the next section.

5.2.4   Perturbations in a dual-Youla parametrization

The dual-Youla parametrization of Lemma 4.2-7 parametrizes all systems P that are internally stabilized by a controller C on the basis of an auxiliary model P_x which is chosen to be stabilized by the controller C. This property can be elaborated in structuring and defining a set of models $\mathcal{P}$ in such a way, that all systems $P \in \mathcal{P}$ are guaranteed to be stabilized by the controller C. A set of models $\mathcal{P}$ that satisfies this property is given in the following definition.

Definition 5.2-4 *Let a nominal model $\hat{P}$ with a rcf $(\hat{N}, \hat{D})$ and a controller C with a rcf (N_c, D_c) form an internally stable feedback connection $\mathcal{T}(\hat{P}, C)$. Then the set of models $\mathcal{P}$ is defined by*

$$\begin{aligned} \mathcal{P}(\hat{N}, \hat{D}, N_c, D_c, \hat{V}, \hat{W}) := \{ P \mid P = (\hat{N} + D_c \Delta_R)(\hat{D} - N_c \Delta_R)^{-1} \\ \text{with } \Delta_R \in R\mathcal{H}_\infty \text{ and } \Delta := \hat{V} \Delta_R \hat{W} \text{ satisfies } \|\Delta\|_\infty < \gamma^{-1} \} \end{aligned} \tag{5.4}$$

and $\hat{V}, \hat{W}$ are stable and stably invertible weighting functions.

The set $\mathcal{P}$ essentially depends on the factorization $(\hat{N}, \hat{D})$ of the nominal model $\hat{P}$, the factorization (N_c, D_c) of the controller C and the weighting functions $\hat{V}$, $\hat{W}$. Similar to the set of models defined above, the unknown, but bounded model perturbation Δ takes into account the incomplete knowledge of the plant P_o that should lie in the above mentioned set of models $\mathcal{P}$.

Remark 5.2-5 *Without loss of generality, the bound on the uncertainty in (5.4) can also be normalized by the weighting functions $\hat{V}$ or $\hat{W}$. Hence, the set $\mathcal{P}$ does not essentially depend on the numerical value of γ, but bounding it by γ^{-1} will simplify notation considerably in the sequel. For similar reasons of notational simplicity, it will be assumed that Δ_R in (5.4) is an unstructured and stable LTI operator.*

Clearly, in case Δ_R in (5.4) is a scalar, the perturbation Δ_R can be bounded by a single scalar stable and stably invertible weighting function. To avoid confusion in case Δ_R is multivariable, the pre- and post-multiplication with respectively $\hat{V}$ and $\hat{W}$ is maintained.

Remark 5.2-6 *To keep track of a multivariable perturbation Δ_R, a notation involving the pair $(\hat{V}, \hat{W})$ is used to indicate the bound on Δ_R.*

Referring back to Procedure 2.5-4, the set $\mathcal{P}_i$ used in the identification of step 1 and the control design in step 2 of Procedure 2.5-4 can be characterized by employing the knowledge of the stabilizing controller C_i that is implemented on the actual plant P_o. Employing the characterization given in Definition 5.2-4, the set of models $\mathcal{P}_i$ used in step 1 and step 2 of Procedure 2.5-4 can be specified. Using $(N_{c,i}, D_{c,i})$ to denote the *rcf* of C_i and a nominal model $\hat{P}_i$ with a *rcf* $(\hat{N}_i, \hat{D}_i)$ that satisfies $T(\hat{P}_i, C_i) \in R\mathcal{H}_\infty$, the set $\mathcal{P}_i$ is given by

$$\mathcal{P}_i(\hat{N}_i, \hat{D}_i, N_{c,i}, D_{c,i}, \hat{V}_i, \hat{W}_i) = \{P \mid P = (\hat{N}_i + D_{c,i}\Delta_R)(\hat{D}_i - N_{c,i}\Delta_R)^{-1}$$
$$\text{with } \Delta_R \in R\mathcal{H}_\infty \text{ and } \Delta_i := \hat{V}_i\Delta_R\hat{W}_i \text{ satisfies } \|\Delta_i\|_\infty < \gamma_i^{-1}\} \tag{5.5}$$

for stable and stably invertible weighting functions $\hat{V}_i$ and $\hat{W}_i$.

The arguments of the set $\mathcal{P}$ in (5.4) or the set $\mathcal{P}_i$ will be omitted in the sequel, since the dependency mentioned above is clear from Definition 5.2-4. However, a distinction must be made between the arguments assumed to be known and those who have to be estimated or identified when constructing the set of models $\mathcal{P}$.

Remark 5.2-7 *From an identification point of view, the arguments equipped with a $\hat{\ }$-symbol ($\hat{N}$, $\hat{D}$, $\hat{V}$ and $\hat{W}$) are the arguments to be identified and have to be obtained by employing a system identification technique. In accordance with Remark 2.5-2, the rcf (N_c, D_c) of the controller C, used to construct the set of models, is assumed to be known.*

Due to the close connection with the dual-Youla parametrization, the set of models (or model uncertainty set) $\mathcal{P}$ in (5.4) contains only models that are stabilized by the currently implemented and known controller C, regardless of the value γ. This advantage, observed also by (Schrama 1992b, pp. 139-141) or Sefton *et al.* (1990), is not shared by alternative uncertainty characterizations, such as the open-loop additive uncertainty description given in (5.1) or the uncertainty sets described in (5.2) or (5.3). As a result, the set of models $\mathcal{P}$, as defined in (5.4), covers all systems that are stabilized by the controller C, used to build up the set of models $\mathcal{P}$.

To anticipate on the results presented in the following sections, it can be noted here that the structure of the set of models $\mathcal{P}$ given in (5.4) will yield an affine expression in Δ_R to evaluate the control objective function $J(P, C) \; \forall P \in \mathcal{P}$. This is an additional motivation to structure the set of models as in (5.4), which will be explained in more detail in the next chapter. First, the definition of the set of models $\mathcal{P}$ as given in (5.4) will be rewritten in a standard form based on LFT's as mentioned in Section 2.2.4.

5.2.5   Representations via LFT's

The use of LFT representations enables the possibility to rewrite the set of models $\mathcal{P}$ of Definition 5.2-4 in a standard form that is beneficial for notational and computational purposes. To apply the LFT framework to the set $\mathcal{P}$ given in (5.4), the perturbation on the nominal model $\hat{P}$, or nominal factorization $(\hat{N}, \hat{D})$, is represented by an LFT with a norm bounded uncertainty $\Delta \in R\mathcal{H}_\infty$. This framework has been discussed in Section 2.2.4 and the general form of model perturbation has again been visualized in Figure 5.1.
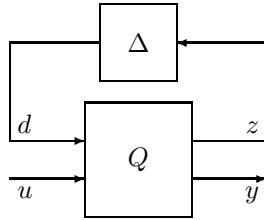
Fig. 5.1: LFT representation of model perturbation.

In Figure 5.1, the signals u and y denote respectively the input and output of any system $P \in \mathcal{P}$, while the uncertainty or allowable perturbation on the nominal model $\hat{P}$ is represented by the mapping Δ between the fictitious signals d and z. In this way, the mapping from u onto y for some $\Delta \in R\mathcal{H}_\infty$ is given by the upper LFT

$$F_u(Q, \Delta) := Q_{22} + Q_{21}\Delta(I - Q_{11}\Delta)^{-1}Q_{12}$$

as previously defined in (2.9).

The entries of the coefficient matrix Q in (2.9) depend on the nominal model $\hat{P}$ and the way in which the allowable model perturbation Δ will affect the nominal model $\hat{P}$. By defining $d = \hat{V}\Delta_R\hat{W}z$, it can be verified that the map form $col(d, u)$ onto $col(z, y)$ for any system $P \in \mathcal{P}$ given in (5.4) can be represented by Figure 5.2.
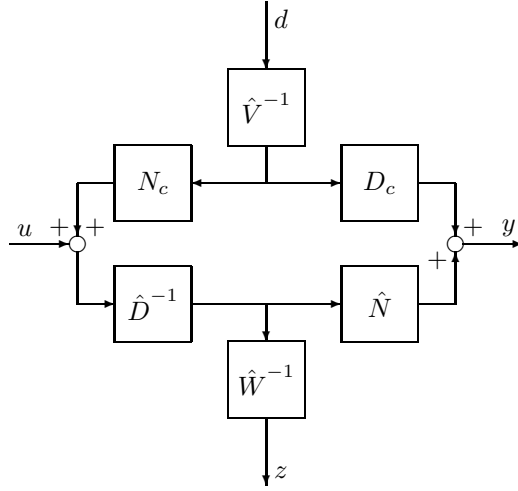
Fig. 5.2: Representation of Q for the set $\mathcal{P}$ given in (5.4).

From Figure 5.2 it can be observed that the nominal map (for $\Delta_R = 0$) equals the nominal model $\hat{P} = \hat{N}\hat{D}^{-1}$. Alternative systems are found by the dual-Youla perturbation Δ_R that relate the fictitious signals d and z via $d = \hat{V}\Delta_R\hat{W}z$ and modifies the nominal map according to the upper LFT given in (2.9). On the basis of Figure 5.2, a characterization of the coefficient matrix Q in (2.9) can be given and an alternative representation of the set of models $\mathcal{P}$ in (5.4) can be obtained.

Corollary 5.2-8 *The set of models $\mathcal{P}$ given in (5.4) can be written as*

$$\mathcal{P} = \{P \mid P = \mathcal{F}_u(Q, \Delta) \ with \ \Delta \in R\mathcal{H}_\infty, \ \|\Delta\|_\infty < \gamma^{-1} \ and$$

$$Q = \left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline Q_{21} & Q_{22} \end{array}\right] = \left[\begin{array}{c|c} \hat{W}^{-1}\hat{D}^{-1}N_c\hat{V}^{-1} & \hat{W}^{-1}\hat{D}^{-1} \\ \hline (D_c + \hat{P}N_c)\hat{V}^{-1} & \hat{P} \end{array}\right] \tag{5.6}$$

Proof: The entries of Q can be found by defining $\Delta = \hat{V}\Delta_R\hat{W}$ and considering the map from $col(d, u)$ onto $col(z, y)$ in Figure 5.2. □

Although the coefficient matrix Q in (5.6) looks complicated, it can be written as

as simple multiplication

$$
Q = \begin{bmatrix} \hat{W}^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix} \begin{bmatrix} \hat{V}^{-1} & 0 \\ 0 & I \end{bmatrix}
$$

where the entries of the (unweighted) coefficient matrix $\bar{Q}$ is given by

$$
\begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix} = \begin{bmatrix} \hat{D}^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ \hat{P} & I \end{bmatrix} \begin{bmatrix} C & I \\ I & 0 \end{bmatrix} \begin{bmatrix} D_c & 0 \\ 0 & I \end{bmatrix}. \tag{5.7}
$$

The multiplication of the transfer functions in (5.7) indicates the construction of the (unweighted) coefficient matrix $\bar{Q}$. Furthermore, it can be observed that $\bar{Q}$ is invertible, as all the matrices in (5.7) are known to be invertible, whereas

$$
\begin{bmatrix} I & 0 \\ \hat{P} & I \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} C & I \\ I & 0 \end{bmatrix}
$$

are unimodular. From the multiplication in (5.7) it can be observed that Q will have a McMillan degree equal to the sum of the McMillan degree of the (nominal) model $\hat{P}$ and the controller C. A state space realization of the coefficient matrix Q, based on the multiplication (5.7) is given in Appendix A.

Although the expressions mentioned in (5.6) and (5.7) are valid for the multivariable case, it is worth mentioning here that in case the (allowable) model perturbation Δ_R in (5.4) is multivariable, it may be beneficial to represent the perturbation Δ in Figure 5.1 in a diagonal form. Especially in the case where stable and stably invertible scalar weighting filters $\hat{V}_{ij}$ are available that upper bound each element (i,j) of Δ_R separately via

$$
\|\hat{V}_{ij}\Delta_{Rij}\|_\infty \le \gamma^{-1}. \tag{5.8}
$$

In this way, the perturbation Δ in Figure 5.1 is given by $\Delta = \text{diag}\{\Delta_{Rij}\hat{V}_{ij}\}$ instead of the matrix multiplication $\hat{V}\Delta_R\hat{W}$ given in (5.4). In that case, the expression for the coefficient matrix Q in (5.6) is slightly modified. To illustrate this modification, consider a 2×2 (unweighted) perturbation Δ that is given by

$$
\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}
$$

which can be linked to a diagonal representation of Δ via

$$
\begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} I & I & 0 & 0 \\ 0 & 0 & I & I \end{bmatrix}}_{T_1} \begin{bmatrix} \Delta_{11} & & & 0 \\ & \Delta_{12} & & \\ & & \Delta_{21} & \\ 0 & & & \Delta_{22} \end{bmatrix} \underbrace{\begin{bmatrix} I & 0 \\ 0 & I \\ I & 0 \\ 0 & I \end{bmatrix}}_{T_2}
$$

Incorporating the abovementioned unitary scaling matrices T_1 and T_2 into the un-weighted coefficient matrix $\bar{Q}$ of (5.7) will modify $\bar{Q}$ into

$$\left[\begin{array}{cc} T_2 & 0 \\ 0 & I \end{array}\right]\left[\begin{array}{cc} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{array}\right]\left[\begin{array}{cc} T_1 & 0 \\ 0 & I \end{array}\right]$$

where the perturbation Δ has the above mentioned diagonal form.

An LFT representation with this modified coefficient matrix can still be used to characterize all models within a set of models $\mathcal{P}$. To avoid the cumbersome notation associated to the diagonal form of the model perturbation Δ, as mentioned in Remark 5.2-6, a notation involving the pair $(\hat{V}, \hat{W})$ is used to indicate the bound on Δ_R.

Finally it can be mentioned that the coefficient matrix Q completely characterizes the set of models in (5.6). Similar to the definition of the set of models $\mathcal{P}$ given in (5.4), the coefficient matrix Q depends on the nominal factorization $(\hat{N}, \hat{D})$ of the nominal model $\hat{P}$, the *rcf* (N_c, D_c) of the controller C and the stable and stably invertible weighting filters $(\hat{V}, \hat{W})$. It should be noted that in the construction of the set of models $\mathcal{P}$ according to Corollary 5.2-8, the condition $T(\hat{P}, C) \in R\mathcal{H}_\infty$ is not required. However, as mentioned in Section 5.2.4, the favourable properties associated with the dual-Youla parametrization are beneficial only if this condition is indeed taken into account. Henceforth, for the construction of the coefficient matrix Q, stable and stably invertible weighting filters $(\hat{V}, \hat{W})$, a nominal factorization $(\hat{N}, \hat{D})$ of a nominal model $\hat{P}$ and a controller C with a *rcf* (N_c, D_c) is needed, where $T(\hat{P}, C) \in R\mathcal{H}_\infty$.

5.3 Performance Characterization

5.3.1 Control objective function

To formalize the performance of a feedback connection of a system P and a controller C, a control objective function $J(P, C) \in R\mathcal{H}_\infty$ has been introduced in Section 2.2.3. As pointed out in this section, the control objective function depends at least on both the controller C and the system P. Although the characterization of performance may involve the specification of additional weighting functions or the use of time domain constraints (Boyd and Barrat 1991), the performance essentially depends on the controller C and the system P that assemble the feedback connection $\mathcal{T}(P, C)$ in Figure 2.2.

Due to the algebraic relations between r_1, u and y_c and between r_2 or v, y and u_c, the various maps present in a feedback connection $\mathcal{T}(P, C)$ can be investigated by the four-block transfer function matrix $T(P, C)$ as defined in (2.5). For reasons of generality it is a natural consequence to choose the control objective function $J(P, C)$ equal to some input/output weighted form of the transfer function matrix $T(P, C)$.

In that case, the performance (level) $\|T(P,C)\|_\infty$ is measured by

$$\|J(P,C)\|_\infty := \|U_2 T(P,C) U_1\|_\infty \tag{5.9}$$

where U_1 and U_2 are (square) weighting functions (not necessarily stable or stably invertible). The weighting functions U_1 and U_2 can be used for additional shaping of the $T(P,C)$ matrix.

Although it is impossible to transform any desirable control design objective into the single norm function $\|U_2 T(P,C) U_1\|_\infty$, the performance characterization (5.9) has wide applicability. It may include a weighted sensitivity or mixed sensitivity characterization by proper modification of the weighting functions U_1 and U_1 (Schrama 1992b). To investigate for example (internal) stability of the feedback connection, it is necessary to study the map from $col(r_2, r_1)$ to $col(u_c, u)$ (Chen and Desoer 1982). On the other hand, disturbance attenuation can be reflected by the map from v to $col(y, u)$ and tracking by the map from r_2 onto $col(y, u)$ (Boyd and Barrat 1991). Therefore, the performance measure mentioned in (5.9) is believed to be fairly general and suitable for most applications.

Considering the notion of performance (robustness) mentioned in Definition 2.2-19 with a $\gamma > 0$ and the set of models $\mathcal{P}$ as given in (5.4), then a controller C is said to satisfy nominal performance if

$$\|J(\hat{P}, C)\|_\infty = \|U_2 T(\hat{P}, C) U_1\|_\infty \leq \gamma \tag{5.10}$$

whereas robust performance is satisfied when

$$\|J(P, C)\|_\infty = \|U_2 T(P, C) U_1\|_\infty \leq \gamma \tag{5.11}$$

for all $P \in \mathcal{P}$.

In light of Procedure 2.5-4, the above mentioned tests have to be performed for some controller C_i or C_{i+1} using an identified set of models $\mathcal{P}_i$ and built around a nominal model $\hat{P}_i$.

As mentioned in Remark 2.5-3, the weighting functions U_1 and U_2 are assumed to be given and fixed in order to compare the performance when updating or redesigning a controller C_i to C_{i+1}. In this way, the numerical value of $J(P, C_i)$ and $J(P, C_{i+1})$ for different systems P that belong to a set of models can be compared.

As mentioned before, this does not imply that the weighting functions U_2 and U_1 in the control objective function are not allowed to be changed. Clearly, the shape and size of the weighting filters U_1 and U_2 will highly depend on required control effort and the dynamics of the plant P_o. Such information becomes available once an identification and controller design (e.g. Procedure 2.5-4) has been performed. Therefore, it is realistic to assume that the weighting filters U_2 and U_1 will also be changed. However, for analysis purposes, the weighting filters U_2 and U_1 are assumed to be fixed.

Finally it can be noted that applying either the controller C_i or C_{i+1} in Procedure 2.5-4 to a set of models $\mathcal{P}_i$ as given in (5.5) requires the evaluation of a robust performance test similar to (5.11). With the LFT framework in Section 5.2.5 to represent the set of models, a performance robustness test can be formulated straightforwardly. This is discussed in the subsequent sections.

5.3.2 LFT representation

Due to the model-based approach, the analysis of the performance when applying a controller to a set of models and the synthesis of a robust controller for a set of models can be handled relatively easily. This can be done by evaluating the worst case performance, or similarly, checking robust performance of a controller when applying it to all the models within the set of models.

In order to be able to check performance robustness, the performance of a controller C applied to any model $P \in \mathcal{P}$ of (5.4) is written in terms of an LFT. In this way, standard results present in literature Zhou *et al.* (1996) can be used to evaluate performance robustness. For a clear understanding of the results, a distinction must be made between the controller applied to the set of models and the controller used in the construction of the set of models $\mathcal{P}_i$ as in (5.5).

Remark 5.3-1 *The controller used in the construction of the set of models $\mathcal{P}_i$ is denoted by C_i, while the controller applied to the set of models $\mathcal{P}_i$ for performance (robustness) analysis purposes is denoted by C.*

It should be noted that the indexing used in Remark 5.3-1 is consistent with the indexing used in Procedure 2.5-4. Hence, C_{i+1} will denote the controller on which the set of models $\mathcal{P}_{i+1}$ is based. With the use of this notation, the following result can be obtained.

Lemma 5.3-2 *Consider the set $\mathcal{P}_i$ defined in (5.5) and a controller C such that the map $J(P,C) = U_2 T(P,C) U_1$ is well-posed for all $P \in \mathcal{P}_i$. Then*

$$\mathcal{P}_i = \{ P \mid J(P,C) = \mathcal{F}_u(M,\Delta) \ with \ \Delta \in R\mathcal{H}_\infty, \ \|\Delta\|_\infty < \gamma_i^{-1} \}$$

where the entries of M are given by

$$
\begin{aligned}
M_{11} &= -\hat{W}_i^{-1}(\hat{D}_i + C\hat{N}_i)^{-1}(C - C_i)D_{c,i}\hat{V}_i^{-1} \\
M_{12} &= \hat{W}_i^{-1}(\hat{D}_i + C\hat{N}_i)^{-1} \begin{bmatrix} C & I \end{bmatrix} U_1 \\
M_{21} &= -U_2 \begin{bmatrix} -I \\ C \end{bmatrix} (I + \hat{P}_i C)^{-1}(I + \hat{P}_i C_i)D_{c,i}\hat{V}_i^{-1} \\
M_{22} &= U_2 \begin{bmatrix} \hat{N}_i \\ \hat{D}_i \end{bmatrix} (\hat{D}_i + C\hat{N}_i)^{-1} \begin{bmatrix} C & I \end{bmatrix} U_1
\end{aligned}
\tag{5.12}
$$

Proof: Consider a LFT representation of $\mathcal{P}_i$ similar to (5.6). Create the feedback connection of Q depicted in Figure 5.1 with a controller C, where $u := r_1 + C(r_2 - y)$ and define signals $col(w_1, w_2)$ and $col(d_1, d_2)$ such that $col(r_2, r_1) = U_1 col(w_1, w_2)$ and $col(e_1, e_2) = U_2 col(y, u)$. Then it can be verified that the map from $col(d, w_1, w_2)$ onto $col(z, e_1, e_2)$ is given by the transfer function M of (5.12). In here the relations

$$\begin{bmatrix} -I & Q_{22} \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} -(I + Q_{22}C)^{-1} & (I + Q_{22}C)^{-1}Q_{22} \\ (I + CQ_{22})^{-1}C & (I + CQ_{22})^{-1} \end{bmatrix}$$

and

$$\begin{bmatrix} -I & Q_{22} \\ C & I \end{bmatrix}^{-1} + \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} Q_{22} \\ I \end{bmatrix} (I + CQ_{22})^{-1} \begin{bmatrix} C & I \end{bmatrix}$$

are used to derive the expressions given in (5.12). Subsequently, the map from $col(w_1, w_2)$ onto $col(e_1, e_2)$ equals the upper LFT $\mathcal{F}_u(M, \Delta)$. □

With Lemma 5.3-2, the performance $\|J(P, C)\|_\infty$ of a controller C applied to any models $P \in \mathcal{P}_i$ can be evaluated by

$$\|\mathcal{F}_u(M, \Delta)\|_\infty = \|M_{22} + M_{21}\Delta(I - M_{11}\Delta)^{-1}M_{12}\|_\infty \tag{5.13}$$

for all $\Delta \in R\mathcal{H}_\infty$ with $\|\Delta\|_\infty \leq \gamma_i^{-1}$. Note that the entries of the transfer function M in (5.12) are determined solely by the controller C, the structure and the variables used to represent the set $\mathcal{P}_i$ of (5.5) and the weightings U_2, U_1 of the performance specification (5.9). As such, a distinction must be made between the controller C_i and its rcf $(N_{c,i}, D_{c,i})$, used in the construction of the set of models $\mathcal{P}_i$, and the controller C applied to the set $\mathcal{P}_i$. As a special entry of M, one can recognise M_{11} as the lower LFT $\mathcal{F}_l(Q, -C)$, whereas M_{22} equals the transfer function $U_2 T(\hat{P}_i, C)U_1$ and is associated to the nominal performance specification.

As indicated in Remark 5.3-1 the feedback controller C in (5.12) denotes the controller C applied to the set of models $\mathcal{P}_i$ for analysis purposes. Hence, substituting $C = C_i$ can be used for the performance assessment in step 1 *a posteriori*, while setting $C = C_{i+1}$ can be employed to check and guarantee performance of C_{i+1} *a priori* in step 2 of Procedure 2.5-4. Similar results can also be obtained for step 3 in Procedure 2.5-4.

As mentioned before, evaluating (5.13) can be done by applying standard results available in the literature (Packard and Doyle 1993, Zhou *et al.* 1996). However, in order to be able to compute the (worst case) performance for the LFT given in (5.13) in a non-conservative way, the concept of μ or structured singular value (Packard and Doyle 1993) is needed. This will be summarized in the next section.

5.3.3 Structured singular value

The structured singular value is a matrix function, denoted by $\mu(M)$, where M can be any (square) complex matrix. It plays a crucial role in the evaluation of performance robustness (Doyle *et al.* 1992), which is the main reason to use it in this thesis.

The definition of $\mu(\cdot)$ depends on an underlying (diagonal) structure (Doyle *et al.* 1992, Zhou *et al.* 1996). This structure, which will be denoted by $\boldsymbol{\Delta}$, is determined by the structure of the uncertainty set and the performance objective function being used. The structured singular value $\mu(\cdot)$ with respect to such a structure $\boldsymbol{\Delta}$ will be denoted by $\mu_{\boldsymbol{\Delta}}(\cdot)$. Using the symbol $\bar{\sigma}(\Delta)$ to denote the maximum singular value of Δ, the definition of $\mu_{\boldsymbol{\Delta}}(\cdot)$ adopted from Doyle *et al.* (1991) reads as follows.

Definition 5.3-3 *For a complex matrix M, the structured singular value $\mu_{\boldsymbol{\Delta}}(M)$ is defined by*

$$
\mu_{\boldsymbol{\Delta}}(M) := \begin{cases} \dfrac{1}{\min\limits_{\Delta \in \boldsymbol{\Delta}}\{\bar{\sigma}(\Delta)\}} & \text{if } \exists \Delta \in \boldsymbol{\Delta} \ \text{s.t. } \det(I - M\Delta) = 0 \\[2ex] 0 & \text{if } \nexists \Delta \in \boldsymbol{\Delta} \ \text{s.t. } \det(I - M\Delta) = 0 \end{cases}
$$

In this thesis, the structure $\boldsymbol{\Delta}$ used in Definition 5.3-3 is restricted to have a diagonal form, having two unstructured uncertainty blocks Δ_1 and Δ_2 only. Now let M be partitioned as

$$
M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{5.14}
$$

then the blocks Δ_1 and Δ_2 are compatible in size with M_{11} and M_{22}, meaning that both $M_{11}\Delta_1$ and $M_{22}\Delta_2$ are square. In this way the structure of $\boldsymbol{\Delta}$ is given by

$$
\boldsymbol{\Delta} := \left\{ \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} \mid \Delta_1, \Delta_2 \in R\mathcal{H}_\infty, \ \|\Delta_1\|_\infty < 1, \|\Delta_2\|_\infty < 1 \right\}. \tag{5.15}
$$

In general $\mu_{\boldsymbol{\Delta}}(M)$ is approximated by computing upper and lower bounds. The upper bound is derived by the computation of non-negative scaling matrices D_l and D_r defined within a set $\mathcal{D}$ that commutes with the structure $\boldsymbol{\Delta}$. One is referred to e.g. (Packard and Doyle 1993) for a detailed discussion on the specification of such a set $\mathcal{D}$ of scaling matrices. Basically, the commutation of $\mathcal{D}$ with $\boldsymbol{\Delta}$ implies that for all $D_l, D_r \in \mathcal{D}$ and for all $\Delta \in \boldsymbol{\Delta}$, $D_r\Delta = \Delta D_l$ and $\mu_{\boldsymbol{\Delta}}(M) = \mu_{\boldsymbol{\Delta}}(D_l M D_r^{-1})$. This gives rise to the computation of the following upper bound.

$$
\mu_{\boldsymbol{\Delta}}(M) \leq \inf_{D_l, D_r \in \mathcal{D}} \bar{\sigma}(D_l M D_r^{-1}) \tag{5.16}
$$

The infimization formulated in (5.16) can be reformulated as a convex optimization problem (Packard and Doyle 1993). However, for the special cases of M and $\boldsymbol{\Delta}$ used in this paper, it is possible to compute $\mu_{\boldsymbol{\Delta}}(M)$ exactly.

Lemma 5.3-4 *Consider the structure* $\mathbf{\Delta}$ *given in (5.15) and* $\mu_{\mathbf{\Delta}}(M)$ *given in Definition 5.3-3, then*

$$\mu_{\mathbf{\Delta}}(M) = \inf_{D_l, D_r \in \mathcal{D}} \bar{\sigma}(D_l M D_r^{-1})$$

Proof: The structure $\mathbf{\Delta}$ consists of two full blocks. Application of theorem 9.1 in Packard and Doyle (1993) or theorem 11.5 in Zhou *et al.* (1996) yields the result. $\square$

The characterization of M in (5.12) will play a crucial role in evaluating robustness issues. With the above mentioned results, stability and performance robustness for a controller C applied to a set of models $\mathcal{P}_i$ of (5.5) can be evaluated.

5.3.4   Evaluating performance for the set of models

The properties of $\mu_{\mathbf{\Delta}}(M)$ as given in Definition 5.3-3 and the result mentioned in Lemma 5.3-4 can now be used to study the upper LFT $\mathcal{F}_u(M, \Delta)$ of (5.13). In this way, both stability and performance robustness can be evaluated by using standard results that are available in the literature (Packard and Doyle 1993, Zhou *et al.* 1996). First, the result on stability robustness is summarized. From this result, the favourable properties associated to a set of models structured as in Definition 5.2-4 will be made clear. Subsequently, performance robustness and the link with Procedure 2.5-4 is mentioned.

Stability robustness

Using the upper LFT $\mathcal{F}_u(M, \Delta)$ given in (5.13), result on stability robustness can be formulated. In this section these results are applied to the set of models $\mathcal{P}_i$ as given in (5.5). First the general result on stability robustness of $\mathcal{F}_u(M, \Delta)$ will be discussed.

Lemma 5.3-5 *Let the stable transfer functions* $M, \Delta \in R\mathcal{H}_\infty$ *construct a basic perturbation model* $\mathcal{F}_u(M, \Delta)$ *and assume that for all* $\Delta \in R\mathcal{H}_\infty$ *with* $\|\Delta\|_\infty < 1$ *the transfer function* $M_{21}\Delta(I - M_{11}\Delta)^{-1}M_{12}$ *does not exhibit unstable pole/zero cancellations*[1]. *Then* $\mathcal{F}_u(M, \Delta)$ *is well-posed and BIBO stable for all* $\Delta \in R\mathcal{H}_\infty$ *with* $\|\Delta\|_\infty < 1$, *if and only if*

$$\|M_{11}\|_\infty \leq 1 \tag{5.17}$$

Proof: Since $M \in R\mathcal{H}_\infty$, and thus $M_{11}, M_{12}, M_{21}, M_{22} \in R\mathcal{H}_\infty$, the small gain theorem (Zames 1963) directly leads to the sufficient condition of the stability of $\mathcal{F}_u(M, \Delta)$. Provided that unstable poles of $(I - M_{11}\Delta)^{-1}$ are not cancelled in $\mathcal{F}_u(M, \Delta)$ for all $\Delta \in R\mathcal{H}_\infty$ with $\|\Delta\|_\infty < 1$, this leads to the necessary condition of (5.17). For a complete proof see Glover and McFarlane (1989) or Zhou *et al.* (1996). $\square$

---

[1]This additional condition, which is often discarded in literature, excludes trivial situations as e.g. $M_{21} = 0$ or $M_{12} = 0$.

With the result of Lemma 5.3-5 and the characterization of M in (5.12), the notion of robust stability according to Definition 2.2-18 can now be formulated for a set of models $\mathcal{P}_i$.

Corollary 5.3-6 *Consider the set $\mathcal{P}_i$ given in (5.5) and a controller C such that the feedback connection $\mathcal{T}(\hat{P}_i, C)$ is well-posed, internally stable and satisfies $T(\hat{P}_i, C) \in RH_\infty$. Then the feedback connection $\mathcal{T}(P, C)$ is well-posed and internally stable for all $P \in \mathcal{P}_i$, if and only if*

$$\left\| \hat{W}_i^{-1}(\hat{D}_i + C\hat{N}_i)^{-1}(C_i - C)D_{c,i}\hat{V}_i^{-1} \right\|_\infty \leq \gamma_i \qquad (5.18)$$

Proof: It can be verified from (5.12), that any internal stabilizing controller C that satisfies $T(\hat{P}_i, C) \in RH_\infty$, will yield stable entries of M, where the weighting functions U_2 and U_1 in (5.12) can be set to identity. Furthermore, it can be verified from (5.12) that for stable and stably invertible filters $\hat{V}_i$ and $\hat{W}_i$, $M_{21}\Delta(I - M_{11}\Delta)^{-1}M_{12}$ cannot exhibit pole/zero cancellations for any stable Δ. Lemma 5.3-5 can be applied by extracting M_{11} from (5.12) and taking into account the scaling γ_i^{-1} in (5.5). For all $P \in \mathcal{P}_i$ this yields the necessary and sufficient condition (5.18). $\qquad\square$

Similar results have also been be found by Schrama (1992*b*, pp. 139). However, in Schrama (1992*b*) only sufficiency of (5.18) was mentioned. Furthermore, a similar result is also obtained when considering a perturbation of the controller C, where the perturbation is described in a dual-Youla parametrization (Schrama 1992*b*). As mentioned before, due to the structure of the set of models $\mathcal{P}_i$ in (5.5), all systems $P \in \mathcal{P}_i$ of (5.5) are stabilized by the controller C_i. This is consistent with the stability robustness result mentioned in Corollary 5.3-6. Checking stability robustness for $C = C_i$ applied to the set $\mathcal{P}_i$ of (5.5) will result in $M_{11} = 0$ and hence (5.18) is satisfied trivially, regardless of the value of γ_i.

Remark 5.3-7 *Independent of the size of the allowable model perturbation Δ_R, stability robustness is satisfied when the controller C_i is applied to the set of models $\mathcal{P}_i$ given in (5.5). However, application of a (newly designed) controller C_{i+1}, that not has been used to construct to the set of models $\mathcal{P}_i$, does require the evaluation of stability robustness by evaluating (5.18) for $C = C_{i+1}$.*

Although the possibility to check stability robustness is helpful before implementing a controller to the plant P_o, satisfying performance robustness is more essential and a stronger requirement than stability robustness. For that purpose, the evaluation of the control objective function for each system $P \in \mathcal{P}_i$ is required.

Performance robustness

Performance robustness is, in general, a much stronger requirement than stability robustness (Doyle *et al.* 1992). In the above mentioned results it was made clear that

the size of set $\mathcal{P}_i$ (measured by the value of γ_i) is not a limiting factor when evaluating stability robustness of the controller C_i applied to $\mathcal{P}_i$. Performance robustness, however, may pose additional requirements on the size of the set of models $\mathcal{P}_i$. To clarify this statement, first the general result on performance robustness of $\mathcal{F}_u(M, \Delta)$ will be discussed.

Lemma 5.3-8 *Consider stable transfer functions $M, \Delta \in R\mathcal{H}_\infty$ where M is partitioned as in (5.14) and $\mu_{\boldsymbol{\Delta}}(M)$ is defined related to the structure $\boldsymbol{\Delta}$ given in (5.15). Then $\mathcal{F}_u(M, \Delta)$ is well-posed, BIBO stable and $\|\mathcal{F}_u(M, \Delta)\|_\infty \leq \gamma$ for all $\Delta \in R\mathcal{H}_\infty$ with $\|\Delta\|_\infty < \gamma^{-1}$, if and only if*

$$\mu_{\boldsymbol{\Delta}}(M) \leq \gamma \qquad (5.19)$$

Proof: By setting $\Delta = \Delta_1$ and adding a fictitious full block uncertainty $\Delta_2 \in R\mathcal{H}_\infty$ with $\|\Delta_2\| < \gamma^{-1}$, the uncertainty structure (5.15) is obtained. Application of the main loop theorem, similar as in theorem 11.7 in Zhou *et al.* (1996) now proves the result. □

The result of Lemma 5.3-8 opens the possibility to evaluate the performance robustness of a controller C applied to a set of models in a non-conservative way. This set of models can be either $\mathcal{P}_i$ as used in step 1 and step 2 of Procedure 2.5-4, or a newly identified set of models $\mathcal{P}_{i+1}$ as used in step 3. The result for evaluating the performance of a controller C applied to the $\mathcal{P}_i$ of (5.5) is stated in the following theorem. Similar results can be derived for $\mathcal{P}_{i+1}$.

Corollary 5.3-9 *Consider the set $\mathcal{P}_i$ defined in (5.5) and a controller C such that $\mathcal{T}(\hat{P}_i, C)$ is well-posed, internally stable and satisfies $U_2 T(\hat{P}_i, C) U_1 \in R\mathcal{H}_\infty$. Then, for all $P \in \mathcal{P}_i$, the feedback system $\mathcal{T}(P, C)$ is well-posed, internally stable and satisfies $\|U_2 T(P, C) U_1\|_\infty \leq \gamma_i$ if and only if*

$$\mu_{\boldsymbol{\Delta}} \left(\begin{bmatrix} \hat{W}_i^{-1} & 0 \\ 0 & U_2 \end{bmatrix} T_{ext}(\hat{P}_i, C_i, C) \begin{bmatrix} -\hat{V}_i^{-1} & 0 \\ 0 & U_1 \end{bmatrix} \right) \leq \gamma_i \qquad (5.20)$$

where $T_{ext}(\hat{P}_i, C_i, C)$ is given by

$$\left[\begin{array}{c|c} Z_i(C - C_i)D_{c,i} & Z_i \begin{bmatrix} C & I \end{bmatrix} \\ \hline \left(\begin{bmatrix} \hat{N}_i \\ \hat{D}_i \end{bmatrix} Z_i C + \begin{bmatrix} I \\ 0 \end{bmatrix} \right) (D_{c,i} + \hat{P}_i N_{c,i}) & \begin{bmatrix} \hat{N}_i \\ \hat{D}_i \end{bmatrix} Z_i \begin{bmatrix} C & I \end{bmatrix} \end{array} \right] \qquad (5.21)$$

where $Z_i = (\hat{D}_i + C\hat{N}_i)^{-1} = \hat{D}_i^{-1}(I + C\hat{P}_i)^{-1}$.

Proof: Lemma 5.3-2 connects $\mathcal{F}_u(M,\Delta)$ with $U_2 T(P,C) U_1$ for all $P \in \mathcal{P}_i$. The expression for $T_{ext}(\hat{P}_i, C_i, C)$ can be found by use of (5.12) and algebraic manipulation. Applying Lemma 5.3-8 yields the necessary and sufficient condition for $\|\mathcal{F}_u(M,\Delta)\|_\infty \leq \gamma_i$ to hold for all $P \in \mathcal{P}_i$. $\qquad\qquad\square$

The analysis result presented in Corollary 5.3-9 can be used in the various steps mentioned in Procedure 2.5-4. Structuring the set of models $\mathcal{P}_i$ as in (5.5), the performance assessment test and the controller and modelling validation tests mentioned in Remark 2.5-6 can be performed with the result mentioned in Corollary 5.3-9.

Remark 5.3-10 *Referring to Procedure 2.5-4, substituting $C = C_i$ in (5.20) can be used for the* a posteriori *performance assessment in step 1. On the other hand, substitution of $C = C_{i+1}$ in (5.20) can be used to check and guarantee the* a priori *performance robustness of C_{i+1} in step 2.*

Recall from Lemma 5.3-4 that for structure $\mathbf{\Delta}^2$ the value of (5.20) can be computed exactly. Similar results can be derived also for the set of models $\mathcal{P}_{i+1}$ as used in step 3 of Procedure 2.5-4.

As already mentioned in the discussion on stability robustness, the substitution of $C = C_i$ in (5.12) yields $M_{11} = 0$ and implies stability robustness for the controller C_i appied to the set of models $\mathcal{P}_i$ regardless of the value of γ_i. However, a requirement on performance robustness limits the allowable value of value of γ_i. For $C = C_i$ the upper LFT $\mathcal{F}_u(M,\Delta)$ modifies into

$$M_{22} + M_{21}\Delta M_{12} \qquad\qquad (5.22)$$

which is an affine expression in Δ. Consequently, performance robustness limits the allowable size of Δ (the uncertainty) as $\|M_{22} + M_{21}\Delta M_{12}\|_\infty$ must be less than or equal to γ_i.

Fortunately, the specific structure of the set of models as given in Definition 5.2-4 gives rise to an affine expression (5.22) in the uncertainty Δ to evaluate the control objective function for each system captured in the set of models. Using system identification, the uncertainty can be reduced and/or modified. As such, the structure of (5.22) will be exploited in the next chapter to formulate a (control relevant) identification problem to estimate the set of models $\mathcal{P}_i$, by employing the knowledge of a stabilizing controller C_i that is implemented on the (unknown) plant P_o. To complete the analysis of this chapter, the robust control design of step 2 in Procedure 2.5-4 will be discussed.

---

[2]In the case of unstructured Δ_R.

5.4   Robust Control Design

5.4.1   Relation to the procedure being followed

The set of models to be used throughout Procedure 2.5-4 has been structured according to Definition 5.2-4. In the previous sections, the favourable properties of such a set of models have been summarized and the various tests mentioned in Remark 2.5-6 have been analyzed. Consequently, once a set of models has been characterized by the variables mentioned in Remark 5.2-7, the possibility can be exploited to (re)design a robust controller on the basis of the set. Basically, this constitutes step 2 in Procedure 2.5-4, where a controller C_{i+1} must be designed on the basis of $\mathcal{P}_i$ that should satisfy performance robustness test in (2.20).

In the sequel of this chapter it is assumed that the variables mentioned in Remark 5.2-7 are available to complete the characterization of the set of models $\mathcal{P}_i$. A more detailed discussion on how to obtain a nominal factorization $(\hat{N}_i, \hat{D}_i)$ and the stable and stably invertible weighting filters $(\hat{V}_i, \hat{W}_i)$ is postponed until chapter 6. Using this set $\mathcal{P}_i$ and the results mentioned in the previous section, the design of a controller C_{i+1} in step 2 of Procedure 2.5-4, that should satisfy (2.20), will be discussed below.

In order to be able to satisfy

$$\|J(P, C_{i+1})\|_\infty \leq \gamma_{i+1} < \gamma_i \ \forall P \in \mathcal{P}_i$$

as mentioned in (2.20), the controller C_{i+1} can be designed by minimizing

$$\min_C \sup_{P \in \mathcal{P}_i} \|J(P, C)\|_\infty. \tag{5.23}$$

In this way, the value γ_{i+1} in

$$\sup_{P \in \mathcal{P}_i} \|J(P, C_{i+1})\|_\infty \leq \gamma_{i+1}$$

is minimized in order to satisfy the requirement $\gamma_{i+1} < \gamma_i$ as mentioned in (2.20). Formally, (5.23) is a robust control design problem, wherein the controller C_{i+1} is being designed such that the worst case performance $J(P, C_{i+1}) \ \forall P \in \mathcal{P}_i$ is being optimized. Computation of such a robust controller can be done by existing techniques based on μ-synthesis (Doyle *et al.* 1992, Packard and Doyle 1993, Zhou *et al.* 1996).

5.4.2   Standard plant description

In order to solve the robust control design problem, consider the weighting functions U_2, U_1 of the performance specification (5.9) and the set of models $\mathcal{P}_i$ structured accordingly to (5.5). On the basis of this information, the robust control synthesis can be formulated due to the analysis performed in Corollary 5.3-9. Substitution

of $C = C_{i+1}$ in Corollary 5.3-9 yields a necessary and sufficient condition for the expression (2.20) to hold. Hence, the minimization (5.23) to synthesise a controller C_{i+1} can be replaced by the minimization

$$\min_{C} \mu_{\mathbf{\Delta}} \left(\begin{bmatrix} \hat{W}_i^{-1} & 0 \\ 0 & U_2 \end{bmatrix} T_{ext}(\hat{P}_i, C_i, C) \begin{bmatrix} -\hat{V}_i^{-1} & 0 \\ 0 & U_1 \end{bmatrix} \right) \tag{5.24}$$

where $T_{ext}(\hat{P}_i, C_i, C)$ is given in (5.21).

Basically, (5.24) is a μ-synthesis problem that can be tackled by using the upper bound (5.16) and solving

$$\min_{C} \inf_{D_l, D_r \in \mathcal{D}} \left\| D_l \begin{bmatrix} \hat{W}_i^{-1} & 0 \\ 0 & U_2 \end{bmatrix} T_{ext}(\hat{P}_i, C_i, C) \begin{bmatrix} -\hat{V}_i^{-1} & 0 \\ 0 & U_1 \end{bmatrix} D_r^{-1} \right\|_{\infty} \tag{5.25}$$

iteratively for the scaling matrices D_l, D_r and the controller C subjected to internal stability of the feedback connection of C and $\hat{P}_i$. This iteration is known as the D-K iteration[3] and for fixed scaling D_l, D_r with $D_l, D_r^{-1} \in R\mathcal{H}_{\infty}$ (5.25) is an $\mathcal{H}_{\infty}$ optimization problem, for which commercially available software exists, see e.g. Zhou *et al.* (1996).

It should be mentioned that this section of the thesis does not aim at giving a complete overview of the properties and the computational procedure associated to μ-synthesis using a D-K iteration. Instead, the results and the computational procedure of the D-K iteration will be used to tackle the robust control design problem. Although convergence of the D-K iteration is not guaranteed, several successful applications have been reported in the literature. Furthermore, it should be stressed that precise minimization of (5.24) is not needed. If suffices to find a controller C_{i+1} that is able to satisfy (5.20).

In order to use the available standard results on $\mathcal{H}_{\infty}$ controller synthesis, the controller C must be extracted from the transfer function M given in (5.12). This can be done by representing the transfer function M as a lower fractional transformation $\mathcal{F}_l(G, C)$, in which the controller C to be computed has been isolated. Subsequently, for fixed D-scalings, an $\mathcal{H}_{\infty}$ controller can be computed by minimizing $\|D_l \mathcal{F}_l(G, C) D_r\|_{\infty}$, as illustrated in Figure 5.3.

From Figure 5.3 it can be seen that the controller C to be computed is extracted from a transfer function G. The transfer G only contains the variables needed to construct the set of models, as mentioned in Remark 5.2-7, and the weightings U_2 and U_1 associated to the performance characterization given in (5.9). An expression for the transfer function G (the standard plant description) is given in the following corollary.

---

[3]The naming D-K iteration is widely used in the literature and is adopted here, although D-C would be more appropriate.
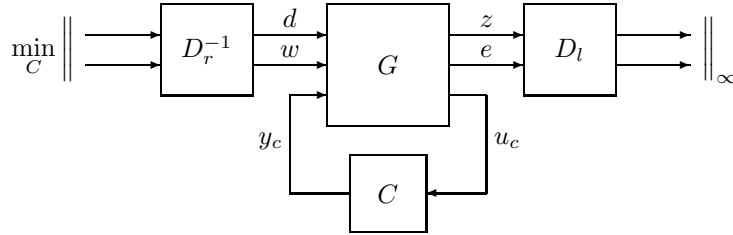
Fig. 5.3: Controller synthesis via $\mathcal{H}_\infty$ optimization for fixed D-scaling.

Corollary 5.4-1 *Consider the map M given in (5.12), then $M = \mathcal{F}_l(G, C)$ where G is given by*

$$
G = \begin{bmatrix} \hat{W}_i^{-1} & 0 & 0 \\ 0 & U_2 & 0 \\ 0 & 0 & I \end{bmatrix} \left[\begin{array}{ccc|c|cc} \hat{D}_i^{-1} N_{c,i} & 0 & \hat{D}_i^{-1} & \hat{D}_i^{-1} \\ (D_{c,i} + \hat{P}_i N_{c,i}) & 0 & \hat{P}_i & \hat{P}_i \\ 0 & 0 & I & I \\ \hline -(D_{c,i} + \hat{P}_i N_{c,i}) & I & -\hat{P}_i & -\hat{P}_i \end{array} \right] \begin{bmatrix} -\hat{V}_i^{-1} & 0 & 0 \\ 0 & U_1 & 0 \\ 0 & 0 & I \end{bmatrix}
$$

Proof: The entries of G can be found by the map from $col(d, w, y_c)$ onto $col(z, e, u_c)$ in Figure 5.3. $\hfill\square$

It can be noted that finding $\mathcal{H}_\infty$ norm-based controllers for a control objective function $J(P, C) = U_2 T(P, C) U_1$ has been studied also in Bongers (1994), Schrama (1992b) or McFarlane and Glover (1992). In these references it has been made clear that minimizing the (unweighted and nominal) performance characterization

$$
\min_C \|T(P, C)\|_\infty \tag{5.26}
$$

can be given a special interpretation. The minimization of (5.26) aims at finding a controller that has a stability robustness margin which is optimal for a set of models structured by additive coprime factor perturbations, similar as given in (5.1).

The control design discussed here is a generalization of the robust controller synthesis as presented in e.g. Bongers (1994) or McFarlane and Glover (1992). It can be verified from Corollary 5.4-1 that by ignoring the map from d onto z (representing the uncertainty), G reduces to

$$
\begin{bmatrix} U_2 & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{cc|c} 0 & \hat{P}_i & \hat{P}_i \\ 0 & I & I \\ \hline I & -\hat{P}_i & -\hat{P}_i \end{array} \right] \begin{bmatrix} U_1 & 0 \\ 0 & I \end{bmatrix}
$$

and $M = \mathcal{F}_l(G, C) = U_2 T(\hat{P}_i, C) U_1$. In the special case of a *diagonal* weighting function $U = diag(U_{in}, U_{out}^{-1})$ with $U_2 = U$ and $U_1 = U^{-1}$, the controller C_{i+1} that

minimizes $\|UT(\hat{P}_i, C_{i+1})U^{-1}\|_\infty$ can be found by loop shaping techniques (Bongers 1994, pp. 107-108). Explicit state space formulae of the optimal controller for this special case can be found in Bongers (1994) or McFarlane and Glover (1992).

As mentioned before, finding a controller by minimizing $\|D_l\mathcal{F}_l(G,C)D_r\|_\infty$ iteratively using the scalings D_l, D_r and the controller C can be solved by available computational algorithms (Zhou *et al.* 1996). As this D-K iteration aims at minimizing (an upper bound) on (5.24), this controller synthesis technique can be used to find a robust and enhanced performing controller C_{i+1} in step 2 of Procedure 2.5-4. With the set of models $\mathcal{P}_i$ structured accordingly to (5.5) and the performance characterization of (5.9), the tools summarized in this chapter constitute the ingredients to evaluate the closed performance assessment test (2.19) and the closed loop validation test of (2.20) and (2.21) in a non-conservative way. Furthermore, the robust control design method on the basis of a set of models $\mathcal{P}_i$ in step 2 of Procedure 2.5-4 has been presented.

6

Control Relevant Identification of a Model Uncertainty Set

6.1 Scope of This Chapter

6.1.1 Estimating a set of models

Having characterized the set of models and its favourable properties in the previous chapter, the problem how to actually obtain the set remains unrequited. The purpose of this chapter is to complete this task and to estimate a set of models by means of system identification techniques. The estimated set of models will be denoted by model uncertainty set to stress that the set of models is obtained via system identification techniques and characterizes the limited knowledge available on the plant P_o.

For the estimation of the model uncertainty set, the fractional approach for identification on the basis of closed-loop data, as outlined in chapter 4, will be employed to complete the fractional characterization of the set of models presented in chapter 5. In this chapter new results on the closed-loop relevant identification of nominal model and model uncertainty set are presented. It is shown how the algebraic approach of fractional model representations can be used to construct a set of models or model uncertainty set on the basis of closed-loop data coming from a feedback connection $\mathcal{T}(P_o, C)$.

Finally it can be noted that the attention is focused on finding a model uncertainty set of limited complexity. As motivated before, the rationale behind this approach is to avoid the synthesis of robust controllers on the basis of highly complex models as much as possible. Such a synthesis inevitably will lead to high order controllers for which the computation may be badly conditioned. For that purpose, the model uncertainty set is based on a nominal model and a model uncertainty characterization that both exhibit a relatively low complexity.

6.1.2   Link to procedure

Referring to Procedure 2.5-4, the estimation of a set of model $\mathcal{P}$ appears both in step 1 and step 3. In this perspective, both steps are similar and require the estimation of a sets of models $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ that should satisfy $P_o \in \mathcal{P}_i$ and $P_o \in \mathcal{P}_{i+1}$ respectively for step 1 and step 3. From an identification point of view, the estimation of the set of models in either step 1 or step 3 of Procedure 2.5-4 is similar and is the identification problem discussed in this chapter.

However, additional to the conditions that the actual plant P_o should lie in the estimated set of models, the sets $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ are subjected to a "quality test" or "validation test" mentioned respectively in (2.19) and (2.21). Estimating the set of models $\mathcal{P}_i$ in step 1 of Procedure 2.5-4 such that γ_i is as small as possible, formally could be achieved by minimizing

$$\sup_{P \in \mathcal{P}_i} \|J(P, C_i)\|_\infty \tag{6.1}$$

subjected to the condition $P_o \in \mathcal{P}_i$. Similarly, the estimation of the set $\mathcal{P}_{i+1}$ in step 3 of Procedure 2.5-4 such that (2.21) will be satisfied can be achieved by solving the minimization

$$\sup_{P \in \mathcal{P}_{i+1}} \|J(P, C_{i+1})\|_\infty \tag{6.2}$$

subjected to the condition $P_o \in \mathcal{P}_{i+1}$.

Clearly, the optimization involved in both the identification of $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ to satisfy the "quality test" or "validation test" is similar. The only difference between step 1 and step 3 of Procedure 2.5-4 is the controller C being used in the optimization. To avoid a separate discussion of the above mentioned optimization problems, the identification problem addressed in this chapter is concerned solely with the estimation of a set of models $\mathcal{P}$ using a controller C. Consequently, applying the appropriate subscripts will reflect either the estimation of the set $\mathcal{P}_i$ using the controller C_i or the estimation of the set $\mathcal{P}_{i+1}$ using the controller C_{i+1}.

6.1.3   Separate estimation of nominal model and uncertainty

The identification of the set of models $\mathcal{P}$ is far from trivial. This is due to the following two problems. Firstly, the actual plant P_o must lie in the estimated set of models $\mathcal{P}$. Secondly, the estimation of the set of models requires an optimization that involves a (worst-case) closed-loop evaluation of the control objective function $J(P, C)$ over all models P that lie in the set of models. The latter makes the estimation of the set of models tuned towards the intended control application and can be labelled as a control relevant identification of a set of models.

According to Definition 2.2-16, the structure of the set of models is determined by a factorization $(\hat{N}, \hat{D})$ of a nominal model $\hat{P}$ and the weighting functions $(\hat{V}, \hat{W})$.

Consequently, the set of models $\mathcal{P}$ can be considered to be parametrized by the variables $(\hat{N}, \hat{D}, \hat{V}, \hat{W})$ as mentioned in Remark 5.2-7. Omitting the indices i and $i+1$ for notational convenience, estimating a set of models $\mathcal{P}$ that takes into account the minimization of (6.1) or (6.2) can be solved by solving the following minimization

$$\min_{N,D,V,W} \sup_{P \in \mathcal{P}} \|J(P,C)\|_{\infty} \qquad (6.3)$$

that is subjected to both $P_o \in \mathcal{P}$ and internal stability of the feedback connection $\mathcal{T}(\hat{P}, C)$. At the current state, the minimization of (6.3) using the variables (N, D, V, W) simultaneously, cannot be solved directly. Therefore, the minimization of (6.3) is tackled by estimating the rcf $(\hat{N}, \hat{D})$ and the pair $(\hat{V}, \hat{W})$ separately.

- *Estimation of a nominal model*
  This involves the estimation of $\hat{P} = \hat{N}\hat{D}^{-1}$ such that (6.3) is being minimized using the rcf (N, D) only, subjected to internal stability of $\mathcal{T}(\hat{P}, C)$. The pair (V, W) in (6.3) is unknown and assumed to vary freely, thereby satisfying the condition $P_o \in \mathcal{P}$.

- *Estimation of uncertainty*
  This consists of the characterization of an upper bound on Δ_R in (5.4) via $(\hat{V}, \hat{W})$ such that (6.3) is being minimized using (V, W) only, subjected to $P_o \in \mathcal{P}$. Subsequently, the variables (N, D) are fixed to the rcf $(\hat{N}, \hat{D})$ estimated previously.

Clearly, by the separate estimation of the rcf $(\hat{N}, \hat{D})$ and the weighting functions $(\hat{V}, \hat{W})$ only an upper bound on (6.3) can be minimized. However, the introduction of a separation between the estimation of a nominal model and an uncertainty characterization makes the control relevant estimation of a set of models more tractable. Currently available tools to estimate a nominal model and to characterize uncertainty can be applied and will be used throughout this chapter. Furthermore, due to the separation being made, the attention can be focused on finding models of limited complexity, to avoid the synthesis of robust controllers on the basis of highly complex models as much as possible.

6.2   Estimation of a Nominal Factorization

6.2.1   Feedback relevant estimation of a nominal factorization

Following the separate identification of nominal model and uncertainty, first the estimation of a nominal model $\hat{P}$ is discussed. The minimization of (6.3) using the rcf (N, D) only, will be used to estimate a nominal model $\hat{P}$ of limited complexity. Hence, to obtain a nominal model $\hat{P}$ or nominal rcf $(\hat{N}, \hat{D})$, the minimization

$$\min_{N,D} \sup_{P \in \mathcal{P}} \|J(P,C)\|_{\infty} \qquad (6.4)$$

must be tackled, where $P_o \in \mathcal{P}$ and (N, D) is a rcf of limited complexity. The controller C in (6.4) indicates the controller currently implemented on the plant P_o. It is perhaps superfluous to mention that the controller C may reflect C_i in step 1 or the controller C_{i+1} in step 3 of Procedure 2.5-4. For reasons of clarity, the attention is focused on step 1, thereby assuming that the controller C_i is implemented on the plant P_o and a factorization $(\hat{N}_i, \hat{D}_i)$ of a nominal model $\hat{P}_i$ must be estimated.

Feedback relevant identification criterion

Clearly, the set $\mathcal{P}$ is still unknown and the variables (V, W) in (6.3) are not used in the minimization of (6.4). As the knowledge of the set of models is incomplete, the minimization of (6.4) cannot be solved directly. Instead, an identification problem to estimate a rcf of a nominal model can be formulated by evaluating $\|J(P, C_i)\|_\infty$ only, using the following triangular inequality (Schrama 1992b).

$$\|J(P, C_i)\|_\infty \leq \|J(P_o, C_i)\|_\infty + \|J(P, C_i) - J(P_o, C_i)\|_\infty \qquad (6.5)$$

As $\|J(P_o, C_i)\|_\infty$ in (6.5) does not depend on the nominal model, the rcf $(\hat{N}_i, \hat{D}_i)$ of a nominal model $\hat{P}_i = \hat{N}_i \hat{D}_i^{-1}$ found by minimizing

$$\min_{N, D} \|J(P, C_i) - J(P_o, C_i)\|_\infty \qquad (6.6)$$

can be used to formulate a control relevant identification of a nominal model $\hat{P}_i$.

Performing the minimization of (6.6) and estimating such a nominal model $\hat{P}_i$ via the estimation of a nominal factorization $(\hat{N}_i, \hat{D}_i)$ can be done by the tools presented in chapter 4 and has also been studied extensively in de Callafon and Van den Hof (1995b) or Van den Hof et $al.$ (1995). To summarize the results of chapter 4, it is worthwhile to mention that Section 4.3.3 indicates how to access a rcf $(N_{o,F}, D_{o,F})$ of the plant P_o by a simple filtering of the signals present in the feedback connection $\mathcal{T}(P_o, C_i)$. The appropriate filtering, as mentioned in (4.24) and Lemma 4.3-5, yields an auxiliary signal x and a rcf $(N_{o,F}, D_{o,F})$ of the plant P_o is accessible by considering the map from x to $col(y, u)$. This result has been summarized in Proposition 4.3-11. Moreover, Proposition 4.3-11 has indicated the fact that the estimation of a factorization is an equivalent open-loop identification problem, since the auxiliary signal x is uncorrelated with the noise acting on the feedback system $\mathcal{T}(P_o, C_i)$.

For a model $P(\theta)$ parametrized by a factorization $(N(\theta), D(\theta))$ with $\theta \in \Theta$, where Θ is given by

$$\Theta := \{\theta \in \mathbb{R}^n \mid (N(\theta), D(\theta)) \in R\mathcal{H}_\infty\} \qquad (6.7)$$

the following result can be used to minimize (6.6).

Lemma 6.2-1 *Let the plant P_o and a controller C_i create an internally stable feedback connection $\mathcal{T}(P_o, C_i)$ and let $(N_{o,F}, D_{o,F})$ be the rcf of P_o as given in*

Proposition 4.3-11, where F is an appropriate filter according to Lemma 4.3-5 on the basis of the controller C_i. Consider any model $P(\theta) = N(\theta)D(\theta)^{-1}$ that satisfies $T(P(\theta), C_i) \in R\mathcal{H}_\infty$ then the minimization of (6.6) with $J(P, C_i)$ as given in (5.9) equals

$$\min_{\theta \in \Theta} \left\| U_2 \left(\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} - \begin{bmatrix} N(\theta) \\ D(\theta) \end{bmatrix} \right) F \begin{bmatrix} C_i & I \end{bmatrix} U_1 \right\|_\infty \tag{6.8}$$

where $(N(\theta), D(\theta))$ is any rcf of $P(\theta)$ that satisfies

$$D(\theta) + C_i N(\theta) = F^{-1}. \tag{6.9}$$

Proof: For the proof of (6.8), consider $(N_{o,F}, D_{o,F})$ to be a *rcf* of P_o. Subsequently, $J(P_o, C_i)$ can be written as

$$J(P_o, C_i) = U_2 \begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} (D_{o,F} + C_i N_{o,F})^{-1} \begin{bmatrix} C_i & I \end{bmatrix} U_1$$

and using the fact that $(D_{o,F} + C_i N_{o,F}) = F^{-1}$ from Corollary 4.3-10, this can be rewritten into

$$J(P_o, C_i) = U_2 \begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} F \begin{bmatrix} C_i & I \end{bmatrix} U_1.$$

In a similar way, with $(N(\theta), D(\theta))$ as a *rcf* of $P(\theta)$, $J(P(\theta), C_i)$ can be written as

$$J(P(\theta), C_i) = U_2 \begin{bmatrix} N(\theta) \\ D(\theta) \end{bmatrix} F \begin{bmatrix} C_i & I \end{bmatrix} U_1$$

under the condition that (6.9) holds. With this condition, $\|J(P_o, C_i) - J(P, C_i)\|_\infty$ can be written as in (6.8).

To prove that indeed there exists a *rcf* $(N(\theta), D\theta))$ of $P(\theta)$ such that (6.9) holds, consider a *rcf* $(\bar{N}(\theta), \bar{D}(\theta)) \in R\mathcal{H}_\infty$ that satisfies

$$\bar{D}(\theta) + C_i \bar{N}(\theta) = \bar{F}^{-1} \tag{6.10}$$

where $\bar{F} \neq F$. Proving that there exists a *rcf* $(N(\theta), D(\theta)) \in R\mathcal{H}_\infty$ that satisfies (6.9) can be done by proving the existence of a stable and stably transfer function Q that modifies the *rcf* $(\bar{N}(\theta), \bar{D}(\theta)) \in R\mathcal{H}_\infty$ into the *rcf* $(N(\theta), D(\theta)) \in R\mathcal{H}_\infty$ via

$$\begin{aligned} N(\theta) &:= \bar{N}(\theta)Q \\ D(\theta) &:= \bar{D}(\theta)Q \end{aligned}. \tag{6.11}$$

Under the condition that $T(P(\theta), C_i) \in R\mathcal{H}_\infty$, where $P(\theta) = \bar{N}(\theta)\bar{D}(\theta)^{-1} = N(\theta)D(\theta)^{-1}$, such a Q with $Q, Q^{-1} \in R\mathcal{H}_\infty$ can always be found and can be seen

as follows. Postmultiplying (6.10) with Q and posing the requirement $\bar{F}^{-1}Q = F^{-1}$ yields

$$(\bar{D}(\theta) + C_i\bar{N}(\theta))Q = F^{-1} \tag{6.12}$$

Lemma 4.3-5 indicates that an appropriate filter F must be given by the form $F = W\tilde{D}_{c,i}$, where $W, W^{-1} \in R\mathcal{H}_\infty$. Furthermore, $(\bar{D}(\theta) + C_i\bar{N}(\theta))$ can be rewritten into $\tilde{D}_{c,i}^{-1}\bar{\Lambda}$, where $\bar{\Lambda}(\theta) := \tilde{D}_{c,i}\bar{D}(\theta) + \tilde{N}_{c,i}\bar{N}(\theta)$ satisfies $\bar{\Lambda}(\theta), \bar{\Lambda}(\theta)^{-1} \in R\mathcal{H}_\infty$ according to Lemma 4.2-6. Using these results, (6.12) can be rewritten into

$$\tilde{D}_{c,i}^{-1}\bar{\Lambda}(\theta)Q = \tilde{D}_{c,i}^{-1}W^{-1}$$

yielding $Q = \bar{\Lambda}(\theta)^{-1}W^{-1}$ and $Q, Q^{-1} \in R\mathcal{H}_\infty$.   □

The result mentioned in Lemma 6.2-1 indicates how to solve the control relevant identification of a nominal model or nominal factorization. The difference in the control objective mentioned in (6.6) is simply restated as a weighted additive difference between coprime factorizations. The weighting is, however, still parametrized in terms of the coprime factorization $(N(\theta), D(\theta))$ to be estimated. As such, a "parametrized weighting" must be used in the minimization of (6.8).

Remark 6.2-2 *It should be noted that the appearance of a "parametrized weighting" is due to the control or closed-loop relevant criterion (6.6) used to formulate an appropriate closed-loop relevant system identification problem. The appearance of the parametrized weighting is not caused by the fractional approach being used to handle the closed-loop identification problem.*

A closed-loop relevant criterion that is defined via an unweighted difference between (input) sensitivity functions

$$\|\frac{1}{1 + CP_o} - \frac{1}{1 + CP(\theta)}\| \tag{6.13}$$

can be mentioned as a simple SISO example. In this case, the closed-loop relevant criterion (6.13) can be written into a weighted additive difference between P_o and P as follows

$$\frac{1}{1 + CP_o}(P_o - P(\theta))\frac{C}{1 + CP(\theta)} \tag{6.14}$$

which indicates the need for a "parametrized weighting".

The fact that the weighting in the additive difference between the coprime factorization $(N_{o,F}, D_{o,F})$ and $(N(\theta), D(\theta))$ in (6.8) depends on the factorization $(N(\theta), D(\theta))$ to be estimated, is mainly due to the condition of finding a low complexity model, for which the McMillan degree can be prespecified. In case the condition on the limited complexity of the nominal model becomes superfluous, a parametrization of the coprime factorization $(N(\theta), D(\theta))$ can be proposed, in which the resulting

additive weighting in (6.8) will not depend on the factorization $(N(\theta), D(\theta))$ to be estimated.

Such a parametrization is found by employing the dual-Youla parametrization, as discussed in Section 4.2.2. Following Lemma 4.2-7, the coprime factorization $(N(\theta), D(\theta))$ of the nominal model $P(\theta)$ can be parametrized via

$$
\begin{aligned}
N(\theta) &= N_x + D_{c,i}R(\theta) \\
D(\theta) &= D_x - N_{c,i}R(\theta)
\end{aligned}
\tag{6.15}
$$

where $R(\theta)$ is now be considered as the unknown transfer function to be estimated. Similar to Lemma 4.2-7, $(N_{c,i}, D_{c,i})$ is a rcf of the controller C_i and (N_x, D_x) is a rcf of an auxiliary model P_x that satisfies $T(P_x, C_i) \in R\mathcal{H}_\infty$. With the parametrization of (6.15) the following corollary can be formulated.

Corollary 6.2-3 *Consider a parametrization of a nominal model $P(\theta) = N(\theta)D^{-1}(\theta)$ where the rcf $(N(\theta), D(\theta))$ is parametrized according to (6.15). Then the minimization of (6.6) with $J(P, C_i)$ as given in (5.9) equals*

$$
\min_{\theta \in \Theta} \left\| U_2 \left(\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} - \begin{bmatrix} N(\theta) \\ D(\theta) \end{bmatrix} \right) F \begin{bmatrix} C_i & I \end{bmatrix} U_1 \right\|_\infty
\tag{6.16}
$$

where $F = (D_x + C_i N_x)^{-1}$ is a fixed, known and appropriate filter according to Lemma 4.3-5.

Proof: The condition (6.9) for the parametrization given in (6.15) is satisfied trivially, as $D(\theta) + C_i N(\theta)$ can be rewritten into

$$
(D_x - N_{c,i}R(\theta)) + C_i(N_x + D_{c,i}R(\theta))
$$

which reduces to $D_x + C_i N_x = F^{-1}$ with $C_i = N_{c,i}D_{c,i}^{-1}$. The fact that F is indeed an appropriate filter according to Lemma 4.3-5 can be seen by rewriting $F = (D_x + C_i N_x)^{-1}$ into

$$
(D_x + C_i N_x)^{-1} = W\tilde{D}_{c,i}
$$

where $W = (\tilde{D}_{c,i}D_x + \tilde{N}_{c,i}N_x)^{-1}$. From Lemma 4.2-6 it can be seen that $W, W^{-1} \in R\mathcal{H}_\infty$ provided that $T(P_x, C_i) \in R\mathcal{H}_\infty$. □

Although the weighting of the additive difference between the rcf $(N_{o,F}, D_{o,F})$ and $(N(\theta), D(\theta))$ in Corollary 6.2-3 is known and fixed, the McMillan degree of the resulting nominal model $P(\theta) = N(\theta)D^{-1}(\theta)$ is influenced by several factors. Clearly, the McMillan degree is influenced by the McMillan degree of the transfer function $R(\theta)$ being estimated. However, the complexity of the nominal model is also increased by the McMillan degree of the controller C_i and the auxiliary model P_x being used.

The filter F, which is solely used to access the coprime factorization $(N_{o,F}, D_{o,F})$ of the plant P_o, now also effects the complexity of the nominal model. This fact has been recognized by system identification procedures that explicitly use a dual-Youla parametrization, see e.g. Tay *et al.* (1989) or Lee *et al.* (1995). A possible solution is to pursue a model reduction to lower the order of the model $P(\hat{\theta})$ being estimated (Lee *et al.* 1993*a*). Clearly, directly estimating a limited complexity nominal model is preferable.

The need for closed-loop experiments

The presence of a weighting filter that is based on a closed-loop transfer function involving the unknown plant P_o, indicates the need for closed-loop experiments. In case the input sensitivity $S_{in} = (I + C_i P_o)^{-1}$ appears as a weighting or as closed-loop transfer function that needs to be accessed, closed-loop experiments, with a controller C applied to the plant P_o, can help providing this information on the input sensitivity function S_{in}.

As an example it can be mentioned that F in (6.8) satisfies $(D_{o,F} + C_i N_{o,F}) = F^{-1}$ only if $D_{o,F} = S_{in} F^{-1}$ and $N_{o,F} = P_o S_{in} F^{-1}$. Hence, closed-loop experiments are needed to access this *rcf* $(N_{o,F}, D_{o,F})$, as the closed-loop transfer function $P_o S_{in}$ or S_{in} is considered to be unknown. Similarly, in (6.14) the weighting filter $(1 + C P_o)^{-1}$ is unknown and such information can be obtained by performing closed-loop experiments. This has for example been illustrated in the two-stage identification, discussed in Section 3.3.4. It was shown in (3.59) or (3.61) that the input sensitivity S_{in} enters automatically as a weighting in the identification criterion due to the closed-loop experiments.

Finally it should be noted that if the controller C applied to the plant P_o is updated from C_i to C_{i+1}, the input sensitivity S_{in} changes. Assuming that full knowledge of the plant P_o, or similarly the input sensitivity $(I + C_i P_o)^{-1}$ is not available, renewed information of the input sensitivity $(I + C_{i+1} P_o)^{-1}$ can be obtained by performing new closed-loop experiments.

Many systems exhibit a poorly damped or even unstable behaviour and closed-loop experiments are unavoidable to obtain reliable data of the system in a limited amount of time. Although in these situations closed-loop experiments are unavoidable, they also serve an important advantage for the purpose of finding control relevant models.

6.2.2   Estimating a limited complexity factorization

Despite the reasonably simple additive weighted difference between the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o and the *rcf* $(N(\theta), D(\theta))$ of the nominal model $P(\theta)$ mentioned in Lemma 6.2-1, the minimization of (6.6), for a nominal model of limited complexity or a prespecified McMillan degree (smaller than the McMillan degree of

the plant P_o), remains a complicated optimization problem. The complication of the optimization is mainly due to the following two items.

Firstly, an $\mathcal{H}_\infty$ norm-based optimization must be used to find a control relevant nominal model. Although such a optimization can be solved by an optimization that uses sub-gradients (Luenberger 1969), the $\mathcal{H}_\infty$ norm-based optimization criterion (6.8) on the basis of experimental (time domain) data can be quite involved. An alternative approach to circumvent the $\mathcal{H}_\infty$ norm-based optimization would be the approximation of the $\mathcal{H}_\infty$ norm in (6.8) by the minimization of an $\mathcal{H}_2$ norm criterion. This is motivated by the fact that an $\mathcal{L}_2$-norm approximation tends to $\mathcal{L}_\infty$-norm approximation, provided that smoothness conditions[1] on the plant P_o are satisfied (Caines and Baykal-Gürsoy 1989).

Similar arguments are used also in Bitmead *et al.* (1990*a*) or Schrama (1992*b*) to approximate an $\mathcal{H}_\infty$ norm-based identification criterion. However, in this thesis an $\mathcal{H}_\infty$ norm-based optimization criterion will be used, where the optimization criterion (6.8) will be evaluated on the basis of experimentally obtained frequency domain data. A more thorough discussion on this topic can be found in Section 6.2.4.

The second item that influences the complexity of the optimization problem is the fact that the weighting, or the filter F, in (6.8) should satisfy the condition mentioned in (6.9). From an optimization point of view, (6.9) constitutes a parametrization restriction that should be dealt with while estimating a *rcf* $(N(\theta), D(\theta))$. As mentioned before, this restriction is due to the fact that a limited complexity nominal factorization $(N(\theta), D(\theta))$ is required. Without such a requirement, the alternative parametrization mentioned in Corollary 6.2-3 can be used to eliminate this restriction.

An approach to deal with the restriction (6.9) is to parametrize the filter F such that it satisfies (6.9). Hence, the filter F needs to be parametrized via

$$F(\theta) = (D(\theta) + C_i N(\theta))^{-1} \tag{6.17}$$

according to (6.9). However, a *rcf* $(N(\theta), D(\theta))$ of a model $P(\theta)$ is not unique, which leads to a non-unique parametrization of the filter $F(\theta)$ in (6.17). This additional freedom can be characterized by the freedom in the *rcf* of the model $P(\theta)$

$$\begin{bmatrix} N(\theta) \\ D(\theta) \end{bmatrix} = \begin{bmatrix} \bar{N}(\theta) \\ \bar{D}(\theta) \end{bmatrix} Q \tag{6.18}$$

where Q is any stable and stably invertible transfer function. It can be seen from (6.18) that the *rcf* $(N(\theta), D(\theta))$ exhibits redundant dynamics Q that eventually cancels out in the operation $P(\theta) = N(\theta)D(\theta)^{-1}$ leading to a low order model $P(\theta) = \bar{N}(\theta)\bar{D}(\theta)^{-1}$. Furthermore, it can be observed from (6.6) that the control relevant identification criterion does not distinguish between different (non-unique) *rcf*'s of the nominal model

---

[1] To verify the smoothness conditions, knowledge of the actual plant P_o is required.

$P(\theta)$. Due to the non-uniqueness of the *rcf* of $P(\theta)$, all stable and stably invertible Q's in (6.18) give the same value of the control relevant identification criterion (6.8).

Similar arguments also hold for the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o that is accessible via closed-loop experiments mentioned in Proposition 4.3-11. However, compared to the non-uniqueness of the *rcf* $(N(\theta), D(\theta))$, the *rcf* $(N_{o,F}, D_{o,F})$ is unique for a given filter F and given by

$$\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} = \begin{bmatrix} P_o S_{in} F^{-1} \\ S_{in} F^{-1} \end{bmatrix} \tag{6.19}$$

as mentioned in Remark 4.3-6. It can be observed from (6.19) that, although the *rcf* $(N_{o,F}, D_{o,F})$ is unique for a given filter F, $(N_{o,F}, D_{o,F})$ may still exhibit redundant dynamics.

The parametrized *rcf* $(N(\theta), D(\theta))$ of $P(\theta)$ is used to estimate or approximate the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o accessable via closed-loop experiments. In order to reduce the additional freedom in the *rcf*'s of plant P_o and model $P(\theta)$, and to avoid the estimation of a high order nominal *rcf* $(N(\theta)Q, D(\theta)Q)$, the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o can be required to be normalized. As mentioned in Section 4.2.1, for such a *nrcf* $(N_{o,F}, D_{o,F})$ the freedom in $(N_{o,F}\bar{Q}, D_{o,F}\bar{Q})$ is limited to a unimodular matrix $\bar{Q}$. In this way, any redundant dynamics the *rcf* $(N_{o,F}, D_{o,F})$ in (6.19) may exhibit, can be eliminated.

The filter F, used to access the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o, can be used to normalize $(N_{o,F}, D_{o,F})$. In Van den Hof *et al.* (1995) such an approach to access and estimate a *nrcf* $(N_{o,F}, D_{o,F})$ of the plant P_o has been proposed. This approach uses the estimate of a (high order) auxiliary model P_x to construct a *nrcf* (N_x, D_x) for the filter F mentioned in Corollary 4.3-9. In this way, the *rcf* $(N_{o,F}, D_{o,F})$ in (6.19) modifies into

$$\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} = \begin{bmatrix} P_o \\ I \end{bmatrix} S_{in}(D_x + CN_x) = \begin{bmatrix} P_o \\ I \end{bmatrix} (I + CP_o)^{-1}(I + CP_x)D_x$$

and will approach the *nrcf* (N_x, D_x) in case P_x is a (high order) accurate estimate of the plant P_o (Van den Hof *et al.* 1995).

Instead of estimating or approximating a *nrcf* $(N_{o,F}, D_{o,F})$, here the attention is focused on finding a limited complexity factorization by addressing the minimization stated in Lemma 6.2-1 and taking into account the restriction mentioned in (6.9). Clearly, with a parametrized filter F as mentioned in (6.17), the optimization problem can not be handled by a "standard" identification procedure. This is caused by the fact that the signal x in (4.24) and used to access the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o, cannot be constructed prior to the identification. To deal with these parametrization problems, an iterative approach of subsequent minimization without the parametrization restriction (6.9) and updating the filter F can be proposed (de Callafon *et al.* 1994, de Callafon and Van den Hof 1995*b*).

Starting off from an initial model estimate $P(\hat{\theta})$ with a $rcf\,(N(\hat{\theta}), D(\hat{\theta}))$ and setting $\hat{\theta}_k = \hat{\theta}$, the iterative scheme reads as follows.

1. In step k, compute Q_k such that

$$\begin{bmatrix} N_n(\hat{\theta}) \\ D_n(\hat{\theta}) \end{bmatrix} := \begin{bmatrix} N(\hat{\theta}) \\ D(\hat{\theta}) \end{bmatrix} Q_k$$

   is a $nrcf\,(N_n(\hat{\theta}), D_n(\hat{\theta}))$ and create $F_k = (D_n(\hat{\theta}_k) + C_i N_n(\hat{\theta}_k))^{-1}$ to simulate the input x to access the rcf of the plant P_o.

2. Estimate $(N(\hat{\theta}_{(k+1)}), D(\hat{\theta}_{(k+1)}))$ by the minimization given in (6.8) and discarding the parameter restriction (6.9).

3. Set $k = k + 1$ and go back to step 1.

The updating of the filter F_k in the aforementioned iterative approach simply means that the signal x in (4.24) is refiltered by the updated filter F_k to access a different (and unique) $rcf\,(N_{o,F}, D_{o,F})$ of the plant P_o given by

$$\begin{bmatrix} N_{o,F} \\ D_{o,F} \end{bmatrix} = \begin{bmatrix} P_o S_{in} F_k^{-1} \\ S_{in} F_k^{-1} \end{bmatrix}$$

The normalization and the computation of Q_k in the first step is to restrict the additional freedom in the rcf of the model $P(\theta)$ and the resulting filter F_k. If the iteration converges then the restriction (6.9) has been satisfied, thus a feedback relevant nominal model $\hat{P}$ of the plant P_o has been obtained. A rigorous proof of the convergence of the iteration is not available but extensive simulations reveal promising results.

6.2.3   Parametrization of stable factorizations

To control the McMillan degree of the nominal model $P(\theta) = N(\theta)D^{-1}(\theta)$ being estimated, the factorization $(N(\theta), D(\theta))$ has to be parametrized in a such a way that $N(\theta)$ and $D(\theta)$ should have common stable modes. In the case of a SISO model $P(q, \theta)$ such a parametrization can be given in the following transfer function form

$$\begin{bmatrix} N(q, \theta) \\ D(q, \theta) \end{bmatrix} = \begin{bmatrix} \dfrac{n(q, \theta)}{f(q, \theta)} \\ \dfrac{d(q, \theta)}{f(q, \theta)} \end{bmatrix}$$

where $n(q, \theta)$ and $d(q, \theta)$ denote the numerator polynomial of respectively $N(q, \theta)$ and $D(q, \theta)$, while $f(q, \theta)$ denotes the common (monic and stable) denominator polynomial of $N(q, \theta)$ and $D(q, \theta)$. To parametrize such a (polynomial) factorization in state space form, the following result can be used.

Theorem 6.2-4 *Let*

$$\begin{bmatrix} \hat{N} \\ \hat{D} \end{bmatrix} \in R\mathcal{H}_\infty$$

be given by a stable and minimal state space representation

$$\left(\bar{\mathsf{A}}, \bar{\mathsf{B}}, \begin{bmatrix} \bar{\mathsf{C}}_N \\ \bar{\mathsf{C}}_D \end{bmatrix}, \begin{bmatrix} \bar{\mathsf{D}}_N \\ \bar{\mathsf{D}}_D \end{bmatrix} \right)$$

with $\det\{\bar{\mathsf{D}}_D\} \neq 0$. *Then the following items hold.*

(i) $\det\{\hat{D}\} \not\equiv 0$

(ii) $\hat{P} = \hat{N}\hat{D}^{-1}$ *is given by the state space representation* $(\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$ *with*

$$\begin{cases} \mathsf{A} &=& \bar{\mathsf{A}} - \bar{\mathsf{B}}\bar{\mathsf{D}}_D^{-1}\bar{\mathsf{C}}_D \\ \mathsf{B} &=& \bar{\mathsf{B}}\bar{\mathsf{D}}_D^{-1} \\ \mathsf{C} &=& \bar{\mathsf{C}}_N - \bar{\mathsf{D}}_N\bar{\mathsf{D}}_D^{-1}\bar{\mathsf{C}}_D \\ \mathsf{D} &=& \bar{\mathsf{D}}_N\bar{\mathsf{D}}_D^{-1} \end{cases} \qquad (6.20)$$

and $(\hat{N}, \hat{D})$ *is a rcf of* $\hat{P}$.

Proof: Due to the non-singular matrix $\bar{\mathsf{D}}_D$, $\hat{D}^{-1}$ allows a state space representation $(\mathsf{A}_D, \mathsf{B}_D, \mathsf{C}_D, \mathsf{D}_D)$ that is given by

$$\begin{cases} \mathsf{A}_D &=& \bar{\mathsf{A}} - \bar{\mathsf{B}}\bar{\mathsf{D}}_D^{-1}\bar{\mathsf{C}}_D \\ \mathsf{B}_D &=& \bar{\mathsf{B}}\bar{\mathsf{D}}_D^{-1} \\ \mathsf{C}_D &=& -\bar{\mathsf{D}}_D^{-1}\bar{\mathsf{C}}_D \\ \mathsf{D}_D &=& \bar{\mathsf{D}}_D^{-1} \end{cases}$$

which proves (i).

Due to the common $\bar{\mathsf{A}}$ and $\bar{\mathsf{B}}$ matrix of $(\hat{N}, \hat{D})$, it can be verified that from the state space representation of the series connection of $\hat{N}$ and $\hat{D}^{-1}$, $n = \dim(\bar{\mathsf{A}})$ uncontrollable states can be omitted. This leads to the state space realization $[\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}]$ of $\hat{P}$ given in (6.20). With (6.20), the matrices $\bar{\mathsf{A}}$, $\bar{\mathsf{B}}$, $\bar{\mathsf{C}}_N$, $\bar{\mathsf{C}}_D$ and $\bar{\mathsf{D}}_N$ can be rewritten as $\bar{\mathsf{A}} = \mathsf{A} - \mathsf{B}\mathsf{K}$, $\bar{\mathsf{B}} = \mathsf{B}\bar{\mathsf{D}}_D$, $\bar{\mathsf{C}}_N = \mathsf{C} - \mathsf{D}\mathsf{K}$, $\bar{\mathsf{C}}_D = -\mathsf{K}$ $\bar{\mathsf{D}}_N = \mathsf{D}\bar{\mathsf{D}}_D$, In this way $\hat{N}(z) = ([\mathsf{C} - \mathsf{D}\mathsf{K}][zI - \mathsf{A} + \mathsf{B}\mathsf{K}]^{-1}\mathsf{B} + \mathsf{D})\bar{\mathsf{D}}_D$ and $\hat{D}(z) = (-\mathsf{K}[zI - \mathsf{A} + \mathsf{B}\mathsf{K}]^{-1}\mathsf{B} + I)\bar{\mathsf{D}}_D$, which is proven to be a *rcf* in Nett *et al.* (1984). $\qquad\square$

The result of Theorem 6.2-4 only indicates the ordering and restrictions on the state space realization of the *rcf* $(N(\theta), D(\theta))$ and does not specify the way in which the parameter θ enters in the state space parametrization. As such, Theorem 6.2-4

include various ways to parametrize the state space realization of the *rcf* $(N(\theta), D(\theta))$. If the parametrization guarantees that the state space realization is stable, minimal and satisfies

$$\det\{\bar{\mathsf{D}}_D(\theta)\} \neq 0 \ \forall \theta \in \Theta \tag{6.21}$$

the result in Theorem 6.2-4 is applicable.

Clearly, restricting the parametrization to yield a stable and minimal state space realization that also satisfies (6.21) requires a non-trivial parametrization of the state space matrices. However, a stable (and minimal) state space estimate with non-singular feedthrough matrix $\bar{\mathsf{D}}_D$ will be found in the generic case. This due to the fact that the map from x onto $col(y, u)$ is stable according to Proposition 4.3-11. Furthermore, the map from x onto u is given by $D_{o,F} = (I + CP_o)^{-1}F^{-1} = (I + CP_o)\tilde{D}_cW^{-1}$, which is non-singular by definition.

Relying on the generic case, in general no special purpose parametrization has to be used in order to parametrize the state space matrices of the *rcf* $(N(\theta), D(\theta))$. In this perspective, classical (global) identifiable parametrizations such as canonical (overlapping) parametrizations of state space matrices (Glover and Willems 1974, van Overbeek and Ljung 1982, Corrêa and Glover 1984, Corrêa and Glover 1986, Janssen 1988) can be employed.

Polynomial based parametrizations

It should be noted that a parametrization with a common $\bar{\mathsf{A}}$ and $\bar{\mathsf{B}}$ matrix can also be realized by parametrizing the *rcf* $(N(\theta), D(\theta))$ via a polynomial Matrix Fraction Description (MFD) (Gevers and Wertz 1984, Van den Hof 1989). In such a polynomial MFD, the *rcf* $(N(\theta), D(\theta))$ can be parametrized in a so-called right MFD

$$\begin{bmatrix} N(q,\theta) \\ D(q,\theta) \end{bmatrix} = \begin{bmatrix} B_N(q^{-1},\theta) \\ B_D(q^{-1},\theta) \end{bmatrix} A^{-1}(q^{-1},\theta) = B(q^{-1},\theta)A^{-1}(q^{-1},\theta) \tag{6.22}$$

where $B(q^{-1}, \theta)$ and $A(q^{-1}, \theta)$ are matrix polynomials that should satisfy certain conditions to impose identifiability of the model structure.

Typically, for a model $P(q, \theta)$ having m inputs and p outputs, the polynomial matrix $B(q^{-1}, \theta)$ for the right MFD in (6.22) can be parametrized by

$$B(q^{-1},\theta) = \sum_{k=d}^{d+b-1} B_k \, q^{-k}$$

where $B_k \in \mathbb{R}^{p+m \times m}$, d denotes the number of leading zero matrix coefficients and b the number of non-zero matrix coefficients in $B(q^{-1}, \theta)$. For the right MFD in (6.22), $A(q^{-1}, \theta)$ is parametrized by

$$A(q^{-1},\theta) = I_{m \times m} + \sum_{k=1}^{a} A_k \, q^{-k}$$

where $A_k \in \mathbb{R}^{m \times m}$ and a denotes the number of non-zero matrix coefficients in the monic polynomial $A(q^{-1}, \theta)$. Hence, for a SISO model $P(q, \theta)$, $B(q^{-1}, \theta)$ is a 2×1 polynomial matrix, while $A(q^{-1}, \theta)$ is a single polynomial.

The parameter θ is determined by the corresponding unknown matrix coefficients in the polynomials. Hence,

$$\theta = \left[\; B_d^T \;\; \cdots \;\; B_{d+b-1}^T \;\; A_1^T \;\; \cdots \;\; A_a^T \;\right]^T \tag{6.23}$$

and $\theta \in \mathbb{R}^{((p+m)b+ma) \times m}$ for the right MFD given in (6.22).

Additionally to the full polynomial parametrization presented here, so-called structural parameters d_{ij}, b_{ij} and a_{ij} with $d := \min\{d_{ij}\}$, $b := \max\{b_{ij}\}$, and $a := \max\{a_{ij}\}$ can be used to specify a non-full polynomial parametrization. In this way, the parameter θ as given in (6.23) may contain prespecified zero entries at specific locations. This may be required in a discrete-time $rcf\,(N(q,\theta), D(q,\theta))$ where the value of d_{ij} has a direct connection with the number of time delays from the jth input to the ith output.

Due to the indeterminate q^{-1}, it can be verified that the right MFD given in (6.22) gives rise to a (strictly) proper transfer function matrix of the $rcf\,(N(\theta), D(\theta))$, regardless of the value of the integers $d_{i,j}$, $b_{i,j}$ or $a_{i,j}$. Hence, there are no restrictions on the size of the structural parameters, other than a limitation on the McMillan degree of the resulting model $(N(\hat{\theta}), D(\hat{\theta}))$. For the connection between the structural parameters and the McMillan degree of $(N(\hat{\theta}), D(\hat{\theta}))$, the following result can be given, see also de Callafon *et al.* (1996*a*).

Lemma 6.2-5 *Consider a parameter $\hat{\theta}$ such that $A_a \neq 0$ and $B_{d+b-1} \neq 0$. Define*

$$\eta := \max\{a, d + b - 1\} \tag{6.24}$$

and $\bar{A}(q, \hat{\theta}) := q^\eta A(q^{-1}, \hat{\theta})$, $\bar{B}(q, \hat{\theta}) := q^\eta B(a^{-1}, \hat{\theta})$. Let n be used to denote the McMillan degree of the multivariable transfer function model $[N(q, \hat{\theta})^T \;\; D(q, \hat{\theta})^T]^T$ obtained by (6.22). Then

$$n = deg\; det\{\bar{A}(q, \hat{\theta})\}$$

if and only if $\bar{A}(q, \hat{\theta})$ and $\bar{B}(q, \hat{\theta})$ are right coprime over $\mathbb{R}[q]$.

Proof: With the condition $A_a \neq 0$, $B_{d+b-1} \neq 0$, it follows that $\bar{A}(q, \hat{\theta}) := q^\eta A(q^{-1}, \hat{\theta})$ and $\bar{B}(q, \hat{\theta}) := q^\eta B(q^{-1}, \hat{\theta})$ are polynomial matrices in q. Subsequently, a state space realization $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ for $[N(q, \hat{\theta})^T \;\; D(q, \hat{\theta})^T]^T$ can be obtained, such that $\dim \bar{A} = deg\; det\{\bar{A}(q, \hat{\theta})\}$ and $(\bar{A}, \bar{B})$ controllable, see e.g Chen (1984). Furthermore, $(\bar{A}, \bar{C})$ is observable if and only if $\bar{A}(q, \hat{\theta})$ and $\bar{B}(q, \hat{\theta})$ are right coprime over $\mathbb{R}[q]$, see theorem 6.1 in Chen (1984). $\qquad\square$

Hence, the structural parameters give rise to (an upper bound) on the McMillan degree of the *rcf* $(N(\hat{\theta}), D(\hat{\theta}))$. Subsequently, Theorem 6.2-4 indicates that the McMillan degree of the *rcf* $(N(\hat{\theta}), D(\hat{\theta}))$ is directly related to the McMillan degree of the model $P(\hat{\theta})$ constructed via $P(\hat{\theta}) = N(\hat{\theta})D^{-1}(\hat{\theta})$. For a more detailed discussion on the exact relation between the McMillan degree, the row/column degree of the polynomial matrices $A(q^{-1}, \theta)$, $B(q^{-1}, \theta)$ and the related observability/controllability indices of a model computed by a polynomial MFD, one is also referred to Gevers (1986) or Van den Hof (1992).

Enforcing stable state space realizations

Alternative canonical parametrizations based on balanced state space realizations (Ober 1987, Ober 1991, Chou 1994) can also be used to define an identifiable parametrization. These parametrizations allow the application of simple and affine constraints on the parameters to enforce a stable, minimal (and balanced) state space realization. As a result, stability and minimality of the state space realization of a *rcf* $(N(\theta), D(\theta))$ can be enforced, if needed.

The parametrization results on stable, minimal and balanced state space realization in Ober (1991) and further elaborated in Chou (1994) are based on the general case of continuous-time systems, having (possibly) multiple common Hankel singular values. For discrete-time systems an indirect state space parametrization can be based on a Möbius transformation, since this transformation preserves both stability, minimality and the balanced property of a continuous-time state space realization. Furthermore, the case of distinct Hankel singular values will be discussed here, which can be considered to be the generic case (Chou 1994).

Lemma 6.2-6 *Let $G(s)$ be defined by*

$$G(s) := \begin{bmatrix} N(s) \\ D(s) \end{bmatrix}$$

where $(N(s), D(s))$ is a rcf of the $p \times m$ rational transfer function $P(s)$, then the following statements are equivalent

1. *$G(s)$ is a $(p+m) \times m$ stable rational transfer function matrix of McMillan degree n, with n distinct Hankel singular values.*

2. *$G(s)$ has a state space representation $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ with*

$$\bar{B} := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \text{ with } b_j = [b_{ij}] \in \mathbb{R}^{1 \times m}$$

$$\text{and } b_{1j} > 0 \text{ for } 1 \leq j \leq n \tag{6.25}$$

$$\bar{C} := \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \text{ with } c_j = u_j [b_j b_j^T]^{1/2} \in \mathbb{R}^{(p+m) \times 1},$$

$$u_j \in \mathbb{R}^{(p+m) \times 1} \text{ and } u_j^T u_j = 1 \text{ for } 1 \leq j \leq n \tag{6.26}$$

$$\bar{A} := [a_{ij}] \in \mathbb{R}^{n \times n} \text{ with } a_{ij} = -\frac{b_j b_j^T}{2\sigma_j} \text{ for } i = j,$$

$$\text{and } a_{ij} = \frac{\sigma_j b_i b_j^T - \sigma_i c_i^T c_j}{\sigma_i^2 - \sigma_j^2} \text{ for } i \neq j, \text{ with}$$

$$\sigma_{j+1} > \sigma_j > 0 \text{ for } 1 \leq j \leq n-1 \tag{6.27}$$

$$\bar{D} := [d_{ij}] \in \mathbb{R}^{(p+m) \times m}$$

which is a balanced state space representation having Hankel singular values σ_j.

Proof: Direct application of theorem 2.1 in Ober (1991) for a system with distinct Hankel singular values. □

The parametrization of the state space matrices $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ in Lemma 6.2-6 looks complicated compared to standard pseudo canonical (overlapping) parametrizations (Ljung 1987) but in fact it contains the same number of parameters and can be implemented relatively easy. Furthermore, (6.25), (6.26) and (6.27) reflect parameter constraints that can be dealt with relatively easy, while in a standard pseudo canonical (overlapping) parametrization more complicated parameter constraints would have been specified to guarantee stability of the state space realization. Unfortunately, (6.26) reflects a non-linear equality constraint in order to parametrize the unitary vectors u_j in the columns of $\bar{C}$.

To remove the non-linear equality constraint (6.26), in Chou (1994) a parametrization has been proposed, that parametrizes almost all unitary vectors in $\mathbb{R}^k$ without equality constraint and is based on rotating actions in $\mathbb{R}^k$. The parametrization is valid only for rotating actions in $\mathbb{R}^k$ with $k > 1$, but according to Lemma 6.2-6, $k = p + m > 1$ and hence it can be applied without objections. Additionally, it has been shown in de Callafon and Van den Hof (1995a) how to parametrize all unitary vectors $\in \mathbb{R}^k$ and the result is given below.

Lemma 6.2-7 *Let $\mathcal{U} := \{\bar{u} \in \mathbb{R}^{p+m} \mid \bar{u}^T \bar{u} = 1\}$, $\phi := [\phi_1 \ \phi_2 \ \cdots \phi_i \ \cdots \ \phi_{p+m-1}]^T \in$*

$\mathbb{R}^{p+m-1}$ and a map $f : \mathbb{R}^{p+m-1} \to \mathbb{R}^{p+m}$ be given by

$$u = u_i, \ \text{where} \ u_i := \prod_{i=1}^{p+m-1} \begin{bmatrix} \cos(\phi_i)u_{i-1} \\ \sin(\phi_i) \end{bmatrix}, \ \text{with} \ u_0 := 1 \qquad (6.28)$$

then $f(\phi) = u \in \mathcal{U}, \ \forall \ \phi \in \mathbb{R}^{p+m-1}$.

Proof: The fact that $u \in \mathcal{U}$ can be found by induction, see also Chou (1994): for $p + m - 1 = 1$ we have $u = u_1 = [\cos(\phi_1) \ \sin(\phi_1)]^T$ and hence $u_1^T u_1 = 1, \ \forall \ \phi_1 \in \mathbb{R}$. For $p + m - 1 = k - 1$ we assume $u_{k-1}^T u_{k-1} = 1$ and for $p + m - 1 = k$ we have $u = u_k^T u_k = \cos(\phi_k)u_{k-1}^T u_{k-1} \cos(\phi_k) + \sin^2(\phi_k) = 1, \ \forall \ \phi_k \in \mathbb{R}$. $\qquad \square$

Additionally it can be shown that the map (6.28) is surjective and hence parametrizes all unitary vectors in $\mathbb{R}^{p+m}$. This is an extension of the results in Chou (1994) and is given in the following lemma.

Lemma 6.2-8 Let the map $f : \mathbb{R}^{p+m-1} \to \mathbb{R}^{p+m}$ be given by (6.28) and let $\mathcal{U}$ be defined by $\mathcal{U} := \{\bar{u} \in \mathbb{R}^{p+m} \mid \bar{u}^T \bar{u} = 1\}$ then $\forall \ u \in \mathcal{U}$ there exists $\phi \in \mathbb{R}^{p+m-1}$ such that $f(\phi) = u$.

Proof: Take any $u := [u_1 \ u_2 \ \cdots \ u_i \cdots u_{p+m}]^T \in \mathbb{R}^{p+m}$ and define k to denote the index of the first non-zero entry u_k in u. Using the map (6.28) we can compute the elements ϕ_i of ϕ as follows:

$$\phi_i := \begin{cases} \dfrac{\pi}{2} & \text{for } 1 \leq i \leq k - 1 \\ \tan^{-1}\left(\dfrac{u_{i+1}}{u_k} \prod_{j=k}^{i-1} \cos(\phi_j) \right) & \text{for } k \leq i \leq p + m - 1 \end{cases} \qquad (6.29)$$

$$\phi_{p+m-1} := \phi_{p+m-1} + \pi \ \text{if} \ f(\phi) = -u \qquad (6.30)$$

where (6.30) is used only to adapt ϕ to the sign of u, if necessary. Clearly, $\phi \in \mathbb{R}^{p+m-1}$ and $f(\phi) = u$, which proves the result. $\qquad \square$

With the result mentioned in Lemma 6.2-8 it is proven that the parametrization of the unitary vectors $u_j \in \mathbb{R}^{p+m}$ in Lemma 6.2-6 with the non-linear equality constraint (6.26) can be replaced by the alternative parametrization in terms of $\phi \in \mathbb{R}^{p+m-1}$ given in Lemma 6.2-7, without any additional (equality) constraints. However, the map (6.28) is surjective so no unique parametrization of the unitary vectors is obtained. A unique parametrization, and hence a bijective map (6.28), can be enforced by putting the constraints

$$\phi_i \in \begin{cases} [-\pi/2, \pi/2] & \text{for } 1 \leq i \leq p + m - 2 \\ [-\pi/2, 3\pi/2) & \text{for } i = p + m - 1 \end{cases} \qquad (6.31)$$

on the elements ϕ_i of ϕ, similar as in Chou (1994). Although the constraints mentioned in (6.31) are needed to obtain a unique parametrization of the unitary vectors u_j, extensive simulations without the constraints (6.31) reveal promising results in de Callafon and Van den Hof (1995a).

To summarize, the parametrization of a rcf $(N(\theta), D(\theta))$ using a balanced state space realization can be formulated as follows. Let $G(\theta)$ be defined as

$$G(\theta) := \left[\begin{array}{c} N(\theta) \\ D(\theta) \end{array} \right]$$

having m inputs and $p + m$ outputs and parametrized according to the results mentioned in Lemma 6.2-6 and Lemma 6.2-7. In this way the parameter θ is defined as the vector

$$\theta = [\sigma_1 \ \sigma_2 \ \cdots \ \sigma_n \ b_1 \ b_2 \ \cdots \ b_n \ \phi_1^T \ \phi_2^T \ \cdots \ \phi_n^T] \in \mathbb{R}^{1 \times n(2m+p)}$$

with the additional constraints given in (6.25) and (6.27), which can be rewritten into

$$\sigma_n - \sigma_{n-1} > 0, \ \sigma_{n-1} - \sigma_{n-2} > 0, \ \cdots, \ \sigma_2 - \sigma_1 > 0, \ \sigma_1 > 0$$
$$b_{11} > 0, \ b_{21} > 0, \ \cdots, \ b_{n1} > 0$$

to ensure a minimal, stable and balanced continuous-time state space realization of the rcf $(N(\theta), D(\theta))$.

6.2.4   Frequency domain based identification

As indicated in Proposition 4.3-11, the access to a rcf $(N_{o,F}, D_{o,F})$ of the plant P_o can be accomplished by considering the map from x to $col(y, u)$, where x is found by the filter operation given in (4.24). Furthermore, it follows from Proposition 4.3-11 that the estimation of a rcf $(N(\theta), D(\theta))$ of a nominal model $P(\theta)$ is an open-loop equivalent identification problem, despite the fact that the data $col(y, u)$ might be gathered under closed-loop controlled conditions. As mentioned before, the rcf $(N_{o,F}, D_{o,F})$ is stable and the signals x and $col(y, u)$ are bounded. Consequently, time domain identification methods based on least squares prediction error methods (Ljung 1987) or time domain based subspace methods (Viberg 1994) can be readily implemented to estimate a stable rcf of the nominal model. The use of such time domain methods to estimate stable rcf's have been reported and applied for example in Zhu and Stoorvogel (1992) and Van den Hof $et\ al.$ (1995). Application of this fractional approach to experimental data coming from a compact disk player mechanism has been reported in de Callafon $et\ al.$ (1994) and Dötsch (1998). A similar time domain based system identification, but then based on the estimation of a lcf of the plant, has also been reported in Iglesias (1990).

Alternatively, frequency domain data(Fourier transformed time domain data or frequency domain measurements) can be used to identify a *rcf*. With such a frequency domain based identification, the nominal *rcf* is estimated on the basis of a finite number and possibly corrupted frequency response measurements. Frequency response samples can be obtained from time domain data which is frequently labelled as a non-parametric system identification, see also Section 3.2.2 and Remark 3.2-5. Such a non-parametric identification can be performed by spectral analysis (Priestley 1981), sine wave testing (Ljung 1987) or by the use of other special purposes periodic signals (Pintelon *et al.* 1994). Furthermore, special purpose hardware such as high speed spectral analyzers are available to obtain a finite number of frequency domain samples. Obviously, a nominal *rcf* still has to be estimated and the use of frequency domain data is regarded only as an intermediate step (Ljung 1993*b*). The frequency domain data can be used to estimate a parametric low complexity *rcf* of a nominal model, by means of curve fitting.

There is a variety of possibilities to motivate the use of frequency domain data for system identification purposes. For an overview of these motivations one can be referred to Pintelon *et al.* (1994), while similarities between the use of time- and frequency domain data using a least-squares prediction error framework is discussed in Ljung (1993*b*). The main reason to use a frequency domain based identification here, can be motivated by referring to the control relevant identification problem mentioned in (6.6) and written in terms of an additive difference between coprime factorizations in Lemma 6.2-1. From this it can be observed that it is preferable to solve a (weighted) $\mathcal{H}_\infty$ norm-based identification criterion.

The use of frequency domain data is helpful for the approximation of the $\mathcal{H}_\infty$ norm criterion mentioned in (6.8). This approximation is based on the fact that for a stable (discrete-time) transfer function matrix G, the $\mathcal{H}_\infty$ norm $\|G\|_\infty$ is given by

$$\|G\|_\infty = \max_{\omega \in [0,\pi]} \bar{\sigma}\{G(\mathrm{e}^{i\omega})\}$$

where $\bar{\sigma}\{\cdot\}$ denotes the maximum singular value. In case the transfer function matrix G is unknown, the value of $\|G\|_\infty$ can, at least, be approximated by computing

$$\max_{j=1,\ldots,l} \bar{\sigma}\{\hat{G}(\mathrm{e}^{i\omega_j})\}$$

where $\hat{G}(\mathrm{e}^{i\omega})$ denote the complex frequency domain samples of G along a prespecified frequency grid $\Omega = (\omega_1, \omega_2, \ldots, \omega_l)$. In case G is used to reflect the stable transfer function mentioned in the $\mathcal{H}_\infty$ norm of (6.8), the use of frequency domain data is beneficial in evaluating the $\mathcal{H}_\infty$ norm-based system identification.

Clearly, a finite number of frequency domain samples $\hat{G}(\mathrm{e}^{i\omega})$, $j = 1, \ldots, l$ that are possibly disturbed, does not uniquely define the underlying system G. Due to the availability of a finite number of frequency points in the grid Ω, additional information

on G has to be introduced to explain the behaviour of the frequency response between two subsequent frequency domain data points (Helmicki *et al.* 1989, Helmicki *et al.* 1991, Partington 1991, de Vries and Van den Hof 1995, Mäkilä *et al.* 1995).

A similar argumentation also holds when using time-domain data for system identification purposes. If a finite number of time-domain samples is available, additional information on the system G that covers the behaviour outside the observed time-domain interval must be introduced (Mäkilä and Partington 1992, Wahlberg and Ljung 1992). In either way, additional information is, in general, based on the assumption of G being analytic and to exhibit a certain "degree of stability". The use and introduction of such additional information will be postponed until Section 6.3, where these effects will be covered during the estimation of the model uncertainty. During the estimation of a nominal model, or nominal factorization, such additional information on the *rcf* $(N_{o,F}, D_{o,F})$ will not be used. Instead, it is assumed that the frequency grid is chosen dense enough.

Assumption 6.2-9 *It is assumed that for the discrete-time systems being considered here, the points ω_j in the frequency grid*

$$\Omega = (\omega_1, \ldots, \omega_j, \ldots, \omega_l), \;\; with \; 0 \leq \omega_1 < \cdots < \omega_j < \cdots < \omega_l \leq \pi \qquad (6.32)$$

are chosen in such a way that the (possibly disturbed) frequency domain samples are dense enough to represent the (continuous) frequency response of the rcf $(N_{o,F}, D_{o,F})$.

As mentioned in Schrama (1992*b*), the property whether or not the frequency domain samples are dense enough is determined by the order of the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o and the order of the *rcf* $(N(\theta), D(\theta))$ of the nominal model $P(\theta)$ to be estimated. Furthermore, as the value of an $\mathcal{H}_\infty$ norm is of importance during the parametric identification of a nominal *rcf*, the subsequent frequency points must be close enough to be able to evaluate and approximate $\max \bar{\sigma}\{\cdot\}$ over the specified frequency grid Ω.

The frequency domain data of the factorization $(N_{o,F}, D_{o,F})$ is denoted by $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$, where ω_j for $j = 1, 2, \ldots, l$ constitutes a (prespecified) frequency grid Ω that is assumed to satisfy the property mentioned in Assumption 6.2-9. Using this notation, the minimization of the H_∞-criterion (6.8) will be approximated by performing a pointwise evaluation of the maximum singular value over the frequency grid Ω.

Taking into account the restriction (6.9), the actual minimization problem using frequency domain data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ can be formalized via

$$\min_{\theta} \; \max_{j=1,2,\ldots,l} \; \bar{\sigma}\{G(\omega_j)\}, \; \text{with}$$

$$G(\omega_j) := U_2(\omega_j) \left(\begin{bmatrix} \hat{N}_{o,F}(\omega_j) \\ \hat{D}_{o,F}(\omega_j) \end{bmatrix} - \begin{bmatrix} N(\theta, \omega_j) \\ D(\theta, \omega_j) \end{bmatrix} \right) F(\omega_j) \begin{bmatrix} C_i(\omega_j) & I \end{bmatrix} U_1(\omega_j) \qquad (6.33)$$

For a fixed filter F, in general, (6.33) is non-convex min-max optimization problem that requires a sophisticated numerical optimization. Such a min-max optimization routine can be found in commercially available software (MatLab 1994). Furthermore, to ensure stability of $(N(\theta), D(\theta))$ during the optimization, the parametrization discussed in Section 6.2.3 can be used (de Callafon and Van den Hof 1995a). A similar min-max optimization problem to perform a curve fitting with a maximum amplitude criterion and a parametrization with guaranteed stability has been proposed in Hakvoort and Van den Hof (1994a). However, the optimization of (6.33) requires the minimization of a maximum singular value of a transfer function matrix G, which coincides with a maximum amplitude only if G in (6.33) is scalar. However, the procedure described in Hakvoort and Van den Hof (1994a) can be used to find an initial estimate for the non-linear optimization involved with (6.33). Alternatively, a straightforward least-squares optimization (de Callafon *et al.* 1996a) can be used. Provided that the resulting *rcf* being found by a least-squares optimization is stable, it can be represented in the parametrization given in Lemma 6.2-6 and used as an initial estimate for the optimization of (6.33).

6.3 Estimation of Model Uncertainty

6.3.1 Uncertainty modelling

In accordance with the separate estimation of nominal model and model uncertainty as discussed in Section 6.1.3, complementary to the nominal factorization $(\hat{N}, \hat{D})$ or corresponding nominal model $\hat{P} = \hat{N}\hat{D}^{-1}$ that has been estimated, a characterization of the model uncertainty is needed to complete the model uncertainty set. Referring to Definition 5.2-4, the set of models is structured according to

$$
\begin{aligned}
\mathcal{P}(\hat{N}, \hat{D}, N_c, D_c, \hat{V}, \hat{W}) := \{P \mid P = (\hat{N} + D_c\Delta_R)(\hat{D} - N_c\Delta_R)^{-1} \\
\text{with } \Delta_R \in R\mathcal{H}_\infty \text{ and } \Delta := \hat{V}\Delta_R\hat{W} \text{ satisfies } \|\Delta\|_\infty < \gamma^{-1}\}
\end{aligned}
\tag{6.34}
$$

where the model uncertainty is represented by a stable perturbation Δ_R in a dual-Youla parametrization. As mentioned in Section 5.2.4, an estimated set of models $\mathcal{P}$ or model uncertainty set that is characterized in this way, facilitates the use of the knowledge of the controller C that is implemented on the plant P_o.

Access to dual-Youla perturbation

The dual-Youla perturbation Δ_R is unknown and is used to represent the incorrectness of the nominal factorization $(\hat{N}, \hat{D})$ and the uncertainty inevitably present in the finite amount and possibly disturbed data used for the system identification procedure. In case Δ_R would be known exactly, the plant P_o would be known. Therefore, the aim is not to construct a parametric model for Δ_R, as the knowledge of the plant has

already been represented by a limited complexity model $(\hat{N}, \hat{D})$. Although estimating a (parametric) model $R(\theta)$ for Δ_R will increase the knowledge of the plant P_o, it will also increase the complexity of the resulting nominal factorization, which is an immediate consequence of using such a dual-Youla parametrization.

As such, the operator Δ_R is considered to be the "uncertainty" and it formulates an (allowable) model perturbation that yields the model uncertainty set $\mathcal{P}$ for which $P_o \in \mathcal{P}$ must hold. Therefore, both the term uncertainty and (allowable) model perturbation is used to portray Δ_R. Instead of finding a model for Δ_R in (6.34), frequency dependent weighting functions $\hat{V}$ and $\hat{W}$ are used to bound the (allowable) model perturbation Δ_R. The pair $(\hat{V}, \hat{W})$ is used to represent and bound the model perturbation Δ_R. Referring to (6.3), the bound on Δ_R must be chosen in such a way that

$$\min_{V,W} \sup_{P \in \mathcal{P}} \|J(P,C)\|_\infty \tag{6.35}$$

is being minimized and P_o is an element of the set of models $\mathcal{P}$ being constructed.

As already pointed out in Section 5.3.4, the substitution of $C = C_i$ in (5.12) yields $M_{11} = 0$. This results in an affine expression

$$M_{22} + M_{21}\Delta M_{12} \tag{6.36}$$

for $\mathcal{F}_u(M, \Delta)$, as mentioned in (6.36). In this affine expression, the matrices M_{22}, M_{21} and M_{12} are known quantities, while the size of the perturbation or uncertainty Δ must be chosen such that $P_o \in \mathcal{P}_i$ and $\|M_{22} + M_{21}\Delta M_{12}\|_\infty$ is minimized. Using system identification, the uncertainty Δ can be reduces and/or modified. As such, the structure of (6.36) can be exploited to formulate a (control relevant) identification problem to bound the uncertainty Δ or Δ_R in (6.34), by employing the knowledge of a controller C_i that is known to stabilize the nominal model $\hat{P}$ and the (unknown) plant P_o.

To accomplish the minimization of (6.35) and $P_o \in \mathcal{P}$, an intermediate step will be used that involves the estimation of a non-parametric (or high order) bound on Δ_R in (6.34). Subsequently, the knowledge of such a bound on the model perturbation Δ_R is used to construct the pair of stable and stably frequency dependent weighting functions $(\hat{V}, \hat{W})$ to complete the formulation of the set of models $\mathcal{P}$ as mentioned in Definition 5.2-4. Summarizing, the following two steps are performed to characterize the model perturbation.

- Given the nominal factorization $(\hat{N}, \hat{D})$, first a high order or non-parametric frequency dependent upper bound $\bar{\Delta}_R$ for Δ_R in (6.34) is estimated. Basically, such a frequency dependent upper bound $\bar{\Delta}_R$ provides the following information

$$|\Delta_R(e^{i\omega})| \le \bar{\Delta}_R(\omega) \text{ for } \omega \in [0, \pi]$$

  and the upper bound $\bar{\Delta}_R$ is used to set up the set of models so that the plant P_o is guaranteed to be an element of the set.

- The pair of frequency dependent stable and stably invertible weighting functions $(\hat{V}, \hat{W})$ are used to formulate a low complexity upper bound for the model perturbation Δ_R. The upper bound $\bar{\Delta}_R$ provides the frequency dependent information on which a pair $(\hat{V}, \hat{W})$ of limited complexity is fitted.

In performing the aforementioned steps, the minimization of (6.35) must be taken into account. Hence, the estimation of the high order or non-parametric upper bound $\bar{\Delta}_R$ must yield a tight upper bound for Δ_R in (6.34) in order to be able to minimize (6.35). This step actually constitutes the so-called uncertainty estimation modelling and is solved here by employing the procedure developed in Hakvoort and Van den Hof (1997). For the same reasoning, the pair of limited complexity weighting filters $(\hat{V}, \hat{W})$ must be computed such that (6.35) is being minimized. This problem will be solved by posing a curve fitting problem in which a frequency dependent upper bound is bounded from above by a stable and stably invertible weighting filter, similar as in Scheid *et al.* (1991). The aforementioned steps will be discussed in more detail in the following sections.

Given an estimate of a nominal factorization $(\hat{N}, \hat{D})$ that yield a model $\hat{P} = \hat{N}\hat{D}^{-1}$ that satisfies $T(\hat{P}, C_i) \in R\mathcal{H}_\infty$, the result of Lemma 4.3-2 can be used to access the signals needed to estimate an upper bound on the model perturbation Δ_R. Since such a factorization $(\hat{N}, \hat{D})$ has been obtained from the estimation of the control relevant identification discussed in Section 6.2.1, Lemma 4.3-2 can be applied readily to formulate the following result.

Corollary 6.3-1 *Consider data coming from a plant P_o operating in an internally stable feedback connection $T(P_o, C)$ as described by (3.2). Let the controller C have a rcf (N_c, D_c) and consider a rcf $(\hat{N}, \hat{D})$ of a nominal model $\hat{P}$ that satisfies $T(\hat{P}, C) \in R\mathcal{H}_\infty$. Then (3.2) can be rewritten into*

$$z = \Delta_R x + S_o e = \Delta_R x + v \tag{6.37}$$

and the signals x and z are given by the filter operation

$$
\begin{aligned}
x &:= (\hat{D} + C\hat{N})^{-1} \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \\
z &:= (D_c + \hat{P}N_c)^{-1} \begin{bmatrix} I & -\hat{P} \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix}
\end{aligned}
\tag{6.38}
$$

where x is uncorrelated with v. Subsequently Δ_R and S_o are given by

$$
\begin{aligned}
\Delta_R &= D_c^{-1}(I + P_o C)^{-1}(P_o - \hat{P})\hat{D} \\
S_o &= D_c^{-1}(I + P_o C)^{-1}H_o
\end{aligned}
\tag{6.39}
$$

where Δ_R and S_o are stable LTI operators.

Proof: Application of Lemma 4.3-2 for $(N_x, D_x) = (\hat{N}, \hat{D})$. $\square$

With the result mentioned in Corollary 6.3-1, it can be seen that the signal x and the signal z, found by a simple filtering of the signals present in the feedback connection $\mathcal{T}(P_o, C)$, can be used to access the model perturbation Δ_R. Furthermore, under the condition that $T(\hat{P}, C) \in R\mathcal{H}_\infty$, the model perturbation Δ_R is a stable LTI operator.

The need for closed-loop experiments

Similar to the concluding remarks mentioned at the end of Section 6.2.1, it can be mentioned here that the presence of the (unknown) closed-loop transfer function $(I + P_oC)^{-1}$ in (6.39) indicates the need for performing closed-loop experiments. As the plant P_o is assumed to be unknown, the output sensitivity function $S_{out} = (I + P_oC)^{-1}$ that appears in (6.39) cannot be reconstructed and closed-loop experiments are needed in acquiring information on this closed-loop transfer function.

Hence, for the estimation of the model uncertainty that has been structured in a set of models that is tuned towards the closed-loop application of the models within the set, closed-loop experiments are beneficial. Similar to the estimation of a feedback relevant nominal model, it can be noted that if the controller C applied to the plant P_o is updated from C_i to C_{i+1}, the output sensitivity S_{out} changes. Assuming that full knowledge of the plant P_o, or similarly the output sensitivity $(I + P_oC_i)^{-1}$ is not available, renewed information of the output sensitivity $(I + P_oC_{i+1})^{-1}$ or the dual-Youla perturbation Δ_R in (6.39) can be obtained by performing new closed-loop experiments.

6.3.2   Estimating an upper bound for the model uncertainty

Stability of Δ_R facilitates the use of many newly developed system identification techniques that are specialized in estimating and characterizing a model uncertainty. The main reason for the preference to deal with a stable system during an uncertainty bounding identification is due to the additional information or *prior* information needed to formulate a reliable uncertainty bound. Usually, such *prior* information is based on the assumption of the system being analytic and to exhibit a certain "degree of stability" (Helmicki *et al.* 1989, Helmicki *et al.* 1991, Partington 1991, de Vries and Van den Hof 1995, Mäkilä *et al.* 1995). Clearly, this requires the system to be stable in order not to violate this *prior* assumption.

Introduction of prior information

The aim of this thesis is not to develop a new methodology for estimating a model uncertainty. Instead, the procedure of probabilistic uncertainty bounding identification, as described in Hakvoort (1994) or Hakvoort and Van den Hof (1997) is being

used. The reason for choosing this procedure is the flexibility with which the model uncertainty estimation can be estimated and the application of the procedure to multivariable systems.

Another motivation for choosing the procedure of probabilistic uncertainty bounding identification of Hakvoort and Van den Hof (1997) is related to the assumptions that need to be made on the noise present on the data. For the uncertainty bounding, assumptions on the noise $v(t) = S_o(q)e(t)$ in (6.37) must be made. In Hakvoort (1994) both deterministic assumptions such as

$$|v(t)| \leq \bar{v}(t) \text{ for } t = 1, \ldots, N \tag{6.40}$$

or stochastic assumptions in accordance with the prediction error framework of Ljung (1987) can be handled. The noise assumption mentioned in (6.40) is frequently used in so-called unknown, but bounded noise parameter bounding or set membership identification (Milanese and Vicino 1991). Such deterministic assumptions lead to deterministic or "hard" error bounds for the uncertainty modelling. The resulting set of models is definitely guaranteed to contain the unknown plant P_o, provided, of course, that the *a priory* information of (6.40) is correct.

Unfortunately, assumptions as (6.40) tend to trivialize the properties of the noise. No averaging properties of the noise are incorporated in (6.40) and the noise $v(t)$ might be completely correlated with the input signal $x(t)$ (Hjalmarsson 1993). Furthermore, the noise bounds $\bar{v}(t)$ may have to be chosen highly conservative to account for outliers in the data or to comprehend the possible stochastic behaviour of the noise. In Hakvoort and Van den Hof (1997) these problems are addressed by considering alternative deterministic noise assumptions. These alternatives include cross-covariance constraints, where a deterministic bound on the correlation between the noise $v(t)$ and the input signal $x(t)$ can be posed as *prior* information.

The procedure works on the basis of time-domain data and typically, the following *prior* assumptions are being made.

- The system Δ_R is assumed to be stable and to exhibit a certain degree of stability. In Hakvoort and Van den Hof (1997) the degree of stability is imposed by first considering $\Delta_R(q)$ to exhibit the series expansion

$$\Delta_R(q) = \sum_{k=0}^{\infty} R_k B_k(q) \tag{6.41}$$

  where $B_k(q)$ for $k = 0, \ldots, \infty$ is some user-defined set of (orthonormal) basis functions. These basis functions can be chosen as a common shift operator $B_k(q) = q^{-k}$ or based upon a Laguerre functions (Wahlberg 1991), Kautz functions (Wahlberg 1994) or the more generalized concept of orthonormal basis functions as developed in Heuberger *et al.* (1995), Ninness and Gómez (1995) and discussed before in Section 3.2.2.

Subsequently, the assumption is made that the unknown coefficients R_k in (6.41) can be bounded by

$$|R_k| \leq \bar{R}_k, \text{ for } k = 0, \ldots, \infty \tag{6.42}$$

where $\{\bar{R}_k\}_{k=0,\ldots,\infty}$ is a (unknown) sequence. The *prior* information now consists of a degree of stability, enforced by assuming that the sequence $\{\bar{R}_k\}_{k=0,\ldots,\infty}$ shows an exponentially decay for k larger than some $k^\star$

$$\bar{R}_k \leq M\rho^k, \ \forall k > k^\star \tag{6.43}$$

for some given $M \geq 0$ and $\rho < 1$.

- The effect of unknown initial conditions, that has influenced the data gathered over a finite time interval, must also be bounded. This effect is handled by assuming that the input signal x in (6.37) is said to satisfy

$$|x(t)| \leq \bar{x}, \ \forall t \leq 0 \tag{6.44}$$

where $t \leq 0$ is used to indicate the time span before the data was captured.

Finally, an assumption on the noise $v(t) = S_o(q)e(t)$ in (6.37) must be made. As mentioned above, stochastic assumptions are used in the uncertainty bounding identification. Consistent with the framework as presented in Ljung (1987) and mentioned previously in Assumption 3.2-2, it is assumed that the *prior* information on the noise $v(t) = S_o(q)e(t)$ in (6.37) satisfies the following assumption, see also Hakvoort and Van den Hof (1997).

Assumption 6.3-2 *The noise $\{v(t)\}$ in (6.37) satisfies $v(t) = S_o(q)e(t)$ where $S_o(q)$ is stable and $\{e(t)\}$ is a sequence of independent random variable with zero means, variances λ, bounded fourth moments and uncorrelated with the input signal $\{x(t)\}$.*

It should be noted that a stochastic assumption on the noise typically will yield a bound $\bar{\Delta}_R$ on the uncertainty Δ_R that holds with a prespecified probability. In Hakvoort and Van den Hof (1997) a procedure for such a probabilistic uncertainty bounding identification is presented and is used in this thesis. As a result, a set of models $\mathcal{P}$ based on the uncertainty bounding identification is said to satisfy

$$P_o \in \mathcal{P}, \ w.p. \geq \alpha < 1$$

provided that the *prior* stochastic assumption on the noise is correct. Specifying a probability α close to 1, will lead to a set of models $\mathcal{P}$ for which $P_o \in \mathcal{P}$ is bound to hold with a high probability. Consequently, a robust controller designed on the basis of the identified set of models, as mentioned in Procedure 2.5-4, will inherit robustness properties that also hold with a certain (high) probability.

With such a probabilistic approach, the property $P_o \in \mathcal{P}$ and the closed-loop validation tests as mentioned in Procedure 2.5-4 in (2.20) and (2.21) can not be evaluated with absolute guarantees. Consequently, robustness properties for a controller C such as mentioned in Definition 2.2-18 and Definition 2.2-19 are said to hold with a certain probability. However, it is believed that this approach is realistic and consistent with many practical applications where robustness and failure can not be assured with absolute guarantees, unless a strongly conservative policy is pursued during the design of the feedback control system.

Computation of probabilistic uncertainty bounds

The objective is to derive a frequency dependent probabilistic uncertainty bound $\bar{\Delta}_R(\omega)$ for the frequency response of Δ_R. As mentioned before, the (probabilistic) uncertainty bound $\bar{\Delta}_R$ is used as a first step to find a tight, non-parametric (or high order) upper bound for the model perturbation Δ_R in (6.34).

Such an estimation of a non-parametric upper bound $\bar{\Delta}_R$ gives rise to an intermediate set of models, denoted by $\mathcal{S}$ in Hakvoort (1994), and is defined as follows.

$$\mathcal{S} := \left\{ \Delta_R(z) \,\middle|\, \Delta_R(e^{i\omega_j}) \in \mathcal{P}(\omega_j),\ \forall \omega_j \in \Omega,\ \left| \frac{\partial \Delta_R(e^{i\omega})}{\partial \omega} \right| \leq \beta,\ \forall \omega \in [0, \pi] \right\} \quad (6.45)$$

The intermediate set of models $\mathcal{S}$ in (6.45) is characterized by convex frequency (uncertainty) response regions, denoted by $\mathcal{P}(\omega_j)$, that are defined in the complex plane and given for a finite number of frequency points in a prespecified frequency grid Ω.

Remark 6.3-3 *Although the formulation of the intermediate set $\mathcal{S}$ in (6.45) implies that $\Delta_R(z)$ is a scalar function, the (probabilistic) uncertainty bounding identification routine mentioned in Hakvoort (1994) is applicable to a multivariable perturbation $\Delta_R(z)$. In that case, convex frequency response regions $\mathcal{P}_{ij}(\omega_j)$ are derived for each element (i, j) of the multivariable perturbation $\Delta_R(z)$.*

In order not to be overwhelmed by the notational issues involved with keeping track off each element (i, j) of a multivariable perturbation $\Delta_R(z)$ and the corresponding convex frequency response region $\mathcal{P}_{ij}(\omega_j)$ most of the result below are written down for a scalar $\Delta_R(z)$. The corresponding convex frequency response regions $\mathcal{P}(\omega_j)$ can be found by formulating convex polytopes that bound, for example, the real and imaginary part of $\Delta_R(e^{i\omega_j})$. Along with a bound on the first derivative β of the frequency response of Δ_R that can be obtained by employing the *prior* assumption on the degree of stability of Δ_R, a uniform bound or frequency dependent bound $\bar{\Delta}_R(\omega)$ on the frequency response of Δ_R can be computed. As mentioned before, this frequency dependent bound $\bar{\Delta}_R(\omega)$ is used to construct a pair of stable and stably invertible weighting filters $(\hat{V}, \hat{W})$ that bound the model perturbation Δ_R.

Clearly, the convex frequency response regions $\mathcal{P}(\omega_j)$ and the bound β on the first derivative need to be computed to construct $\mathcal{S}$. It is beyond the scope of this thesis

how to derive $\mathcal{P}(\omega_j)$ and β. For a detailed discussion, one is referred to Hakvoort (1994). In this thesis only the main results are presented and can be find below.

Lemma 6.3-4 *Let $\Delta_R(q)$ be any stable transfer function satisfying (6.41), where the coefficients R_k are bounded by $\bar{R}_k$ as in (6.42). Let β be given by*

$$\beta := \sum_{k=0}^{\infty} \bar{R}_k \left\| \frac{\partial B_k(z)}{\partial z} \right\|_{\infty}$$

then

$$\left| \frac{\partial \Delta_R(\mathrm{e}^{i\omega})}{\partial \omega} \right| \leq \beta, \ \forall \omega \in [0, \pi]$$

Proof: For a proof, one is referred to Hakvoort (1994, pp. 104). $\qquad\qquad\square$

The result mentioned in Lemma 6.3-4 links the bound $\bar{R}_k$ on the (unknown) generalized impulse coefficients R_k of $\Delta_R(q)$ with a bound on the first derivative of the frequency response of Δ_R. With the *prior* information M and ρ in (6.43) that specify the exponential decay rate of the sequence $\{\bar{R}_k\}_{k=0,\dots,\infty}$, the bound β in Lemma 6.3-4 can be computed (Hakvoort 1994).

To characterize the convex frequency uncertainty regions $\mathcal{P}(\omega_j)$ using a probabilistic uncertainty estimation routine, first a parametric model $\hat{\Delta}_R$ is estimated. The parametric model $\hat{\Delta}_R$ is used solely to specify a complex central orientation point of $\mathcal{P}(\omega_j)$. In this perspective, the model $\hat{\Delta}_R$ can be considered as the "carrier" of the convex frequency uncertainty regions $\mathcal{P}(\omega_j)$.

With Δ_R expressed in the orthonormal expansion (6.41), it is advantageous to exploit this linear regression structure. As discussed in Section 3.2.2, a model $\hat{\Delta}_R$ parametrized via an ORTFIR (Heuberger *et al.* 1995) structure yields such a linearly parametrized model structure. Writing the model $\hat{\Delta}_R$ in the ORTFIR expansion with n elements

$$\hat{\Delta}_R(q, \hat{r}) = \sum_{k=0}^{n} \hat{r}_k B_k(q) \tag{6.46}$$

the coefficients $\{\hat{r}_k\}_{k=1,\dots,n}$ are used to denote the generalized pulse coefficients that have been obtained on the basis of a convex least-squares minimization

$$\hat{r} = \arg \ \min_{r} \|\varepsilon(t, r)\|_2 \tag{6.47}$$

where

$$\varepsilon(t, r) = z(t) - \sum_{k=0}^{n} r_k B_k(q) x(t)$$

denotes the output error and x and z are respectively the input and output data as given in (6.37).

The solution to the convex least-squares optimization of (6.47) can be characterized by rewriting $\hat{\Delta}_R(q,r)x(t)$ into

$$\hat{\Delta}_R(q,r)x(t) = \sum_{k=0}^{n} r_k B_k(q)x(t) = \sum_{k=0}^{n} r_k x_k(t)$$

with

$$x_k(t) = \sum_{h=0}^{h-1} b_{k,h} q^{-h} x(t)$$

and where $b_{k,h}$ denotes the h-th Markov parameter of the k-th basis function $B_k(q)$. Hence, $x_k(t)$ denotes the FIR-filtered signal $x(t)$ using the FIR response filter based on the k-th basis function $B_k(q)$. Stacking the signals $x_k(t)$ in the vector representation

$$X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_k(t) \end{bmatrix} \tag{6.48}$$

the solution $\hat{r}$ of the optimization (6.47) is found via a standard linear regression problem (Ljung 1987)

$$\begin{bmatrix} \hat{r}_1 \\ \vdots \\ \hat{r}_k \end{bmatrix} = \left[\frac{1}{N} \sum_{t=1}^{N} X(t) X^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^{N} X(t) z(t) \tag{6.49}$$

where $X(t)$ is defined as in (6.48).

Before presenting the result on the computation of the probabilistic frequency domain uncertainty regions $\mathcal{P}(\omega_j)$, for notational convenience the following expressions are introduced. First of all, using the n basis functions $B_k(q)_{k=1,\ldots,n}$ of the model $\hat{\Delta}_R$ in (6.46) and the expression for $X(t)$ in (6.48) the following two signal $\xi_{re}(t)$ and $\xi_{im}(t)$ are defined.

$$\xi_{re}(t) = [\mathrm{Re}\{B_0(e^{i\omega_j})\} \;\cdots\; \mathrm{Re}\{B_n(e^{i\omega_j})\}] \left[\frac{1}{N} \sum_{t=1}^{N} X(t) X^T(t) \right]^{-1} X(t) \tag{6.50}$$

$$\xi_{im}(t) = [\mathrm{Im}\{B_0(e^{i\omega_j})\} \;\cdots\; \mathrm{Im}\{B_n(e^{i\omega_j})\}] \left[\frac{1}{N} \sum_{t=1}^{N} X(t) X^T(t) \right]^{-1} X(t) \tag{6.51}$$

The above mentioned signals are simply filtered version of the signal $x(t)$. However, they play a crucial role in the notation and characterization of the uncertainty bounding, as mentioned below.

A computable bound for the effect of unknown initial conditions and the fact that only n basis functions have been used in the model $\hat{\Delta}_R(q, \hat{r})$ of (6.46) can be characterized by means of the signals $\xi_{re}(t)$ and $\xi_{im}(t)$ (Hakvoort 1994). For notational convenience, the bounds that includes the effect of initial conditions is written as follows

$$
\begin{aligned}
f_{re} &= \sum_{k=0}^{\infty} \bar{R}_k \sum_{h=0}^{\infty} \left| \sum_{t=1}^{N} b_{k,h+t} \xi_{re}(t) \right| \bar{x} \\
f_{im} &= \sum_{k=0}^{\infty} \bar{R}_k \sum_{h=0}^{\infty} \left| \sum_{t=1}^{N} b_{k,h+t} \xi_{re}(t) \right| \bar{x}
\end{aligned}
\tag{6.52}
$$

where $\bar{x}$ was given in (6.44). The bound that incorporates the tail contribution of the ORTFIR expansion of the model $\hat{\Delta}_R(q, \hat{r})$ is denoted by

$$
\begin{aligned}
d_{re} &= \sum_{k=n+1}^{\infty} \bar{R}_k \left| \sum_{t=1}^{N} \xi_{re}(t) \sum_{h=0}^{t-1} b_{k,h} q^{-h} x(t) \right| \\
d_{im} &= \sum_{k=n+1}^{\infty} \bar{R}_k \left| \sum_{t=1}^{N} \xi_{im}(t) \sum_{h=0}^{t-1} b_{k,h} q^{-h} x(t) \right|
\end{aligned}
\tag{6.53}
$$

where $b_{k,h}$ is again used to denote the h-th Markov parameter of the k-th basis function $B_k(q)$ in (6.41). Similar to the bound given in (6.53), the real and imaginary part of the frequency response of the tail contribution of the ORTFIR expansion of $\Delta_R(q)$ needs to be bounded. Following Hakvoort (1994), these bounds are given by

$$
\begin{aligned}
\delta_{re}(\omega_j) &= \sum_{k=n+1}^{\infty} \bar{R}_k |\mathrm{Re}\{B_k(\mathrm{e}^{\,i\omega_j})\}| \\
\delta_{im}(\omega_j) &= \sum_{k=n+1}^{\infty} \bar{R}_k |\mathrm{Im}\{B_k(\mathrm{e}^{\,i\omega_j})\}|
\end{aligned}
\tag{6.54}
$$

It should be noted that the computation of the expressions in (6.52), (6.53) and (6.54) involves the evaluation of infinite sums. However, the assumption on the exponential decay rate of the sequence $\bar{R}_k$ for $k = k^{\star}, \ldots, \infty$, as mentioned in (6.43), make the bounds computable (Hakvoort 1994).

To conclude the enumeration of notations associated to the signals mentioned in (6.50) and (6.51), the notation for the cross-correlation between $\xi_{re}(t)$, $\xi_{im}(t)$ and the noise $v(t)$ in (6.37) is defined as follows.

$$
\Lambda_{\xi_{re}\xi_{im}}^{N} := \frac{1}{N} \sum_{t=1}^{N} \begin{bmatrix} \xi_{re}(t) \\ \xi_{im}(t) \end{bmatrix} v(t) \sum_{t=1}^{N} \begin{bmatrix} \xi_{re}(t) & \xi_{im}(t) \end{bmatrix} v(t)
\tag{6.55}
$$

Furthermore, for notational convenience, $\Lambda^N_{\xi_{re}\xi_{im}}$ is split up in

$$\Lambda^N_{\xi_{re}\xi_{im}} = \begin{bmatrix} \lambda^N_{\xi_{re}\xi_{re}} & \times \\ \times & \lambda^N_{\xi_{im}\xi_{im}} \end{bmatrix} \tag{6.56}$$

where the upper left matrix-element of the matrix $\Lambda^N_{\xi_{re}\xi_{im}}$ associated to ξ_{re} is denoted by $\lambda^N_{\xi_{re}\xi_{re}}$ and the lower right matrix-element associated to ξ_{im} is denoted by $\lambda^N_{\xi_{re}\xi_{re}}$. Having set up the above mentioned notations, the following main result can be formulated for the frequency response confidence regions.

Lemma 6.3-5 *Consider the ORTFIR estimate $\hat{\Delta}_R(q, \hat{r})$ given in (6.46) and (6.49) and suppose $\{v(t)\}$ and $\{x(t)\}$ in (6.37) satisfy Assumption 6.3-2. Let the computable bounds f_{re} and f_{im} be given by (6.52), d_{re} and d_{im} be given by (6.53), $\delta_{re}(\omega_j)$ and $\delta_{im}(\omega_j)$ be given by (6.54). Furthermore, let the covariance information $\lambda^N_{\xi_{re}\xi_{re}}$ and $\lambda^N_{\xi_{im}\xi_{im}}$ be given by (6.55) and (6.56). Consider a positive real number $\alpha < 1$, then for $N \to \infty$ the following expressions hold*

$$\left| Re\{\hat{\Delta}_R(e^{i\omega_j}, \hat{r}) - \Delta_R(e^{i\omega_j})\} \right| \leq \mathcal{R}_\alpha(\omega_j), \ w.p. \ \geq \alpha$$
$$\left| Im\{\hat{\Delta}_R(e^{i\omega_j}, \hat{r}) - \Delta_R(e^{i\omega_j})\} \right| \leq \mathcal{I}_\alpha(\omega_j), \ w.p. \ \geq \alpha \tag{6.57}$$

where $\mathcal{R}_\alpha(\omega_j)$ and $\mathcal{I}_\alpha(\omega_j)$ are given by

$$\mathcal{R}_\alpha(\omega_j) = c_{\mathcal{N},\alpha}\sqrt{\frac{\lambda^N_{\xi_{re}\xi_{re}}}{N}} + \frac{d_{re}}{N} + \frac{f_{re}}{N} + \delta_{re}(\omega_j)$$
$$\mathcal{I}_\alpha(\omega_j) = c_{\mathcal{N},\alpha}\sqrt{\frac{\lambda^N_{\xi_{im}\xi_{im}}}{N}} + \frac{d_{im}}{N} + \frac{f_{im}}{N} + \delta_{im}(\omega_j) \tag{6.58}$$

and $c_{\mathcal{N},\alpha}$ denotes a real number that corresponds to the probability α in the standard Normal distribution.

Proof: For a proof one is referred to the work by Hakvoort (1994)   □

The result in Lemma 6.3-5 provides the possibility to formulate a probabilistic convex frequency response regions $\mathcal{P}(\omega_j)$ formed by the bounds for the real and imaginary parts mentioned in (6.57) and (6.58). A typical shape of a (probabilistic) box-shaped frequency response region $\mathcal{P}(\omega_j)$ for one frequency ω_j out of the frequency grid Ω is depicted in Figure 6.1. As can be seen from this figure, a rectangular shaped confidence region in the complex plane is obtained around the complex frequency response value $\hat{\Delta}_R(e^{i\omega_j}, \hat{r})$. As can be seen from (6.57), the different sources that contribute to the size of probabilistic convex frequency response regions can be distinguished and possibly individually influenced.
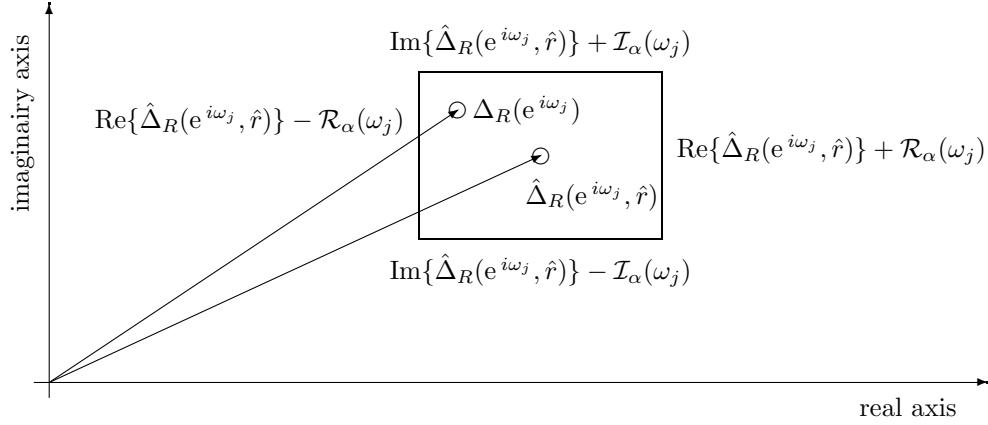
Fig. 6.1: Rectangular (probabilistic) frequency response region for one frequency ω_j out of the frequency grid Ω.

The result of Lemma 6.3-4 and Lemma 6.3-5 conclude the characterisation of the intermediate set of models $\mathcal{S}$. Consequently, a set of model $\mathcal{S}$ is obtained for which $P_o \in \mathcal{S}$ is known to hold with a certain (high) probability, provided that the *prior* information is correct. Once the bound β on the first derivative of the frequency response and the convex frequency response regions $\mathcal{P}(\omega_j)$ are available, a frequency dependent upper bound $\bar{\Delta}_R(\omega)$ for the model perturbation Δ_R can be computed (Hakvoort 1994).

6.3.3   Parametric approximation of upper bound

The availability of a (non-parametric) frequency dependent upper bound $\bar{\Delta}_R(\omega)$ of $\Delta_R(e^{i\omega})$ can be used to find stable and stable invertible weighting filter $(\hat{V}, \hat{W})$ of limited complexity that form a tight overbound

$$\|\hat{V}\bar{\Delta}_R\hat{W}\|_\infty \leq \gamma^{-1} \tag{6.59}$$

for $\bar{\Delta}_R(\omega)$. It should be noted that in case of a multivariable perturbation Δ_R, the (probabilistic) uncertainty bounding identification summarized above yields a frequency dependent upper bound $\bar{\Delta}_{Rij}(\omega)$ for each element (i,j) of the perturbation Δ_R. It was already observed in Section 5.2.5 that such detailed information can be used to find scalar stable and stably invertible weighting filters $\hat{V}_{ij}$ that bound each element (i,j) of $\bar{\Delta}_R$

$$\|\hat{V}_{ij}\bar{\Delta}_{Rij}\|_\infty \leq \gamma^{-1} \tag{6.60}$$

as previously mentioned in (5.8).

While estimating limited complexity weighting filters $(\hat{V}, \hat{W})$ or $\hat{V}_{ij}$ respectively in (6.59) or (6.60), the objective mentioned in (6.35) must still be taken into account. Clearly, the possibility to construct high order weighting filters will allow a tight overbound of $\bar{\Delta}_R$. However, the weighting filters reappear in the lower LFT $\mathcal{F}_l(G, C)$ given in Corollary 5.4-1. Since $\mathcal{F}_l(G, C)$ is used in the design of a robust controller, it is beneficial to limit the complexity of the weighting filters, similar to the motivation to limit the complexity of the nominal model or nominal factorization $(\hat{N}, \hat{D})$.

Estimating a scalar weighting function that will form a tight (weighted) upper bound of $\bar{\Delta}_{Rij}$ in (6.60) has for example been studied in Scheid *et al.* (1991). The linear programming spectral overbounding and factorization (LPSOF) algorithm presented in Scheid *et al.* (1991) can be used to find low complexity stable and stably invertible weighting filters that overbound the frequency dependent information $\bar{\Delta}_R$. For reasons of completeness, the LPSOF is summarized below.

The concept that $\hat{V}_{ij}$ must be a weighted tight upper bound can be formalized by posing the optimization

$$\min_{\theta} \kappa(\theta)$$

with

$$\kappa(\theta) := \max_{\omega_j \in \Omega} \{(V_{ij}(e^{i\omega_j}, \theta) - \bar{\Delta}_R(\omega_j))Q(\omega_j)\} \tag{6.61}$$

where Ω is a frequency grid as given in (6.32) and $Q(\omega)$ defines some user specified weighting.

In case the stable and stably invertible weighting filter $V_{ij}(e^{i\omega}, \theta)$ has a McMillan degree n, evaluating $V_{ij}(e^{i\omega}, \theta)^{\star}V_{ij}(e^{i\omega}, \theta)$ will lead to an expression of the form

$$V_{ij}(e^{i\omega}, \theta)^{\star}V_{ij}(e^{i\omega}, \theta) = \frac{b(\omega_j, \theta)}{a(\omega_j, \theta)} \tag{6.62}$$

where

$$\begin{aligned}
b(\omega_j, \theta) &= b_0 + b_1 \cos(\omega_j) + \cdots + b_n \cos(n\omega_j) \\
a(\omega_j, \theta) &= 1 + a_1 \cos(\omega_j) + \cdots + a_n \cos(n\omega_j)
\end{aligned}$$

and $[b_0 \ \cdots \ b_n \ a_1 \ \cdots \ a_n]$ corresponds to the unknown parameter θ in the weighting filter $V_{ij}(e^{i\omega}, \theta)$ to be estimated. Hence, the condition that $V_{ij}(e^{i\omega}, \theta)$ should overbound $\bar{\Delta}_R(\omega_j)$ can be restated as

$$\frac{b(\omega_j, \theta)}{a(\omega_j, \theta)} \geq \bar{\Delta}_R(\omega_j)^2, \ \forall \omega_j \in \Omega. \tag{6.63}$$

The condition that the overbound $b(\omega_j, \theta)/a(\omega_j, \theta)$ admits a spectral factorization and can be represented by a stable and stably invertible weighting filter $V_{ij}(e^{i\omega}, \theta)$ can be guaranteed by the additional condition

$$a(\omega_j, \theta) > 0 \ \forall \omega_j \in \Omega. \tag{6.64}$$

Using (6.62) and employing an equation error form

$$\delta(\theta) := \max_{\omega_j \in \Omega} \{(b(\omega_j, \theta) - \bar{\Delta}_R(\omega_j)^2 a(\omega_j, \theta))Q(\omega_j)\} \tag{6.65}$$

it has been recognized in Scheid *et al.* (1991) that the optimization

$$\min_{\delta, \theta} \delta(\theta)$$

subjected to the constraints mentioned in (6.63) and (6.64) is a linear programming problem for a fixed δ. In this linear programming problem, a feasible solution θ must be found. By systematically lowering the value of δ, a (weighted) tight overbound of $\bar{\Delta}_R(\omega_j)$ can be found. This LPSOF algorithm is also used in this thesis to find the low complexity parametric stable and stably invertible weighting filters, that bound the perturbation Δ_R element wise.

6.4   Summary on the Model Uncertainty Set Estimation

With the tools described in this chapter, the issue of estimating a set of models or model uncertainty set has been addressed. It has been made clear that the set of models, as given in (6.34) depends on the nominal factorization $(\hat{N}, \hat{D})$ and the weighting functions $(\hat{V}, \hat{W})$ that bound the model uncertainty. During the estimation and construction of the set of models, the closed-loop criterion

$$\min_{N, D, V, W} \sup_{P \in \mathcal{P}} \|J(P, C)\|_\infty \tag{6.66}$$

has been taken into account to find a model uncertainty set, suitable for enhanced robust control design.

As a minimization, using the variables (N, D, V, W) simultaneously is intractable, a separate control relevant identification of a nominal factorization and a model uncertainty is proposed for the estimation of the set of models. The separate estimation estimation involves the following two steps.

- The estimation of a nominal factorization $(\hat{N}, \hat{D})$ such that (6.66) is being minimized using the *rcf* (N, D) only, subjected to internal stability of $\mathcal{T}(\hat{P}, C)$. The estimation of $(\hat{N}, \hat{D})$ is done with the optimization posed in (6.33) and involves a non-linear optimization for which frequency domain measurements of the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o is being used.

- The estimation of the model uncertainty consists of the characterization of a frequency dependent upper bound $\bar{\Delta}_R(\omega)$ on Δ_R in (6.34) via $(\hat{V}, \hat{W})$ such that (6.66) is being minimized using (V, W) only, subjected to $P_o \in \mathcal{P}$. This frequency dependent upper bound $\bar{\Delta}_R(\omega)$ is found by the probabilistic uncertainty

bounding identification of Hakvoort and Van den Hof (1997) that is applied to the set of models as described in (6.34) on the basis of closed-loop time domain observations. Subsequently, stable and stably invertible weighting filters that bound $\bar{\Delta}_R(\omega)$ are found with the LPSOF algorithm of Scheid *et al.* (1991).

The model uncertainty set estimation forms an integral part of the identification steps in Procedure 2.5-4. In the next chapter, the main ingredients of Procedure 2.5-4, as presented in this thesis, will be outlined to indicate the applicability of the procedure.

Part IV

Application

7

Application of the Model-Based Procedure

7.1 Introduction

With the contents of the previous chapters, the ingredients and results to perform the model-based procedure mentioned in Section 2.5.2 have been presented. Most of these results have been outlined in chapter 5 and chapter 6 and merge the results available in the fields of system identification and robust controller design to constitute a framework for a model-based approach to the design of robust and enhanced performing controllers for a unknown plant, on the basis of observations coming from the plant. For reasons of clarity, a short summary of these results is outlined in this chapter. The different results and ingredients have to be merged to complete the model-based procedure so as to address the problem of designing an improved controller for an unknown, possibly unstable, system.

Firstly in Section 7.2, the main ingredients, as mentioned in Section 2.5.3, are summarized. These ingredients include the control objective function being chosen, the way in which the set of models is being structured, the procedure for estimating the model uncertainty set and the robust control design methodology used to develop enhanced performing robust controllers on the basis of the estimated model uncertainty set.

Subsequently, in Section 7.3 the flow chart of the model-based procedure mentioned in Figure 2.7 is repeated. This flow chart is used to refer to the different contributions developed and merged in this thesis. In this way, the path is outlined in which the estimation of the model uncertainty set, the robust controller design and the closed-loop validation test are organized to progressively find an improved performing robust controller.

The next chapter is devoted to the practical application of the scheme. This application involves the identification of a set of models and the design and implementation of a feedback controller for a positioning mechanism present in a wafer stepper.

7.2   Main Ingredients

Following the enumeration mentioned in Section 2.5.3, the following summary of the main ingredients in the model-based procedure can be given.

The control objective function

As indicated in Definition 2.2-13 and Remark 2.2-15, the control objective function $J(P,C)$ is used to characterize the performance of the feedback connection $\mathcal{T}(P,C)$. Furthermore, $J(P,C)$ is crucial in the closed-loop validation tests as mentioned in (2.20) and (2.21) and the way the feedback controller C is being designed.

For reasons of generality, in chapter 5 of this thesis, the control objective function $J(P,C)$ is chosen to be some input/output weighted form of the transfer function matrix $T(P,C)$ and given in (5.9). The transfer function matrix $T(P,C)$ in (2.5) represents the various maps between the signals in a feedback connection $\mathcal{T}(P,C)$, while the weightings are introduced to impose an additional shaping of the $T(P,C)$ matrix. Although it is impossible to transform any desirable control design objective into the single norm function $\|J(P,C)\|_\infty$, the performance characterization (5.9) has wide applicability. It may include a weighted sensitivity or mixed sensitivity characterization by proper modification of the weighting functions.

The structure of the set of models

The set of models is used to represent the incomplete knowledge of the plant P_o, caused mainly by the availability of only finite time, possibly disturbed observations of the plant. The set of models $\mathcal{P}$ should allow the evaluation of the closed-loop performance assessment test (2.19) and the closed-loop validation test of (2.20) and (2.21) in a non-conservative way.

The knowledge that the plant P_o is stabilized by a controller C, is used to capture the limited knowledge available on the plant P_o. This knowledge is used to structure the set of models $\mathcal{P}$ in chapter 5 yielding a set of models as given in Definition 5.2-4. The favourable properties of such a set of models have been summarized and the various tests mentioned in Remark 2.5-6 have been analyzed.

Due to the close connection with the dual-Youla parametrization, the set of models (or uncertainty set) $\mathcal{P}$ in (5.4) contains only models that are stabilized by the currently implemented and known controller C. The set of models is characterized by the transfer functions mentioned in Remark 5.2-7; a nominal factorization $(\hat{N}, \hat{D})$ and weighting functions $(\hat{V}, \hat{W})$.

For the evaluation of the closed-loop validation tests (2.20), (2.21) and the closed-loop performance assessment test (2.19), the results mentioned in Section 5.3.4 can be used. Rewriting the structure of the set of models $\mathcal{P}$ in (5.4) using an LFT framework, the concept of the structured singular value $\mu(\cdot)$ and the result mentioned in

Lemma 5.3-8 opens the possibility to evaluate the performance robustness of a controller C applied to a set of models $\mathcal{P}$ in a non-conservative way. This has been indicated in the analysis result presented in Corollary 5.3-9 and can be used in the various steps mentioned in Procedure 2.5-4. Structuring the set of models $\mathcal{P}$ of (5.4) as in (5.5), the performance assessment test and the controller and modelling validation tests mentioned in Remark 2.5-6 can be performed with the result mentioned in Corollary 5.3-9.

Identification of a set of models

With a set of models $\mathcal{P}$ being structured as in Definition 5.2-4, a procedure to estimate a set of models has been presented in chapter 6. The problems associated to an identification based on closed-loop data have been reviewed in chapter 3, whereas the fractional approach outlined in chapter 4 was shown to be beneficial to address the closed-loop identification problem.

Referring to Procedure 2.5-4, the estimation of a set of model $\mathcal{P}$ appears both in step 1 and step 3. In this respect, both steps are similar and require the estimation of a sets of models $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ that should satisfy $P_o \in \mathcal{P}_i$ and $P_o \in \mathcal{P}_{i+1}$ respectively for step 1 and step 3. The optimization involved in both the identification of $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ to satisfy the "quality test" or "validation test", mentioned respectively in (2.19) and (2.21), is similar. The only difference between step 1 and step 3 of Procedure 2.5-4 is the controller C being used in the optimization. Due to the closed-loop evaluation of the validation tests, a so-called control relevant identification is inevitable.

For the actual estimation of a set of models $\mathcal{P}$, the identification of the transfer functions $(\hat{N}, \hat{D}, \hat{V}, \hat{W})$ that characterize the set $\mathcal{P}$, a separate estimation of a nominal factorization $(\hat{N}, \hat{D})$ and uncertainty bounding weighting functions $(\hat{V}, \hat{W})$ is proposed. The separation between the estimation of a nominal model $\hat{P} = \hat{N}\hat{D}^{-1}$ and an parametric upper bound $(\hat{V}, \hat{W})$ for the allowable model perturbation makes the control relevant estimation of a set of models more tractable. Currently available tools to estimate a nominal model and to characterize uncertainty are applied and used throughout chapter 6. As such, the control relevant estimation of a nominal factorization is discussed in Section 6.2 and uses the fractional approach discussed in chapter 4. With the fractional approach and a set of models structured via a perturbation in a dual-Youla parametrization, the estimation of a (parametric) upper bound on the model perturbation Δ_R is discussed in Section 6.3. Both identification techniques yield the transfer functions $(\hat{N}, \hat{D}, \hat{V}, \hat{W})$ that characterize a set of models.

Robust control design methodology

Once a set of models has been characterized by the transfer functions mentioned in Remark 5.2-7, the possibility can be exploited to (re)design a robust controller on the basis of the set. Basically, this constitutes step 2 in Procedure 2.5-4, where a

controller C_{i+1} must be designed on the basis of $\mathcal{P}_i$ that should satisfy performance robustness test in (2.20) to outperform the previously designed or existing feedback controller C_i.

The problem to design such a controller C_{i+1} can be labelled as a robust control design problem, wherein the newly controller C_{i+1} is being designed such that the worst-case performance $J(P, C_{i+1}) \; \forall P \in \mathcal{P}_i$ is being optimized. Computation of such a robust controller can be done by existing techniques based on μ-synthesis and have been summarized in Section 5.4 of chapter 5. This controller synthesis technique can be used to find a robust and enhanced performing controller C_{i+1} in step 2 of Procedure 2.5-4.

7.3   Flow Chart of the Model-Based Procedure

The main ingredients mentioned above complete the model-based procedure. The possibilities to deal with an unstable system and data coming from a closed-loop system have been outlined in chapter 4. These possibility use a fractional model approach for both the estimation of a nominal model and the structure of the model set to be estimated. The identification of a set of models presented in chapter 6 is used to represent the limited knowledge available on the plant P_o, while the robust control design methodology outlined in chapter 5 is used to deal with this model uncertainty. However, to ensure that the feedback controller designed is improving the performance of the controlled plant, specific conditions have to be met during the modelling and control design phase.

Referring to Procedure 2.5-4, the conditions (2.20) and (2.21) reflect respectively a controller and a modelling (closed-loop) validation test in order to enforce (2.17). A controller must be designed that should satisfy the closed-loop validation test (2.20) before implementing it on the plant. After that, a set of models should be estimated that should satisfy the validation test (2.21).

A schematic overview of these validation tests in Procedure 2.5-4 has been depicted in Figure 2.7 and redrawn in Figure 7.1 for reference purposes. Figure 7.1 also indicates how the system identification and the control design are linked in order to improve the performance of the feedback controlled plant progressively.

Following the flow chart, starting with an initial controller C_i, closed-loop experiments of the feedback connection $\mathcal{T}(P_o, C_i)$ are obtained. The closed-loop data is used to estimate a set of models $\mathcal{P}_i$ with the model uncertainty set estimation results presented in chapter 6. That is, a nominal model and an upper bound of the model uncertainty is estimated to construct a set of models as given in (6.34). Subsequently, with the controller C_i and the estimated set of models, the performance level γ_i can be determined *a posteriori* with the result mentioned in Corollary 5.3-9.

With the given set of models $\mathcal{P}_i$, the controller C_i and the *a posteriori* determined performance level γ_i, an iteration of subsequent robust controller design and model

uncertainty set estimation can be initiated. As indicated in Figure 7.1, first the set $\mathcal{P}_i$ is used to redesign the feedback controller with the μ-synthesis summarized in chapter 5. The newly designed controller C_{i+1} must satisfy the closed-loop controller validation test, as mentioned in (2.20). As the current controller C_i is not designed for the set of models $\mathcal{P}_i$, an (optimal) redesign of the controller C_{i+1} on the basis of the set $\mathcal{P}_i$ via a μ-synthesis is most likely to perform better than the controller C_i.
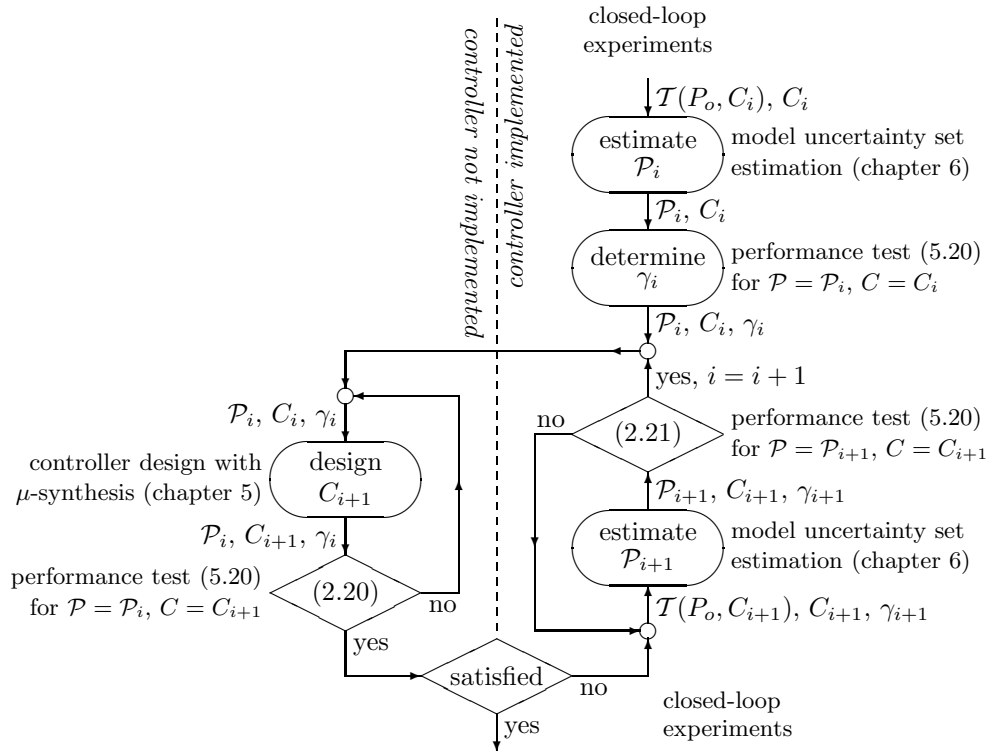
Fig. 7.1: Schematic flow chart of the model uncertainty set estimation, feedback controller design and the controller (2.20) and modelling (2.21) validation test in Procedure 2.5-4.

In case C_{i+1} passes the test (2.20), it can be implemented on the plant P_o to create a new feedback connection $\mathcal{T}(P_o, C_{i+1})$. In case the performance is satisfactory, the iteration can be stopped, otherwise the new feedback connection delivers new closed-loop data that can be used to update the knowledge on the plant P_o. As the controller C_i is changed to C_{i+1}, a new (optimal) set of models $\mathcal{P}_{i+1}$ can be estimated that is based on a control objective function which involves C_{i+1}.

The newly gathered closed-loop data and the knowledge of the new feedback con-

troller can be used to (re)estimate a set of models $\mathcal{P}_{i+1}$, as indicated in Figure 7.1. In case the updated information on the plant P_o, represented in the set of models $\mathcal{P}_{i+1}$, passes the closed-loop model validation test (2.21), the newly identified set of models $\mathcal{P}_{i+1}$ can be reused for robust controller design, from which the iteration repeats.

Obviously, the iteration is terminated as either the obtained performance level is satisfactory or the controller design or modelling phase is not able to come up with respectively a controller or set of models $\mathcal{P}$ that passes the test (2.20) or (2.21). In the next chapter, the application of this iteration is illustrated for a multivariable positioning mechanism present in a wafer stepper.

8

Identification and Control of a Wafer Stage

8.1 Summary of Application

This chapter discusses the application of the model-based procedure summarized in the previous chapter to an commercially available three degree of freedom positioning mechanism present in a wafer stepper. The positioning mechanism is labelled as the wafer stage and in Section 8.2 a description of this mechanism is given. Subsequently, in Section 8.3, the control specifications and the experiments for identification and feedback controller implementation and evaluation are outlined. In Section 8.4 the results on the identification of the set of models is presented in more detail and basically describes the application of step 1 of Procedure 2.5-4. In Section 8.5, step 2 of Procedure 2.5-4 is applied by designing and implementing an enhanced performing feedback controller. Finally, in Section 8.6 results are presented when performing a subsequent iteration on the steps step 1 and step 2 of Procedure 2.5-4. As mentioned in Section 2.5.2, such an iteration of renewed estimation of a set of models and an updated design of a feedback controller is able to yield a feedback controlled system in which the performance can be enhanced progressively. Finally, the chapter is ended by concluding remarks in Section 8.7.

8.2 Description of a Wafer Stepper

8.2.1 Application of wafer steppers

For the mass production of a high quality Integrated Circuit (IC), so-called wafer steppers are being used. Wafer steppers combine a high accuracy positioning and a sophisticated photolithographic process to manufacture IC's via a fully automated process.

The architecture of the IC is captured in a mask or recticle that bears the image of the circuit. Ultraviolet light passes through the mask and is reduced and projected

via a cascade of lenses on a silicon disk, called the wafer. By means of the photolithographic process, the architecture is exposed via a photo resist on the surface of the wafer.

In most of the applications, the wafer is used to manufacture multiple IC's and the number of circuits on the wafer may vary from 80 till 200. Due to economical reasons, such as the size of the mask and the dimension of the lenses in the wafer stepper, only a limited amount of IC can be processed on the wafer during the illumination phase of the photolithographic process.

In order to expose the complete surface of the wafer with multiple circuits, the mask must be projected sequentially onto the wafer. For that purpose, the wafer is placed on a moving table that needs to be moved or stepped accurately in at least 3 Degrees Of Freedom (3DOF) for the sequential illumination of the IC's on the wafer. The phrase "wafer stepper" originates from the subsequential moving and exposure of the wafer during the IC manufacturing on the wafer.

The performance of a wafer stepper is characterized by the number of processed wafers per hour (throughput) and the number of acceptable IC's per wafer (yield). Clearly, the movements and positioning of the wafer during the IC manufacturing will highly influence the throughput and the yield of a wafer stepper. Each step movement must be performed as fast as possible, while the wafer must be positioned accurately with small residual vibrations before illuminating another IC on the wafer. A fast and accurate servo positioning mechanism is able to decrease the time needed to make a step with the wafer during the sequential illumination of the IC's on the wafer. As such, the positioning mechanism will play an important role in influencing the performance of a wafer stepper.

Inevitably, there is an economical need to move and expose the IC's as fast as possible to increase the throughput of the wafers. The exponential growth of the transistor density on IC's (Stix 1995) puts additional high requirements on the accuracy of the positioning mechanism. Similar as in (de Roover 1997), the requirements on fast and precise positioning of the wafer are important motivations in the application discussed in this thesis. The modelling and servo control of the positioning mechanism will serve as an illustrative example for the framework introduced in this thesis.

8.2.2 Description of the wafer stage

The servo positioning mechanism used to move the wafer during the IC manufacturing is called the wafer stage. The wafer stage discussed in this thesis is an integral part of a linear direct drive type wafer stepper and is denoted by the Silicon Repeater 3rd generation (SIRE3). In the wafer stage of the SIRE3, the wafer is positioned in three degrees of freedom (3DOF) on the horizontal plane of the granite block. Three laser measurements are used to determine the horizontal position of the wafer

chuck, whereas three linear motors are used to position the wafer chuck in 3DOF. A schematic overview of the mechanical servo positioning mechanism of the SIRE3 is depicted in Figure 8.1.
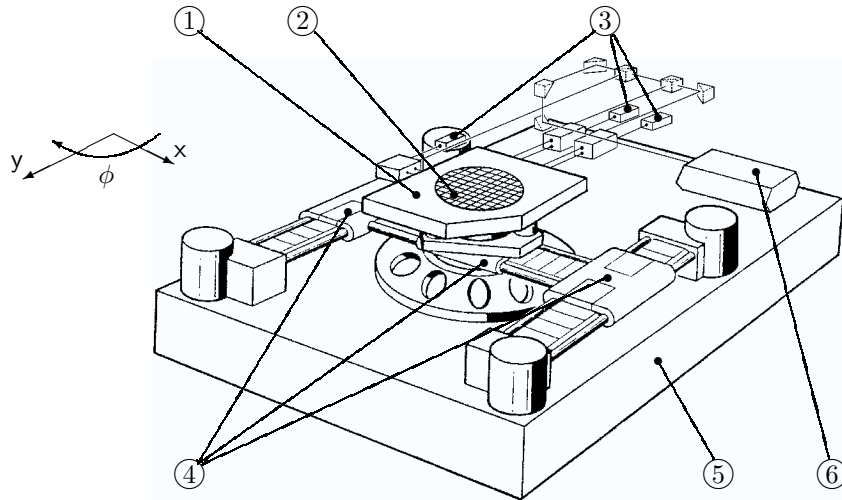
Fig. 8.1: Schematic view of a wafer stage; 1: mirror block, 2: wafer chuck, 3: laser interferometers, 4: linear motors, 5: granite block, 6: laser.

In the wafer stage depicted in Figure 8.1, the table used to store and move the wafer, is called the wafer chuck. The wafer chuck is equipped with an air bearing and placed on a large suspended granite block to reduce the effect of external vibrations. The position of the wafer chuck on the horizontal surface of the granite block is measured by means of laser interferometry.

This makes the servo positioning mechanism of the wafer stepper a multivariable system, having three inputs and three outputs. The inputs reflect the currents to the three linear motors, whereas the outputs are constructed by measuring the position of the wafer chuck both in x-, y-direction (translation) and the ϕ-direction (rotation), as indicated in Figure 8.1.

8.2.3   Experimental set up

An experimental set up of a SIRE3 wafer stage has been provided by Philips Research Laboratories in Eindhoven, the Netherlands. The experimental set up is available in the laboratory of the Mechanical Engineering Systems and Control Group at Delft University of Technology, the Netherlands. A picture of the set up is given in Figure 8.2 and includes a wafer stage, an interface for both the laser interferometers

and the linear motors and a host computer equipped with a digital signal processor (DSP).

Fig. 8.2: Experimental set up with (from left to right) the wafer stage, interface and host computer with digital signal processor.

In the wafer stage, laser interferometry is used to determine the horizontal position of the wafer chuck, whereas three linear motors are used to position the wafer chuck in 3DOF. Below, a short description is given of the components in the experimental set up that are important for changing, measuring and controlling the position of the wafer chuck. A more detailed discussion of the experimental set up can also be found in de Roover (1997).

Actuation

As indicated in Figure 8.1, three linear motors are used to position the wafer chuck in 3DOF over the surface of the granite block. The linear motors are placed in an H-shape to accomplish motion in 3DOF over the surface of the granite block.

The three linear motors are electromagnetic (voice-coil) motors and are built up from a stator and a slider. The stator consist of an iron core surrounded by alternately segmented copper coils with opposite polarity. The slider embraces the stator with eight permanent magnets connected by an iron yoke and moves along the stators by means of a ball bearing.

A close-up picture of the wafer stage is given in Figure 8.3, where the layout and position of the three linear motors can be seen in more detail.
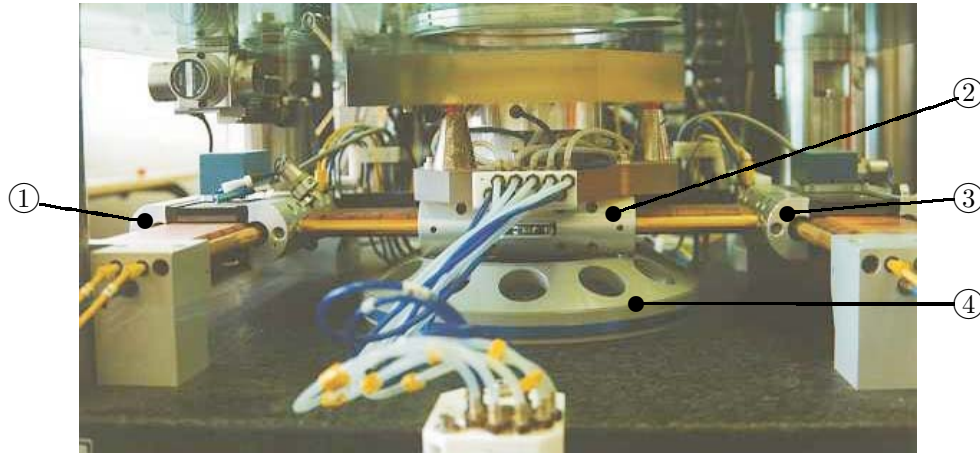
Fig. 8.3: Close-up of wafer stage displaying the three linear motors (1: y_1-motor, 2: x-motor, 3: y_2-motor) and 4: the wafer chuck with air bearing.

As indicated in Figure 8.1, in the H-stage two linear motors are placed in line with the y-direction. In Figure 8.3 these motors are indicated by the y_1- and the y_2-motor. One linear motor is placed in line with the x-direction and is indicated by the x-motor. Evidently, the x-motor is used to move the chuck in the x-direction. A simultaneous and equivalent activation of the y_1- and y_2-motor will yield an activation in y-direction. A difference in the current to the y_1- and y_2 motor may be used to obtain a (small) rotation of the wafer chuck to move in the ϕ-direction.

Sensing

To measure the position of the wafer stage moving over the surface of the granite block, laser interferometry is being used. The three degrees of freedom in the positioning of the wafer stage (x- and y-translation and ϕ-rotation) are measured by three position sensors. A position measurement in x-direction is taken, while two parallel measurements y_1 and y_2 in y-direction are used to measure the y-translation and the ϕ-rotation.

The three position measurements are performed by a Hewlett-Packard Helium-Neon Laser Transducer System (Hewlett-Packard 1985). The light coming from the Helium-Neon laser is split into three beams and projected onto an ultra flat mirror block that is an integral part of the wafer chuck. A photo of the stage with mirror block and parts of the Laser Transducer System can be seen from the top view of the wafer stage in Figure 8.4.

The mirror block with the three interferometers (one in x-direction, two in y-
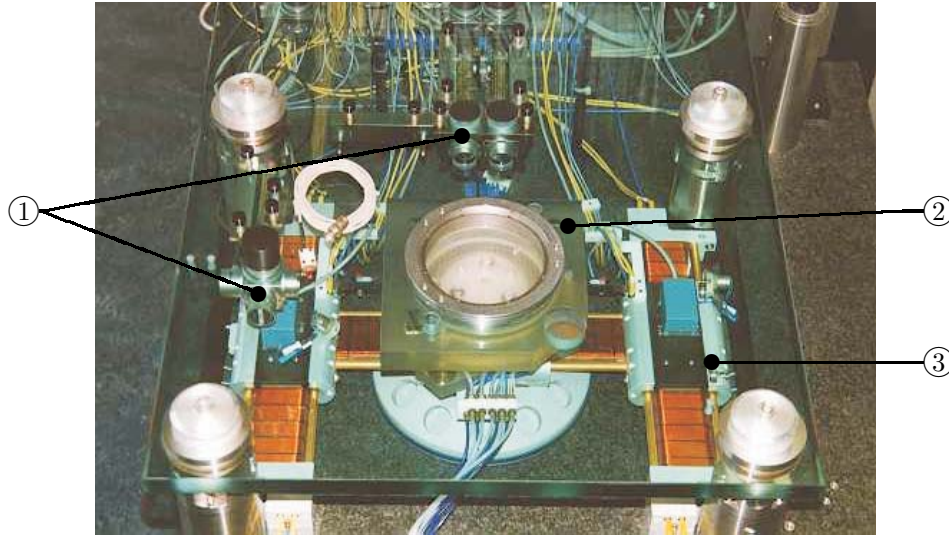
Fig. 8.4: Top view of wafer stage with 1: interferometers, 2: mirror block and 3: H-
          shape of the three linear motors.

direction) can be seen from the top view and is also redrawn in Figure 8.5.
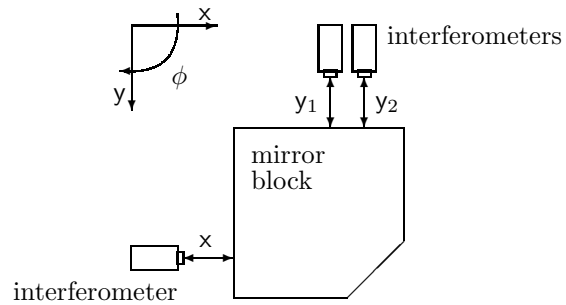
Fig. 8.5: The three position measurements on the mirror block

The three position measurements can be used to reconstruct the position of the
stage in x-, y- and ϕ-direction. This idea has been illustrated in Figure 8.5, where
the three position measurements of mirror block are depicted schematically. From
the three (relative) position measurements given in Figure 8.5, the three degrees of

freedom of the stage in x-, y- en ϕ- direction can be obtained via

$$\begin{bmatrix} x \\ y \\ \phi \end{bmatrix} = T \begin{bmatrix} x \\ y_1 \\ y_2 \end{bmatrix}$$

where T is a transformation matrix defined by

$$T := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{bmatrix} \tag{8.1}$$

that tries to statically decouple the two measurements y_1 and y_2 in the y-direction. It should be noted that

$$\phi = \frac{y_1 - y_2}{2}$$

is not the actual rotation of the stage, as this would require the computation involving a sinusoidal term. However, the rotation of the stage is limited and for small difference between y_1 and y_2, e.g. small rotations, the value of ϕ is proportional to the angular rotation of the stage.

Each beam in the Laser Transducer System is reflected from the mirror block and interferes with the light coming from the laser. The interference of the laser light is used to measure the velocity of the stage by sampling the Doppler shift in frequency of the interference signal. Subsequently, a relative position signal is obtained by integration and is done summing up the samples obtained. The sampling of the Doppler shift can be done at various resolutions and depends on the maximum velocity of the stage (Hewlett-Packard 1985). In the application discussed throughout this thesis, a fixed resolution of 1/48 of the 633 nm wavelength of the Helium-Neon laser is being used. As such, the three position measurements x, y_1 and y_2 in Figure 8.5 are in increments of 13.1875 nm, whereas the velocity of the stage is limited to 0.12 m/s at this resolution.

Signal processing

The signals coming from the laser interferometers and the signals to the linear motors are digested by a digital signal processor (DSP). The DSP is based on the TMS320C30 Texas Instruments processor and is a floating-point processor running at 33MHz. The DSP is used to generate and gather signals for identification purposes and is able to implement digital controllers for controlling the position of the wafer stage. The DSP is installed in a host computer to provide an interface to the data acquisition and controller implementation.

For the implementation of digital controllers, the DSP reads the data coming from the interferometers using a special digital I/O card, while a 16 bit DAC is

used to generate an analogue signal for the linear motors. External reference signals
on the input and the output signal of the controller can be used for closed-loop
excitation purposes. Summarizing, the closed-loop configuration of the wafer stage
servo positioning mechanism with the DSP can be depicted by the block diagram
given in Figure 8.6.
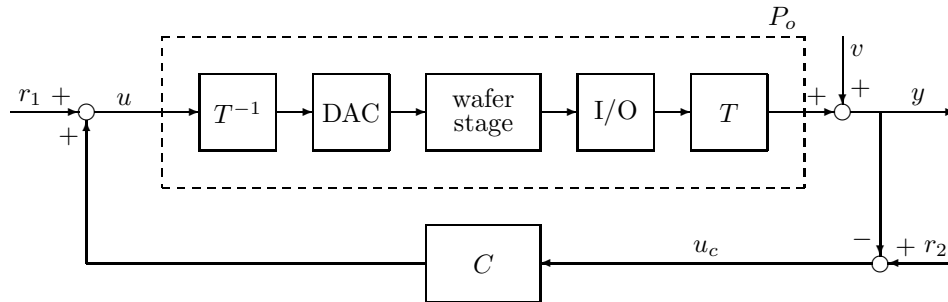
Fig. 8.6: Closed-loop block diagram of wafer stage with digital controller implemen-
          tation.

The transformation matrix T and the inverse T^{-1} in Figure 8.6 has been defined
in (8.1). As mentioned before, the transformation T is used to statically decouple
the two measurements y_1 and y_2 in the y-direction. Additionally, the transformation
matrix T and T^{-1} enables to process the reference signals and the measured signals
in terms of position signals in x-, y- and ϕ-direction.

As such, the input u is used to indicate the (transformed) actuator input. The
input signal u consists of three input signals u_x, u_y and u_ϕ reflecting respectively
actuation signals in x-, y- and ϕ-direction. In a similar way, the (transformed) laser
interferometer output signal y can be considered to consists of the three position
signals y_x, y_y and y_ϕ.

From the block diagram of Figure 8.6, the (unknown) plant P_o and the (known
digital) controller C can be distinguished. The configuration is similar to the feedback
connection $\mathcal{T}(P, C)$ of a system P and a controller C as previously displayed in
Figure 2.2. With the three actuation input signals and the three laser interferometer
output signals, the unknown plant P_o is a multivariable (3×3) dynamical process.

The ability to add (digital) reference signals into the feedback loop of Figure 8.6
is directly implemented in the DSP. As such, the reference signals r_1 and r_2 serve two
important purposes.

- The reference signals r_1 and r_2 in Figure 8.6 can be used to command a move-
  ment or step of the wafer chuck in a desired direction. In this way the reference
  signals can be used to evaluate the performance of the feedback controlled posi-
  tioning mechanism system by applying a reference signal r_2 and a feedforward

signal r_1 to move the wafer chuck.

- To avoid problems associated to loss of (closed-loop) identifiability, see also Assumption 3.2-2, the reference signals r_1 and r_2 are also used to excite the feedback connection. During the gathering of time domain data for identification purposes, both the reference signals r_1 and r_2 will be used.

Basically, the aforementioned feedback connection is used as an experimental configuration for respectively the identification, control design and analysis of the servo positioning mechanism. During identification, the reference signals are used for (closed-loop) excitation purposes. To evaluate the performance of the servo mechanism in terms of speed (throughput) and accuracy (yield), the reference signals are specified as a command reference signal r_2 and a feedforward signal r_1 (de Roover 1997). A more detailed discussion of the reference signals being used during the experiments is postponed until Section 8.3.

8.2.4   Dynamics and modelling of the wafer stage

To increase the throughput and yield of the wafer stepper, the subsequent positioning steps of the wafer stage during wafer illumination and exposure must be performed as fast as possible. However, fast and aggressive movements of the wafer stage will excite many of the resonance modes the mechanical positioning mechanism may exhibit. To accurately control these residual vibrations, it is important to know the dynamical behaviour of the servo positioning mechanism.

As such, modelling of the positioning mechanism is concerned with finding a model that describes the dynamical behaviour of the wafer stage. The model is used to predict the dynamical or time domain based relation between the input signal u and the possibly disturbed output signal y as given in Figure 8.6. Basically, the plant P_o in Figure 8.6 is used to denote the unknown dynamical system to be modelled. Once a model of P_o is available, a feedback controller C can be (re)designed to control the positioning mechanism and to reduce any residual vibrations that the mechanical system may exhibit.

Dynamics of wafer stage

To give an indication of the dynamical behaviour and the resonance modes of the positioning mechanism, an estimate of the frequency response of P_o has been depicted in the Bode amplitude plot of Figure 8.7. The experimentally obtained frequency response of P_o is denoted by $\hat{G}(\omega_j)$, where ω_j denotes the frequencies along a pre-specified frequency grid $\Omega = (\omega_1, \omega_2, \ldots, \omega_l)$.

To anticipate on the experiment design discussed in Section 8.3.2, it can be mentioned here that the estimated frequency response $\hat{G}(\omega_j)$ is estimated by conducting closed-loop experiments with periodic reference signals r_1 and r_2 around the center

position of the wafer stage. By means of a closed-loop spectral analysis, a frequency response estimate $\hat{G}(\omega_j)$ of the unknown plant P_o is obtained in 300 points within the frequency range between 10 and 1000 Hz. In closed-loop spectral analysis, a cross spectrum between the reference signals and the input signal u and a cross spectrum between the reference signals and the output signal y is estimated to come up with an estimate of the frequency response between u and y. An amplitude Bode plot of the data $\hat{G}(\omega_j)$ is depicted in Figure 8.7 by a dotted line.
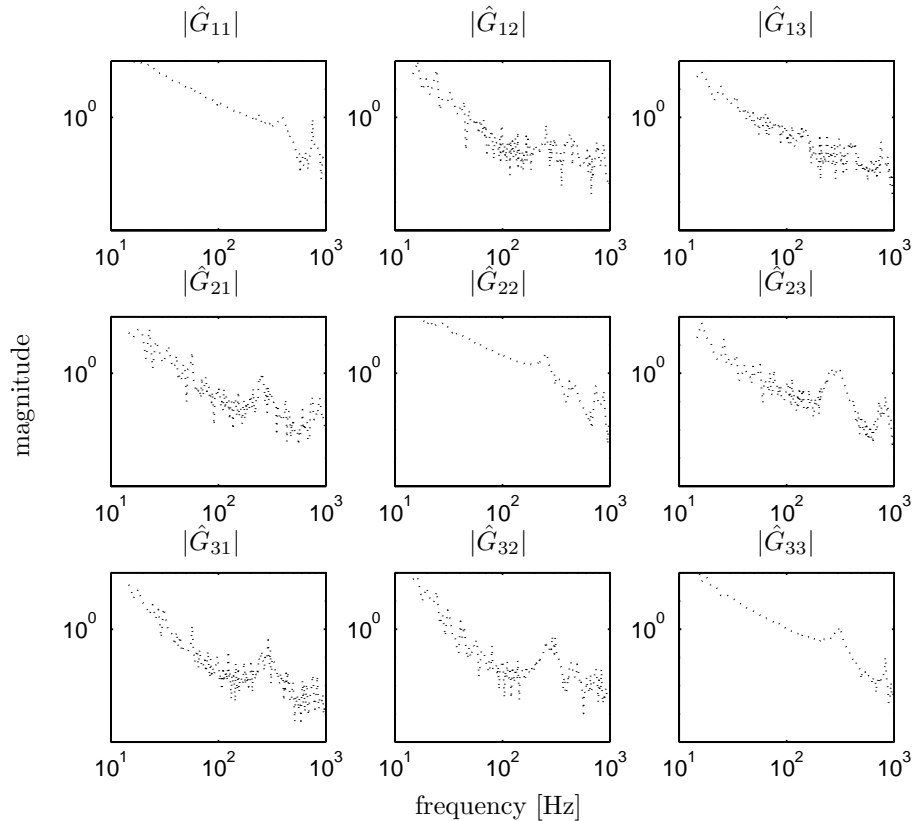
Fig. 8.7: Magnitude Bode plot of experimentally obtained frequency response $\hat{G}(\omega_j)$ of the nine transfer functions of P_o from actuator input u to laser interferometer output y in Figure 8.6.

Although only an amplitude Bode plot of $\hat{G}(\omega_j)$ has been given in Figure 8.7, the data does give an indication of the dynamical behaviour of the nine transfer functions between actuator input signal $u = [u_x\ u_y\ u_\phi]^T$ and the laser interferometer output signal $y = [y_x\ y_y\ y_\phi]^T$. The nine amplitude Bode plots are labelled in correspondence with the ordering within the multivariable transfer function P_o, when considering the

map from the actuator input u to the laser interferometer output y.

It can be observed from Figure 8.7 that the low frequent behaviour (below 100 Hz) of P_o is dominated by a doubly integrating action. This is due to the fundamental proportional relation between a force applied to a free rigid body in motion and its acceleration. Due to the presence of a double integrator in the relation between force (actuator input) and position (laser interferometer output), the positioning mechanism is marginally stable and can not be operated in an open-loop configuration. Closed-loop experiments with a controller C that yields a stabilizing feedback connection $\mathcal{T}(P_o, C)$ is inevitable.

Above the 200 Hz, several peaks in the frequency response estimate indicate the resonance modes present in the mechanical system. The cross-talk or interaction (the non-diagonal elements) are reasonable small compared to the diagonal elements in the frequency response estimate. However, still dominantly present, the interaction terms may influence the multivariable behaviour of the wafer stage considerably. Modelling the (dominant) resonance modes and the interaction effects so as to design a (multivariable) feedback controller for attaining an improved servo positioning performance is preferable and the aim of the application discussed in this thesis.

Some existing models of wafer stage

An approach to model the dynamical behaviour of the servo positioning mechanism using an analytical procedure based on first principles has been presented in de Roover and van Marrewijk (1995). In this approach, the interconnection of multiple rigid bodies is used to derive a (non-linear, 18th order) dynamical model of the mechanical servo system. A more elaborate discussion on the modelling of the wafer stage as a multi body mechanical system can also be found in de Roover (1997).

First principles modelling provides physical insight in the behaviour of the wafer stage and has explained the nature and cause of some of the resonance modes of the mechanical system. Unfortunately, only a qualitative description of the dynamical behaviour has been obtained (de Roover 1997). A more accurate qualitative description of the (linear) input/output behaviour of the mechanical system is preferable, especially in the case when the model is to be used for the design of a feedback compensator.

An alternative to first principles modelling of a mechanical system, as advocated in this thesis, is the use of a system identification technique to find a linear model suitable for control design. Especially for (fast) electromechanical system where the gathering of data can be performed within a reasonable time, the use of a system identification technique may be beneficial. Because high speed mechanical servo systems allow a relatively fast and easy data acquisition, frequency domain identification techniques are well suited for determining its dynamics. In that case, an estimated frequency response can be used for curve fitting purposes to estimate a linear parametric model

(Ljung 1993*b*, Bayard 1994, Pintelon *et al.* 1994, Kollár 1994). Due to the nature of the high speed mechanical servo system, frequency domain based system identification techniques will also be used throughout this chapter.
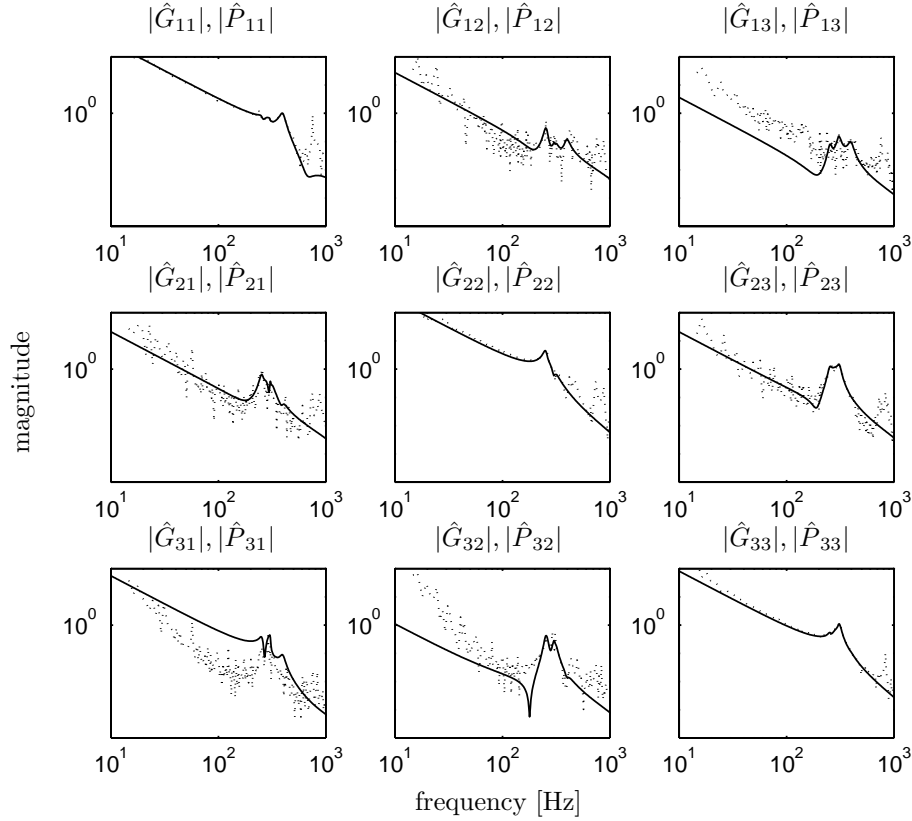
Fig. 8.8: Magnitude Bode plot of experimentally obtained frequency response $\hat{G}(\omega_j)$ ($\cdots$) and 18th order linear model $\hat{P}$ (—) found by curve fitting from (de Callafon *et al.* 1996*a*).

In de Callafon *et al.* (1996*a*), such a curve fit technique has been presented and applied to the frequency response data depicted in Figure 8.7. By means of a weighted two-norm minimization, a (linear, 18th order) model $\hat{P}$ has been derived directly on the basis of the frequency response data. For details one is referred to Appendix B of this thesis where the computational aspects of the curve fitting procedure are summarized. For illustrative purposes, the result obtained in de Callafon *et al.* (1996*a*) has been depicted in the amplitude Bode diagram of Figure 8.8. As can be seen from Figure 8.8, mostly the low frequent and dominant resonance modes have been captured by the 18th order model.

Although the above mentioned procedures of analytical and experimental modelling have produced models for the wafer stage, the question wether or not the models are suitable for control design has not been addressed. Furthermore, a quantification of the error made during modelling or, alternatively, a characterization of an allowable model perturbation so as to design a robust performing feedback controller has not been presented yet.

As mentioned in this thesis, both items must be addressed to be able to design and guarantee a robust and enhanced performing feedback controller for the wafer stage. Both items are addressed in the model-based procedure of Procedure 2.5-4 where a set of models $\mathcal{P}$ is estimated in view of the intended control design application. The set of models has to be estimated such that it is suitable for subsequently designing enhanced performing controllers.

8.3 Design Objectives and Experiments

8.3.1 Control Implementation and Specifications

Controlling the positioning mechanism of the wafer stage aims at minimizing the servo error, while moving the wafer chuck as fast as possible. The servo error is indicated in Figure 8.6 by the signal u_c. The signal u_c serves as an input signal for the feedback controller C.

To be able to estimate a set of models that captures the dynamics of the wafer stage and that is suitable to design an (improved) robust controller, first the specifications for the intended control application must be characterized. The specifications for designing a controller, assumed to be reflected by a single control objective function $J(P, C)$ in Definition 2.2-13, will effect the way in which the control-relevant estimation of a set of models $\mathcal{P}$ has to be performed. Before going into the details associated with the identification of the set of models, first the control design specifications for the design of an improved controller are outlined.

Reference step trajectory specification

The position steps that have to be performed with the wafer stage typically require a displacement of 10^{-2} m, which is the approximate size of a single IC. The design specification for the SIRE3 wafer stepper is to bring the servo error within a bound of approximately $52 \cdot 10^{-9}$ m (4 times the measurement resolution of 13.1875 nm) as soon as possible after a positioning step has been carried out. This is due to the fact that the wafer mounted on the wafer chuck must be kept in a constant position before a chip can be illuminated and exposed on the surface of the wafer. Henceforth, controlling the positioning of the wafer chuck requires the combined design of both a feedback controller and the appropriate reference r_2 and feedforward signal r_1 (de Roover et $al.$ 1996, de Roover 1997). In this thesis however, the attention is focused on the

identification of a set of models, denoted by $\mathcal{P}$, to improve the design of the feedback controller only.
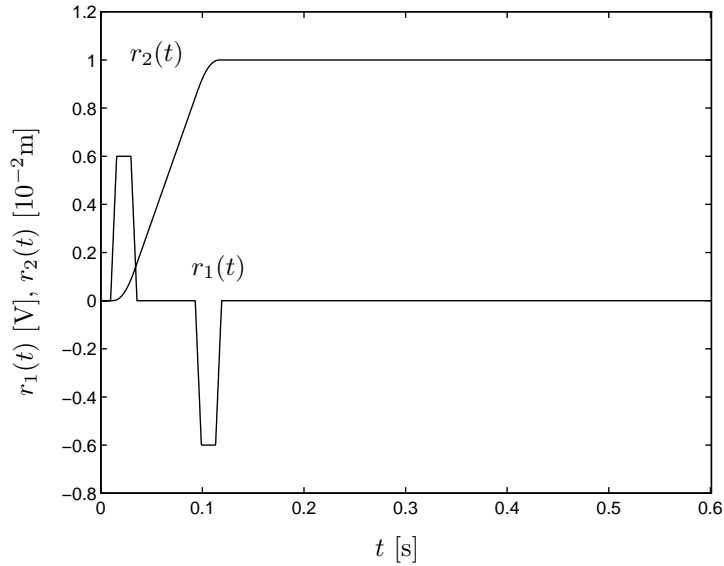
Fig. 8.9: Shape of reference signal r_2 and feedforward signal r_1 for a positioning step in either x- or y-direction.

In order to compare feedback controllers designed on the basis of a set of models $\mathcal{P}$ being estimated, for analysis purposes the signals r_2 and r_1 in Figure 8.6 are fixed to some pre-specified desired trajectory. This pre-specified trajectory is based on the dominating open loop dynamical behaviour of the wafer stage. As already indicated in Figure 8.7 and Figure 8.8, the dominant open-loop behaviour is given by a double integrator, relating the force generated by the linear motors to the position of the wafer chuck. Based on this relatively simple model, r_2 will denote a desired position profile, whereas r_1 denotes (a scaled) acceleration profile obtained by computing the second derivative of r_2. A typical shape of the reference signal r_2 and the feedforward signal r_1 to position the wafer chuck in either the x- or y- direction over approximately 1cm is depicted in Figure 8.9.

In Figure 8.9, the position profile r_2 is obtained by taking into account a maximum jerk (derivative of acceleration) and a maximum speed of the wafer chuck. As a result, the ramp of the position profile r_2 consists of three parts. Firstly, a third order polynomial for optimal acceleration at maximum jerk. Secondly, a linear interpolation at maximum speed. Finally again a third order polynomial for optimal de-acceleration at maximum jerk of the wafer chuck. The resulting acceleration profile r_1 is the second derivative of r_2.
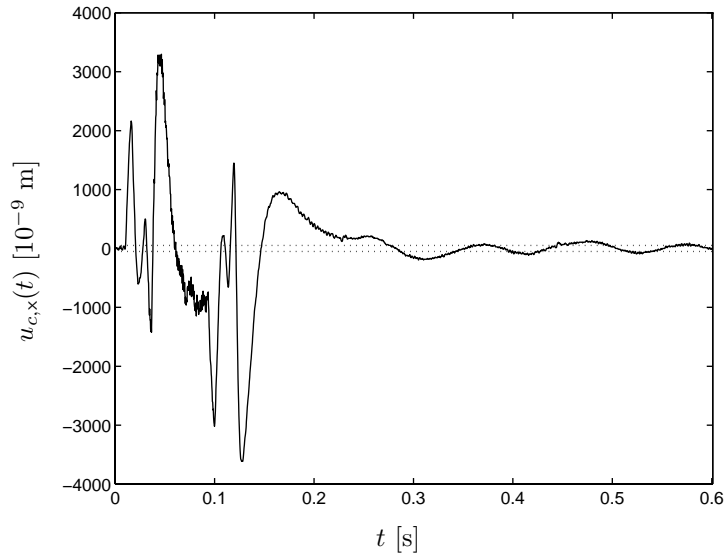
Fig. 8.10: Servo error u_c response to a step in x-direction (—) and required accuracy
interval ($\cdots$).

Application of both reference signals in either an x- or y-direction is labelled as a
(wafer) step respectively in x- or y-direction. Using these specified reference signals r_1
and r_2, a servo error pattern can be measured. Anticipating on the results mentioned
in Section 8.4.1, it can be mentioned here that in the initial experimental set up,
the feedback controller consists of three parallel PID controllers. The three PID
controllers are used to control the positioning in x- y- and ϕ-direction individually.
For this initial experimental set up, the servo error $u_{c,x}$ depicted in Figure 8.10 for a
step in the x-direction is obtained.

It can be observed from Figure 8.10 that the servo error $u_{c,x}$ is hardly within the
bounds of 52nm indicated by the dotted lines. According to Figure 8.9 the reference
step trajectory, specified by the signals r_1 and r_2, ends at approximately 0.12 s.
However, the servo error is still not settled at 0.25 s. Furthermore, the servo error
$u_{c,x}$ exhibits a low frequent vibration after the reference step trajectory has ended.
As a result, the settling time of the servo error needs to be improved. Both an
enhancement of the speed of decay and a reduction of the low frequent vibration of
the servo error is desired to improve the behaviour of the servo mechanism.

Weighting functions in control design

To accomplish a faster settling and an improved low frequent vibration suppression
of the servo error, the feedback controller C_i currently implemented in the initial

experimental set up of the wafer stage needs to be improved. For that purpose, the control design procedure mentioned in Section 5.4 will be used.

In the control design procedure, the control design specifications are captured by a (single) control objective function $J(P, C)$. As mentioned in Definition 2.2-13, the (nominal) performance of the controller is measured by a weighted $\mathcal{H}_\infty$ norm of the $T(P, C)$ matrix. The weighting functions U_2 and U_1 in (5.9) can be used to translate the control design specifications into frequency dependent weightings and scalings used during the controller design.

It can be observed from the results presented in chapter 6 that the weighting functions U_1 and U_2 in the control objective function

$$J(P, C) = U_2 T(P, C) U_1$$

are not only crucial during the design of a controller, but also for the identification of a set of models $\mathcal{P}$ on which the controller needs to be designed. The control relevancy of an estimated set of models $\mathcal{P}$ (or a nominal model $\hat{P}$ only) is influenced by the criterion on which the controller needs to be designed, see e.g. (6.8). As such, it is important to make a statement on the weighting functions U_2 and U_1 to be used during both the identification and the control design.

Remark 8.3-1 *To accomplish an attenuation of the low frequent disturbances in the servo error in Figure 8.10, it is desired to equip the controller with additional low frequent gain, such as an integrator. For a fast settling of the servo error, it is advisable to increase the bandwidth of the feedback loop. These design specifications should be captured by the weighting functions U_2 and U_1.*

As the orders of the weighting functions U_2 and U_1 directly influence the order of the controller derived via the existing $\mathcal{H}_\infty$ norm-based controller computation software, it is advisable to keep the order of the weighting functions as low as possible. Choosing such low order, multivariable, weighting filters U_2 and U_1 is a non trivial task. To simplify the procedure of finding the filters the following simplifications have been made.

- Firstly, the multivariable weighting functions U_2 and U_1 are chosen as simple (block) diagonal weighting filters.

  Although non-diagonal weighting filters allow more freedom in specifying control design specifications, diagonal weighting filters simplify the selection of U_2 and U_1 considerably. Furthermore, it will be shown here that the relatively simple diagonal weighting filters allow enough freedom to design and improve the performance of the feedback control system of the wafer stage.

- Secondly, the weighting filters U_2 and U_1 are related via

$$U_1 = U_2^{-1} \tag{8.2}$$

and U_2 is chosen according to the block diagonal structure

$$U_2 = \begin{bmatrix} U_l & 0 \\ 0 & U_r^{-1} \end{bmatrix} \tag{8.3}$$

where U_l and U_r are weighting filters to be specified.

The reason to choose U_2 and U_1 according to the above mentioned relation is to link the choice of the weighting filters with a loop shaping control design procedure. This is due to the following argument, see also Bongers (1994). It can be verified that the minimizing value of the norm $\|U_2 T(P, C)U_1\|_\infty$ obtained by the minimization

$$\min_C \|U_2 T(P, C)U_1\|_\infty \tag{8.4}$$

with U_2 and U_1 satisfying (8.2) and (8.3) is equivalent to the value obtained by a so-called loop shaped minimization. For this to hold, the loop shaped minimization is given by

$$\min_{C_l} \|T(P_l, C_l)\|_\infty \tag{8.5}$$

where P_l is given by a loop shaped weighted version of P

$$P_l = U_l P U_r \tag{8.6}$$

and the controller C_l in (8.5) is related to the controller C in (8.4) via

$$C = U_r C_l U_l \tag{8.7}$$

Modifying the system P via the loop shaping of (8.6), computing a controller C_l via (8.5) and bringing the loop shape weighting filters back into the controller in (8.7) is a loop shape control design procedure (Bongers 1994).

Unfortunately, the loop shape procedure increases the order of the controller C being computed, since the loop shape weighting filters are used twice. Firstly, to loop shape the system P, thereby increasing the order of the system P_l and the resulting optimal controller C_l being computed. Secondly, to recompute the controller C from C_l, thereby increasing the order of C again.

Instead of a loop shape design procedure, as mentioned in Section 5.4, a (direct) minimization of the weighted $T(P, C)$ matrix in (8.4) is used. Furthermore, a worst-case optimization is performed by considering $\|U_2 T(P, C)U_1\|_\infty$ for all systems P that lie in a certain set of models $\mathcal{P}$. Although a loop shape control design procedure will not be used here, the interpretation of U_l and U_r in (8.3) being loop shape weighting function, simplifies the design and choice of the weighting filters U_2 and U_1 considerably.

Taking into account the above mentioned simplifications of the weighting functions U_2 and U_1, the control design specifications mentioned in Remark 8.3-1 are implemented by the choice of the following (discrete-time) weighting filters.

$$U_l(q) = \begin{bmatrix} 3.5 \cdot 10^{-4} & -7.8 \cdot 10^{-6} & -2.5 \cdot 10^{-5} \\ -4.6 \cdot 10^{-5} & 2.8 \cdot 10^{-4} & -8.4 \cdot 10^{-5} \\ -1.1 \cdot 10^{-5} & 5.8 \cdot 10^{-6} & 1.0 \cdot 10^{-3} \end{bmatrix}$$

$$U_r(q) = 0.6 \begin{bmatrix} \frac{(q-0.9)(q-0.9)}{(q-0.99999)(q-0.99)} & 0 & 0 \\ 0 & \frac{(q-0.91)(q-0.91)}{(q-0.99999)(q-0.99)} & 0 \\ 0 & 0 & 0.75\frac{(q-0.95)}{(q-0.9999)} \end{bmatrix}$$

$$(8.8)$$

Clearly, the shape of the weighting filters cannot be chosen without some prior knowledge about the actual plant P_o. For that purpose, the data and the relatively simple 18th order model depicted in Figure 8.8 were used to make a statement about the weighting filters. The filters U_l and U_r have to be substituted in (8.2) and (8.3) to get the closed loop weighting filters U_2 and U_1.

As mentioned above. the choice of U_l and U_r are based on the interpretation of being loop shaping filters. With this interpretation, the filter U_l is chosen as a constant matrix to (approximately) decouple the plant P_o around the frequency of 90 Hz. The 90 Hz is chosen as the desired closed-loop bandwidth of the servo system and U_l is found by computing the real-valued pseudo inverse of the frequency response of the plant P_o at 90 Hz, see also Maciejowski (1989). The diagonal filter U_r is chosen to incorporate the control design specifications mentioned in Remark 8.3-1. As such, discrete-time poles around 1 are introduced to improve the attenuation of low frequent disturbances present in the servo error depicted in Figure 8.10. The scaling of the transfer functions in U_r is used to attain a bandwidth of approximately 90 Hz.

8.3.2   Experiment Design

Experiments on the wafer stage have to be done in the closed-loop setting depicted in Figure 8.6. Thereby, the role of the reference signals in Figure 8.6 is twofold. As mentioned in Section 8.3.1, the reference signals r_1 and r_2 are used to specify respectively an acceleration reference and a position reference profile to perform a position step with the wafer stage. From an identification point of view, the reference signals are used to excite the closed-loop system to avoid problems associated with closed-loop identifiability issues.

Although the signals r_1 and r_2 can be used to conduct identification experiments, mostly low frequent information is contained in the signals depicted in Figure 8.9. To get enough information on the system in the frequency range from approximately 10 Hz till 1 kHz, different reference signals have to be used. Furthermore, the flexibility

of the experimental set up and the speed of the mechanical system allow the use of different reference signals that can be used to obtain a frequency response estimate of the positioning mechanism.

To use the flexibility in specifying the experiment design in the experimental set up, the reference signals are specified periodic signals that consist of a sum of sinusoids

$$r(t) := \sum_{i=1}^{l} sin(\omega_i t + \phi_i) \tag{8.9}$$

specified at predefined frequency grid $\Omega = \{\omega \mid \omega = \omega_i, \ i = 1, 2, \ldots, n\}$. The phase shift ϕ_i of the sinusoids has to be chosen properly to avoid high signal amplitudes due to the cumulative effect of adding multiple sinusoids.

A derivation of a sequence of phase shifts $\{\phi_i\}$, $i = 1, 2, \ldots l$ that minimizes $\sup_t |r(t)|$ is known as a so-called Schroeder-phased sum of sinusoids. Unfortunately, six different reference signals[1] have to be specified. In case all six reference signals are based on the same frequency grid Ω, a Shroeder-phased sum of sinusoids will yield six reference signals that are all the same.

To avoid complications associated to having six reference signals that are identical, the phase shifts ϕ_i in the sequence $\{\phi_i\}$, $i = 1, 2, \ldots l$ of (8.9) are chosen independently from a uniform distribution over the interval $(-\pi, \pi]$. In this way, the six reference signals are generated independently and uncorrelated and a random phased sequence of sinusoids is generated with favourable properties (Pintelon *et al.* 1994).

For the identification experiments of the wafer stage, the following technical details can be mentioned. The sampling of the continuous-time signals is set at a sampling time $\Delta T = 3.0 \cdot 10^{-4}$ s. The reference signals are specified as a periodic signal having a period of 2048 data points or $T \approx 0.614$ s. As a result, the frequency resolution Δf of the frequencies that can be distinguished in the periodic signal is fixed at $\Delta F = 1/T \approx 1.628$ Hz. The periodic signal is a random phased sequence of 200 sinusoids. The frequency grid Ω of the 200 sinusoids is distributed (approximately logarithmically) between $9\Delta f \approx 14.65$ Hz and $714\Delta f \approx 1162.11$ Hz.

To give an indication of the reference signals being used, a time domain plot and the spectrum in the frequency range between 100 Hz and 1 kHz of one of the six references signals is given in Figure 8.11. It can be seen from the time domain plot that the random phased sequence of 200 sinusoids does not exhibit a high signal level due to the cumulative effect of adding multiple sinusoids. Although the signal looks "noisy", the spectrum is very well defined for the 200 frequency points in the frequency grid Ω.

With the reference signals described above, three different experiments are performed with independently generated random phased sequence of 200 sinusoids having the same frequency grid Ω. The experiments are all performed in a closed-loop and

---

[1]Both r_1 and r_2 consist of three reference signals, respectively in x-, y- and ϕ-direction.
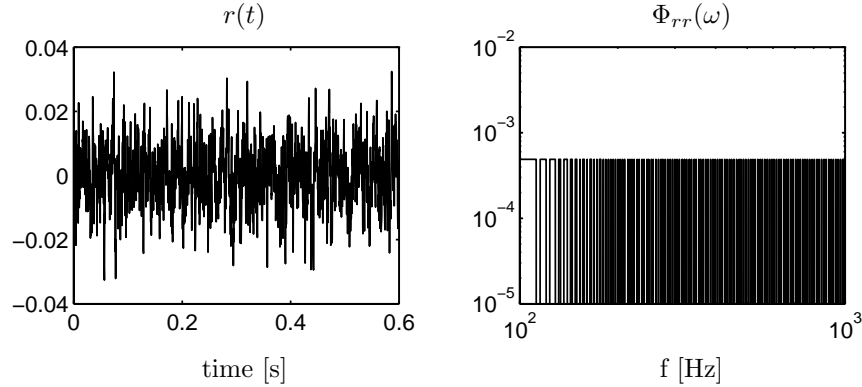
Fig. 8.11: Time domain plot (left) and spectrum (right) of the reference signal r_1 in
x-direction configured as a sum of 200 sinusoids with random phase.

for averaging purposes, a time span of 50 periods of the periodic signal is used to
capture the signals u and y depicted in Figure 8.6 for each experiment. With the
knowledge of the controller used during the closed-loop experiments, the data $\{u, y\}$
being measured and the weighting functions U_2 and U_1 mentioned in (8.2), (8.3) and
(8.8) the control relevant identification of a set of models $\mathcal{P}_i$ commences.

8.4   Initiating the Suboptimal Design

8.4.1   The initial controller

In the initial experimental set up, the feedback controller depicted in Figure 8.6 is real-
ized by three parallel PID controllers to control the positioning in x- y- and ϕ-direction
individually. Following the notation as introduced in Procedure 2.5-4, this controller
is denoted by C_i and is the initial controller to be used in the subsequent procedure of
closed-loop identification and model-based controller design. For reference purposes,
an amplitude Bode plot of the initial controller C_i is given in Figure 8.12.

Although the feedback controller can be designed as a multivariable controller,
it can be seen from Figure 8.12 that the initial controller C_i only contains diagonal
elements. Evidently, the aim is to redesign the controller C_i to attain an improved
feedback controller C_{i+1}. The design of the feedback controller is done on the basis
of a set of models $\mathcal{P}_i$ being estimated.

8.4.2   Identification of a nominal model

To start the estimation of a set of models $\mathcal{P}_i$, first a nominal model $\hat{P}_i$ is estimated. As
mentioned in Section 6.1.3, the control relevant estimation of a set of models is split
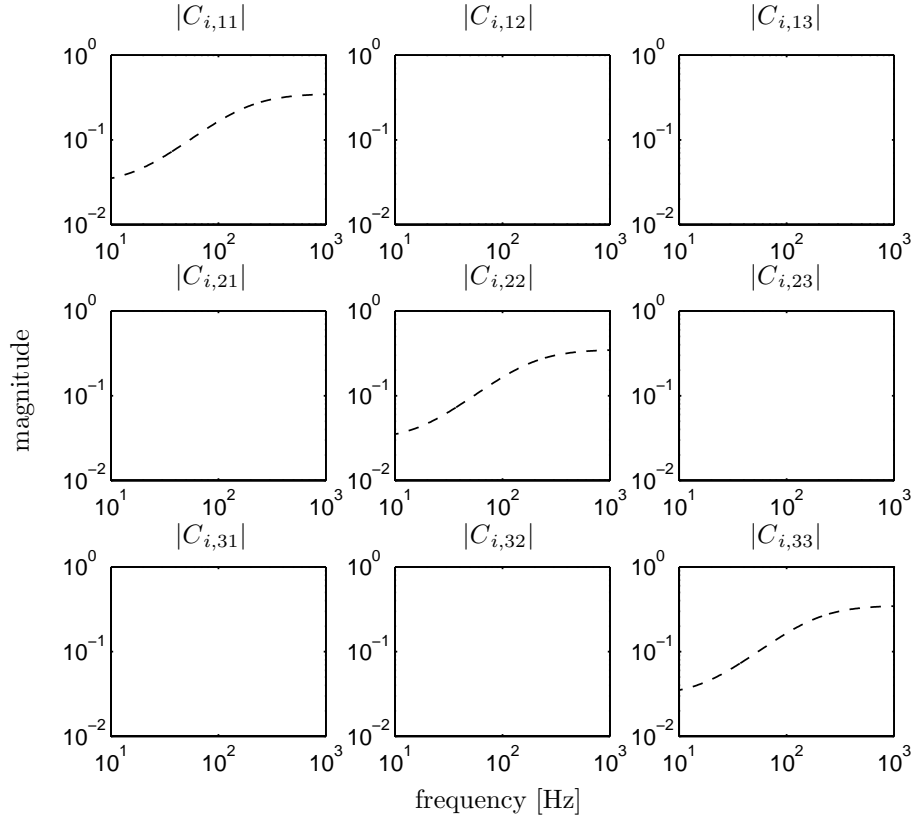
Fig. 8.12: Magnitude Bode plot of initial controller C_i used for closed loop experiments.

up in two parts. Using the fractional approach presented in chapter 4 to deal with data obtained under closed-loop conditions, first a nominal model $\hat{P}_i$ is estimated that should satisfy $T(\hat{P}_i, C_i) \in R\mathcal{H}_\infty$. Subsequently, using the same fractional approach, a model uncertainty is estimated to complete the set of models.

Access to a right coprime factorization

The fractional approach to closed-loop identification presented in chapter 4 of this thesis begins with the construction of a filter F to access a *rcf* of the plant P_o. For that purpose, the data $\{u, y\}$ needs to be filtered by a filter F that should satisfy the conditions mentioned in Lemma 4.3-5. With the knowledge of the controller C_i used during the experiments, the filtered signal x given in (4.24) can be generated. Subsequently, a *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o can be accessed, as mentioned in Corollary 4.3-10.

According to Corollary 4.3-9, an auxiliary model P_x can be used to construct such a filter F to access a *rcf* of the plant P_o. With the knowledge of the existing 18th order model depicted in Figure 8.8, such an auxiliary model P_x is readily available. The 18th order model P_x satisfies $T(P_x, C_i) \in R\mathcal{H}_\infty$ and can be used to construct the filter F. Computing a *nrcf* (N_x, D_x) of the model P_x and using the knowledge of the controller C_i, a filter F and the signal x can be constructed. It should be noted that the signal x satisfies

$$x = F \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} r_2 \\ r_1 \end{bmatrix}.$$

With the signals r_1 and r_2 specified as periodic reference signals, the signal x is again a periodic input signal.

As mentioned in Proposition 4.3-11, the fractional approach enables one to consider the map from the closed-loop data x to $col(y, u)$ as an open-loop map. This open-loop map enables the access to a *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o. Due to the periodic nature of the signals and the amount of data obtained from the high speed mechanical positioning mechanism, it is obvious and convenient to represent the data in an estimated frequency response.

Due to the open-loop nature of the map from x to $col(y, u)$, a spectral estimate can be obtained straightforwardly. An amplitude Bode plot of the spectral estimate $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ of the *rcf* $(N_{o,F}, D_{o,F})$ of the plant P_o that is accessible on the basis of the filter F being constructed is depicted in Figure 8.13. The ordering of the frequency responses is in accordance with the following remark.

Remark 8.4-1 *The map from x to $col(y, u)$ is given by the rcf $(N_{o,F}, D_{o,F})$ where $N_{o,F}$ and $D_{o,F}$ are stacked row wise, see Corollary 4.3-10. Since the plant P_o representing the wafer stage has three inputs and three outputs, the stacked configuration of the rcf $(N_{o,F}, D_{o,F})$ has three inputs and six outputs. As a result, the top nine amplitude Bode plots in Figure 8.13 indicate a spectral estimate of $\hat{N}_{o,F}(\omega_j)$ of the rcf $N_{o,F}$, while the bottom nine transfer functions denote a spectral estimate $\hat{D}_{o,F}(\omega_j)$ of $D_{o,F}$.*

For illustration purposes, only the amplitude Bode plot of the frequency domain data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ is given in Figure 8.13. The complex frequency response data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ can be used to estimate a factorization of the nominal model $\hat{P}_i$.

Estimating a nominal factorization

For the control relevant estimation of a factorization $(\hat{N}_i, \hat{D}_i)$ of the nominal model $\hat{P}_i$, the optimization stated in (6.8) must be performed. As motivated in Section 6.2.4, frequency domain data is helpful for the approximation of the $\mathcal{H}_\infty$ norm criterion by a point wise evaluation of (6.8).
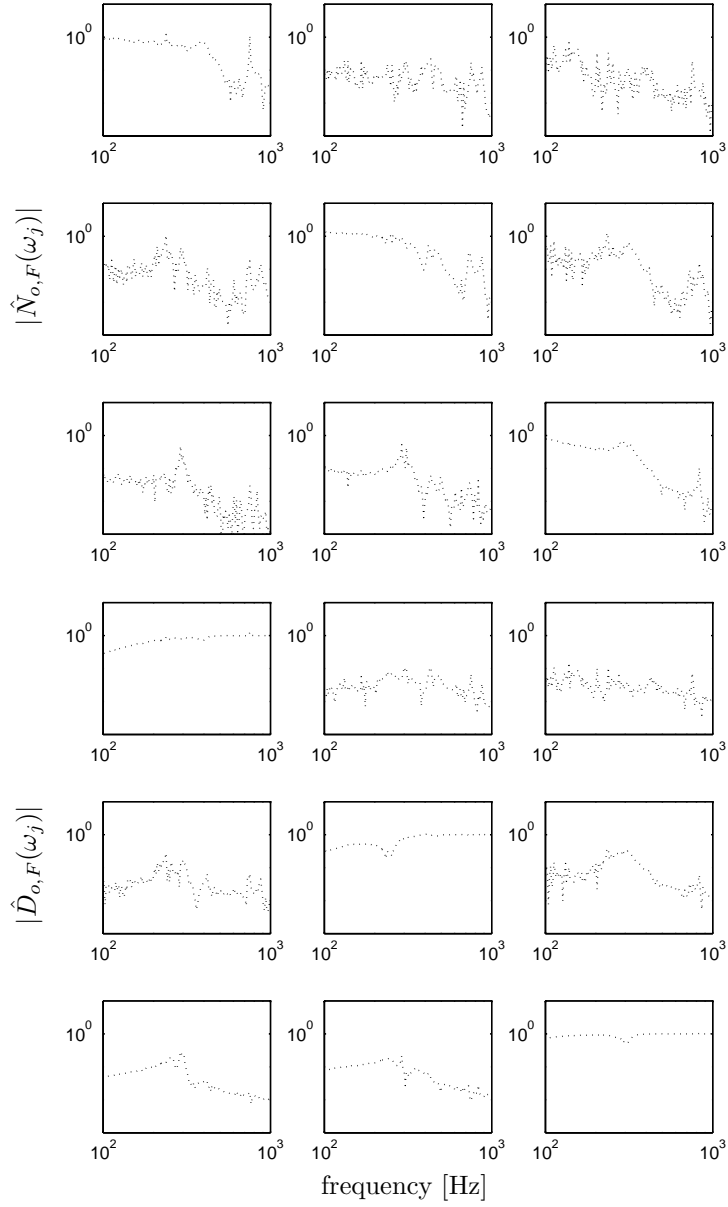
Fig. 8.13: Magnitude Bode plot of experimentally obtained frequency response $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ of the rcf $(N_{o,F}, D_{o,F})$ of the plant P_o.

However, Assumption 6.2-9 states that the (possibly disturbed) frequency domain samples $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ must be chosen over a frequency grid Ω which is dense enough to represent the frequency response of the *rcf* $(N_{o,F}, D_{o,F})$. With the frequency grid of 200 data points in the range from approximately 14 Hz till 1.2 kHz this condition is assumed to be satisfied for the wafer stage. Using the point wise evaluation of the maximum singular value along the frequency grid Ω, the optimization stated in (6.8) is replaced by the optimization problem given in (6.33). For the optimization, use is made of the frequency response estimate $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ given in Figure 8.13, the knowledge of controller C_i used during the experiments and the weighting functions U_2, U_1 given in (8.2), (8.3) and (8.8).

Even for a fixed filter F, (6.33) is non-convex min-max optimization problem that requires a sophisticated numerical optimization. A min-max optimization routine found in commercially available software (MatLab 1994) is used for that purpose. To ensure stability of the nominal *rcf* being estimated during the optimization, the factorization $(N(\theta), D(\theta))$ to be estimated is parametrized accordingly to the parametrization discussed in Section 6.2.3.

To find an initial estimate of a *rcf* to start up the non-linear min-max optimization, first a straightforward least-squares curve fit routine is applied to the frequency domain data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$. During the least-squares curve fitting, a polynomial based parametrization is used to parametrize the *rcf* $(N(\theta), D(\theta))$. As already indicated in (6.22), the *rcf* $(N(\theta), D(\theta))$ is parametrized according to

$$\begin{bmatrix} N(q, \theta) \\ D(q, \theta) \end{bmatrix} = B(q^{-1}, \theta) A^{-1}(q^{-1}, \theta)$$

and the iterative procedure of least-squares optimization presented in de Callafon *et al.* (1996*a*) is used to solve the least-squares curve fitting problem. Computational details can also be found in Appendix B of this thesis. Subsequently, the stable estimate $(N(\hat{\theta}), D(\hat{\theta}))$ of the least squares optimization is reparametrized in the state space parametrization mentioned in Section 6.2.3 and used as an initial estimate for the non-linear min-max optimization.

The final result of the optimization has been depicted in the amplitude Bode plots of Figure 8.14. For reasons of completeness, the phase bode plot are also given in Figure 8.15. For comparison with the data being measured, the plots of the frequency response data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ have also been given in these figures.

The resulting nominal *rcf* $(\hat{N}_i, \hat{D}_i)$ being estimated is a 27th order factorization and the nominal model $\hat{P}_i = \hat{N}_i \hat{D}_i^{-1}$ satisfies $T(\hat{P}_i, C_i) \in R\mathcal{H}_\infty$. The order of the nominal *rcf* is determined by estimating several *rcf*'s of different order. The 27th order *rcf* was able to find a low value of the criterion (6.33) at the price of finding a reasonably low complexity model.

The Bode plots depicted in Figure 8.14 and Figure 8.15 give insight how the estimated factorization $(\hat{N}_i, \hat{D}_i)$ compares with the data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$. Addition-
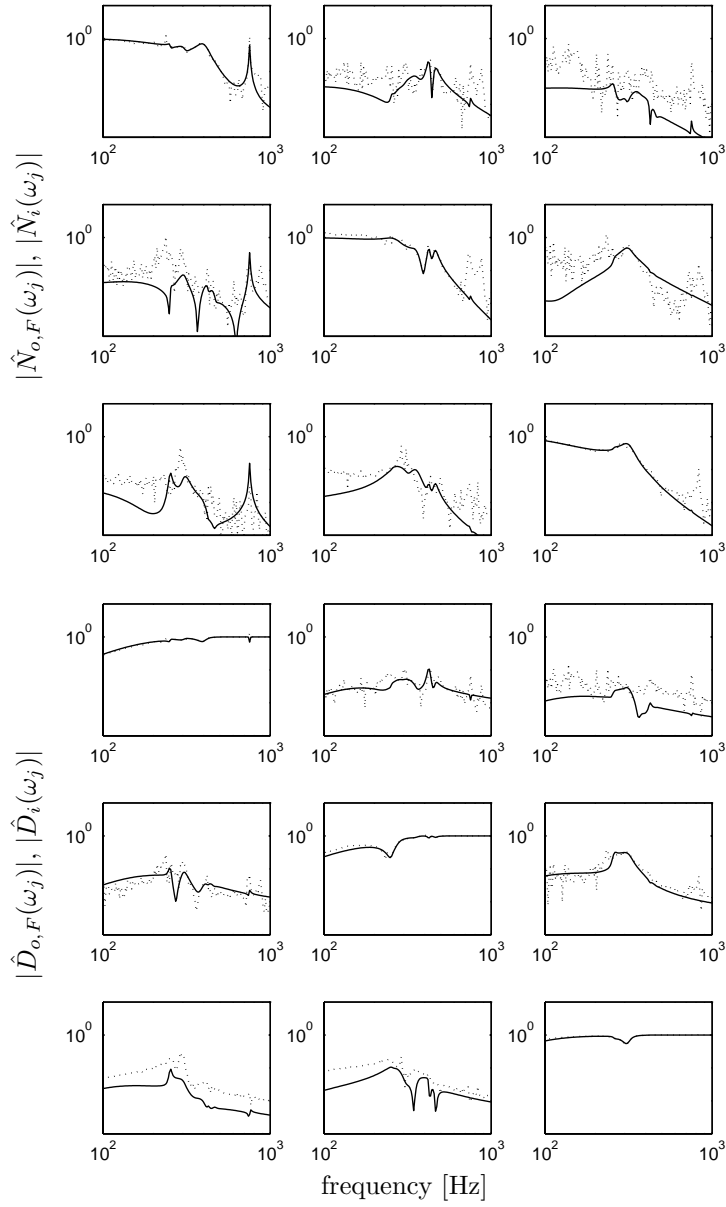
Fig. 8.14: Magnitude Bode plot of spectral estimate $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ $(\cdots)$ and 27th order nominal factorization $(\hat{N}_i(\omega_j), \hat{D}_i(\omega_j))$ (—).
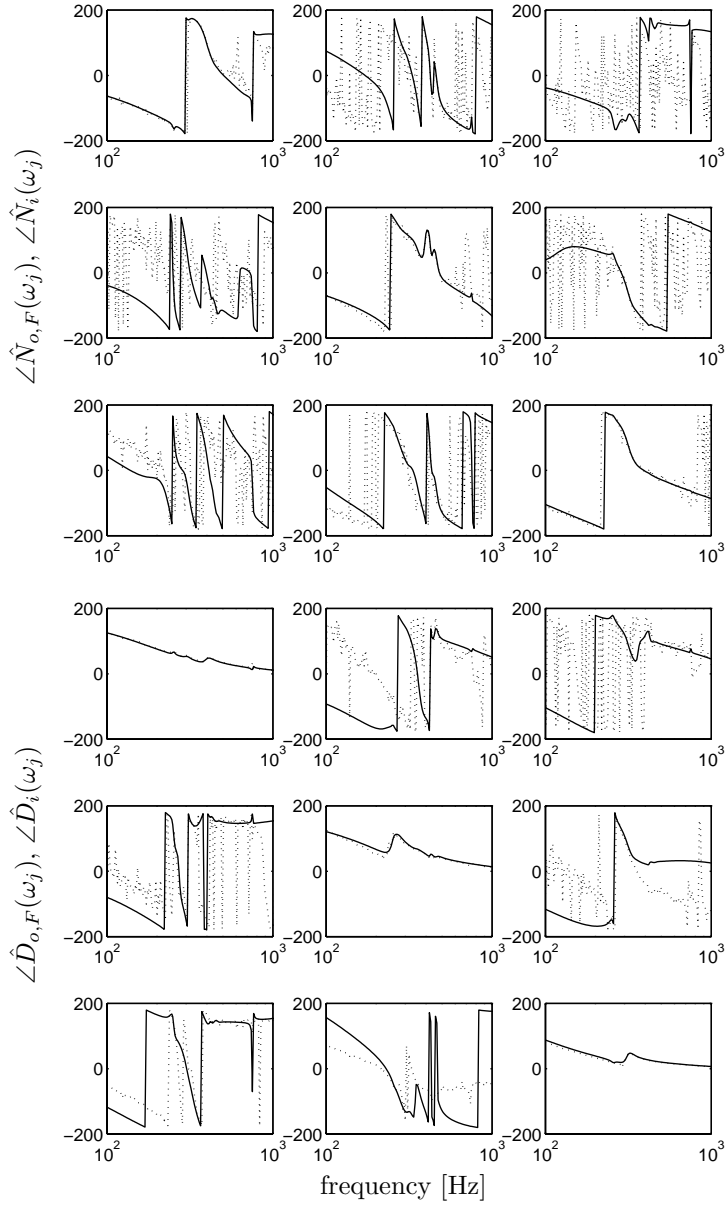
Fig. 8.15: Phase Bode plot of spectral estimate $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ $(\cdots)$ and 27th order nominal factorization $(\hat{N}_i(\omega_j), \hat{D}_i(\omega_j))$ (—).

ally, a Bode plot of the nominal model $\hat{P}_i = \hat{N}_i\hat{D}_i^{-1}$ being estimated can be compared with the frequency domain data $\hat{G}(\omega_j)$. The frequency domain data $\hat{G}(\omega_j)$, as given in Figure 8.7, can be recomputed from the data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ via

$$\hat{G}(\omega_j) = \hat{N}_{o,F}(\omega_j)\hat{D}_{o,F}^{-1}(\omega_j).$$

To complete the comparison of the frequency domain data with the nominal factorization being estimated, a Bode plot of $\hat{P}(\omega_j)$ and $\hat{G}(\omega_j)$ has been depicted in Figure 8.16.

From Figure 8.14 and Figure 8.15 it can be observed that the diagonal elements of the data of $\hat{N}_i(\omega_j)$ and $\hat{D}_i(\omega_j)$ has been fitted reasonably well. However, the non-diagonal elements have not been fitted very well. This is due to the fact that the non-diagonal elements are smaller in amplitude. Furthermore, the non-diagonal elements currently do not have a significant contribution to the criterion

$$\bar{\sigma}\left\{U_2(\omega_j)\left(\begin{bmatrix} \hat{N}_{o,F}(\omega_j) \\ \hat{D}_{o,F}(\omega_j) \end{bmatrix} - \begin{bmatrix} N(\theta,\omega_j) \\ D(\theta,\omega_j) \end{bmatrix}\right)F(\omega_j)\begin{bmatrix} C_i(\omega_j) & I \end{bmatrix}U_1(\omega_j)\right\} \quad (8.10)$$

for which the maximum over the frequency grid

$$\Omega = (\omega_1,\ldots,\omega_j,\ldots,\omega_l), \text{ with } 0 \le \omega_1 < \cdots < \omega_j < \cdots < \omega_l \le \pi$$

needs to be minimized according to (6.33). As mentioned before, the controller C_i used in (6.33) or (8.10) denotes the controller currently implemented on the plant P_o. At the current stage, the controller C_i only consists of the three SISO parallel placed PID controllers depicted in Figure 8.12. Therefore, the controller C_i has only diagonal terms that influence the optimization of (6.33), giving an explanation to the good fit of the diagonal terms of the data.

To give an indication of the result of the 27th order nominal estimate $(\hat{N}_i, \hat{D}_i)$ in terms of the criterion mentioned in (8.10), a plot of (8.10) has been depicted in Figure 8.17. With the frequency domain data $(\hat{N}_{o,F}(\omega_j), \hat{D}_{o,F}(\omega_j))$ and the 27th order nominal estimate, the criterion can be evaluated along the frequency grid Ω. From Figure 8.17 it can be observed that the optimization (6.33) has tried to minimize the maximum of (8.10) over the frequency grid Ω.

8.4.3   Completing the set of models

The set of models, used to reflect the limited knowledge of the plant P_o, is structured according to

$$\mathcal{P}(\hat{N}, \hat{D}, N_c, D_c, \hat{V}, \hat{W}) := \{P \mid P = (\hat{N} + D_c\Delta_R)(\hat{D} - N_c\Delta_R)^{-1}$$

$$\text{with } \Delta_R \in R\mathcal{H}_\infty \text{ and } \Delta := \hat{V}\Delta_R\hat{W} \text{ satisfies } \|\Delta\|_\infty < \gamma^{-1}\}$$
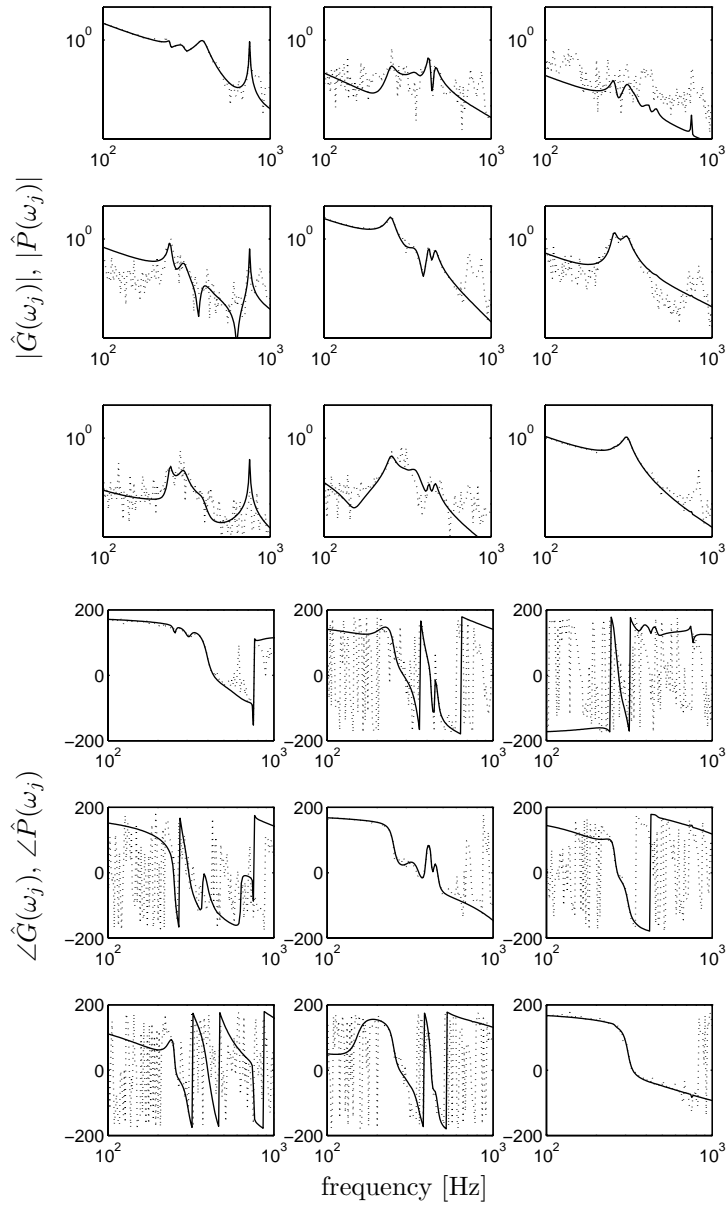
Fig. 8.16: Magnitude Bode plot (top nine plots) and phase Bode plot (bottom nine plots) of spectral estimate $\hat{G}(\omega_j) = \hat{N}_{o,F}(\omega_j)\hat{D}_{o,F}^{-1}(\omega_j)$ ($\cdots$) and 27th order nominal model $\hat{P}_i(\omega_j) = \hat{N}_i(\omega_j)\hat{D}_i^{-1}(\omega_j)$ (—).
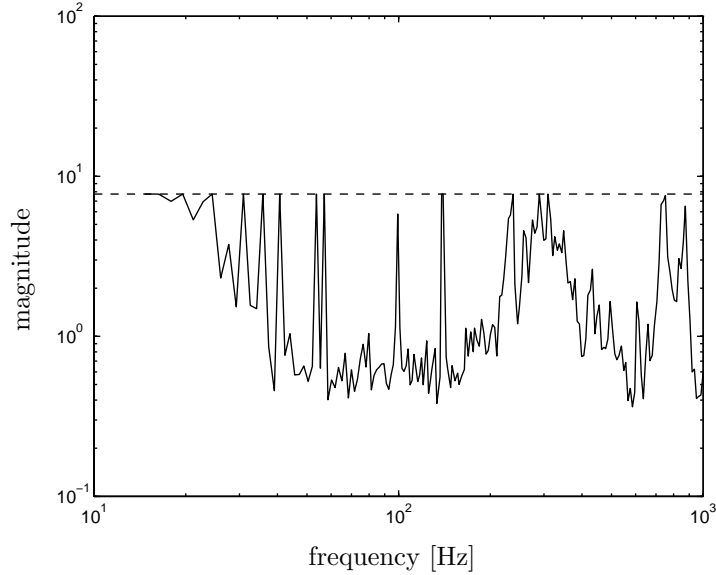
Fig. 8.17: Evaluation of criterion (8.10) over the frequency grid Ω.

and has been mentioned previously in Definition 5.2-4. As mentioned in Remark 5.2-7, a nominal factorization $(\hat{N}, \hat{D})$ and weighting functions $(\hat{V}, \hat{W})$ are needed to complete the set of models. The pair (N_c, D_c) is assumed to be known and is found by computing a *rcf* of the (known) controller that is being used in the feedback connection of Figure 8.12 during the closed-loop experiments.

The results on the estimation of a nominal factorization $(\hat{N}_i, \hat{D}_i)$ of a nominal model $\hat{P}_i = \hat{N}_i \hat{D}_i^{-1}$ that satisfies $T(\hat{P}_i, C_i) \in R\mathcal{H}_\infty$ have been presented in the previous section. The factorization $(\hat{N}_i, \hat{D}_i)$ is used as the nominal factorization for the set of models $\mathcal{P}_i$. Hence, to complete the characterization of this set of models $\mathcal{P}_i$, frequency dependent weighting functions $(\hat{V}_i, \hat{W}_i)$ that upper bound the stable model perturbation Δ_R have to be determined. As indicated in the estimation procedure depicted in Section 6.3, first a frequency dependent (non-parametric) upper bound $\bar{\Delta}_R(\omega)$ for the unknown, but stable and bounded model perturbation $\Delta_R(e^{i\omega})$ is being estimated. Subsequently, low order frequency dependent weighting functions $(\hat{V}, \hat{W})$ are determined that upper bound frequency dependent (non-parametric) upper bound $\bar{\Delta}_R(\omega)$. The results are presented next.

Upper bound for the model perturbation

To estimate a frequency dependent upper bound $\bar{\Delta}_R(\omega)$ for the unknown, but stable and bounded model perturbation Δ_R, the signals x and z mentioned in Corollary 6.3-1 are needed. These signals x and z can be obtained by filtering of closed-loop measured

signals. The filtering is given by the filter operations

$$
\begin{aligned}
x &:= (\hat{D}_i + C_i \hat{N}_i)^{-1} \begin{bmatrix} C_i & I \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \\
z &:= (D_{c,i} + \hat{P}_i N_{c,i})^{-1} \begin{bmatrix} I & -\hat{P}_i \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix}
\end{aligned}
\tag{8.11}
$$

and the resulting signals x and z yield access to the stable model perturbation Δ_R via

$$
z = \Delta_R x + v \tag{8.12}
$$

where v is uncorrelated with x. According to Corollary 6.3-1, Δ_R is then given by

$$
\Delta_R = D_{c,i}{}^{-1}(I + P_o C_i)^{-1}(P_o - \hat{P}_i)\hat{D}_i \tag{8.13}
$$

in which the (unknown) plant P_o, the nominal model $\hat{P}_i$ and the controller C_i play a role. Additionally, it can be observed that both the factorization $(\hat{N}_i, \hat{D}_i)$ of the nominal model and the factorization $(N_{c,i}, D_{c,i})$ of the controller play a role in "shaping" the model perturbation Δ_R. The same information of these factorizations can also be found in the construction of the signals x and z in (8.11).

Additionally, the knowledge of the controller C_i and a *rcf* $(N_{c,i}, D_{c,i})$ of C_i is needed to construct respectively the signals x and z. As the factorization of the controller is not unique, a choice is made to compute a *nrcf* $(N_{c,i}, D_{c,i})$ of the controller C_i. Subsequently, the signals x and z in (8.11) can be created and a spectral estimate of Δ_R in (8.13) can be computed.

The spectral estimate of Δ_R is displayed in Figure 8.18 as a dotted line for reference purposes. It should be noted that the exact knowledge of Δ_R is not being used to characterize a set of models $\mathcal{P}$. The aim is to estimate a frequency dependent upper bound $\bar{\Delta}_R(\omega)$ of Δ_R that guarantees $P_o \in \mathcal{P}$ to hold with a specific probability α. This is done by the probabilistic uncertainty estimation routine of Hakvoort (1994) that has been summarized in Section 6.3.2.

As Δ_R is a multivariable transfer function, the uncertainty estimation routine is performed three times for a row-wise evaluation of Δ_R. An upper bound on the first row of Δ_R is found by estimating upper bounds for the three transfer functions from the three dimensional input signal x to the first signal in the three dimensional output signal z. Analogously, upper bounds for the transfer functions in the second and third row of Δ_R are found, by considering respectively the second and the third signal in the three dimensional output signal z. The discussion below presents the procedure being followed for each row-wise evaluation of Δ_R, but the final result consists of nine different frequency dependent upper bounds for the transfer functions in Δ_R.

Additional *prior* information on Δ_R must be introduced in order to be able to estimate a frequency dependent upper bound $\bar{\Delta}_R(\omega)$ of Δ_R. Similar to the listing presented in Section 6.3.2, the following *prior* information is used.
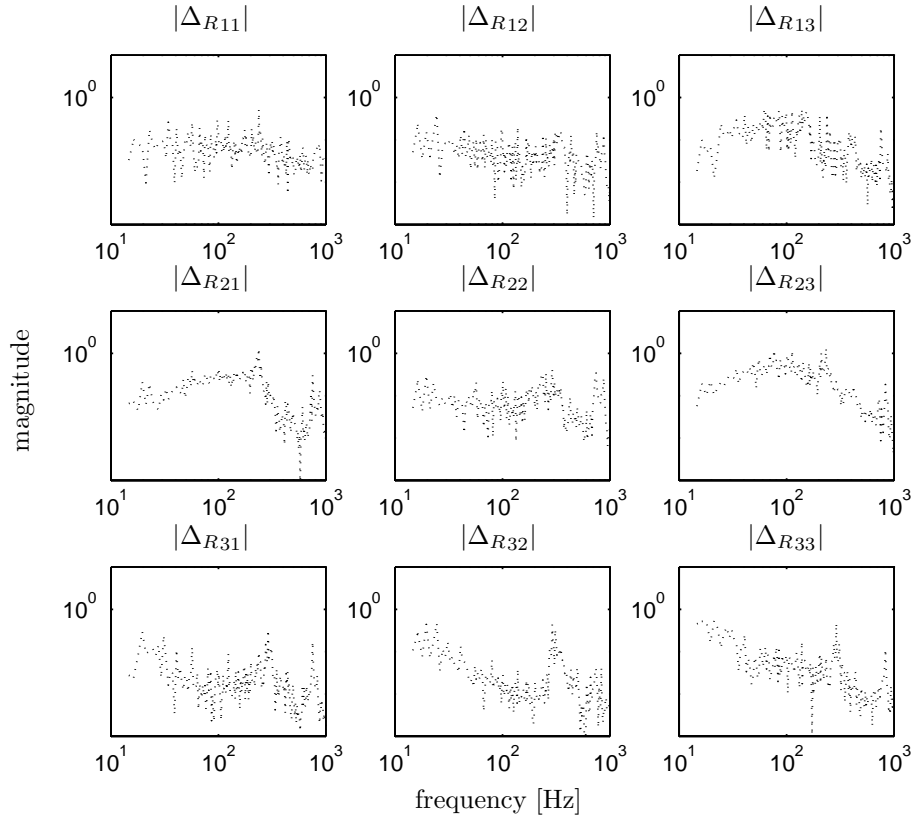
Fig. 8.18: Magnitude Bode plot of spectral estimate of model perturbation Δ_R given in (8.13).

- To be able to write Δ_R in terms of a series expansion

$$\Delta_R(q) = \sum_{k=0}^{\infty} R_k B_k(q)$$

and to gain information on the degree of stability of Δ_R, first a (high order) LS-estimation is performed on the spectral estimate of Δ_R depicted in Figure 8.18. The LS-estimation is performed with the frequency identification techniques presented in Appendix B and provides information on the poles of Δ_R. Subsequently, the information obtained on the dynamics of Δ_R can be used to construct a set of orthonormal basis functions $B_k(q)$ for $k = 0, \ldots, \infty$ (Heuberger et al. 1995). Additionally, an an upper bound $\bar{R}_k$

$$|R_k| \leq \bar{R}_k, \text{ for } k = 0, \ldots, \infty$$

for the coefficients R_k is obtained. Using these estimates, the *prior* information needed

$$\bar{R}_k \leq M\rho^k, \ \forall k > k^\star$$

can be formulated in terms of the parameters $M \geq 0$ and $\rho < 1$.

- The *prior* assumption

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

needed to bound the effects of initial conditions, is readily obtained. As indicated in Section 8.3.2, the reference signals r_1 and r_2 are defined as periodic signals. Using the algebraic relation $r_1 + Cr_2 = u + Cy$ mentioned in Corollary 3.2-4, it can be seen that x in (8.11) is also a periodic signal. Due to the periodic nature of the input signal, the effect of initial conditions can be eliminated and $\bar{x}$ is set to 0.

- As mentioned above, a probabilistic uncertainty bounding identification is used. For that purpose, the assumption on the noise v in (8.12) or (6.37) is in accordance with Assumption 6.3-2.

As mentioned in Section 6.3, for the estimation of (probabilistic) uncertainty bounds, the intermediate set of models $\mathcal{S}$ in (6.45) is constructed. The intermediate set $\mathcal{S}$ is determined by a bound β on the first derivative of the frequency response of Δ_R and a set of convex (probabilistic) frequency response regions $\mathcal{P}(\omega_j)$ defined along a frequency grid Ω. With the above mentioned *prior* information and the results mentioned in Lemma 6.3-4 and Lemma 6.3-5, a bound β on the first derivative of the frequency response of Δ_R and a set of rectangular frequency response regions $\mathcal{P}(\omega_j)$ is determined. The frequency response regions $\mathcal{P}(\omega_j)$ are computed around a ORT-FIR based parametric estimate $\hat{\Delta}_R(e^{i\omega_j})$ that acts as a "carrier" of the uncertainty regions $\mathcal{P}(\omega_j)$. For the computation of the regions $\mathcal{P}(wj)$, the same frequency grid Ω is used, as mentioned during the experiment design in Section 8.3.2 and consists of 200 points distributed (approximately logarithmically) between 10 Hz and 1200 Hz.

To give an indication of the convex frequency response regions $\mathcal{P}(\omega_j)$ being estimated, in Figure 8.19 the regions $\mathcal{P}(\omega_j)$ for the $(1,1)$ element of the model perturbation Δ_R have been depicted. To produce a clear picture, only 25 frequency points ω_j from the frequency grid Ω have been used to generate the plot given in Figure 8.19. It should be noted that regions $\mathcal{P}(\omega_j)$ have been estimated for all frequency points $\omega_j \in \Omega$ and for all nine elements of the model perturbation Δ_R. A plot of the remaining frequency response regions $\mathcal{P}(\omega_j)$ is omitted, as similar figures are obtained.

The real and imaginary bounds of the rectangular shaped regions $\mathcal{P}(\omega_j)$ in Figure 8.19 hold with a probability of 99% each. It can be seen from this figure that, similar to Figure 6.1, rectangular shaped regions are estimated for each frequency
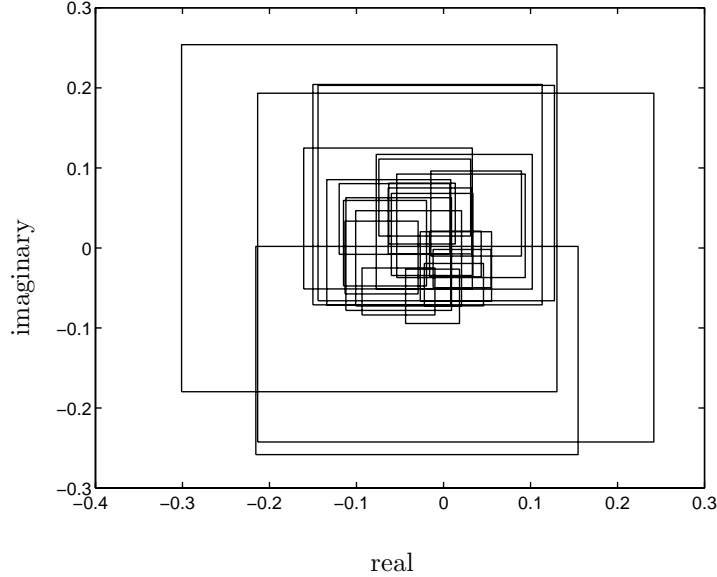
Fig. 8.19: Estimated rectangular uncertainty regions $\mathcal{P}(\omega_j)$ for the $(1,1)$ element of $\Delta_R(\mathrm{e}^{\,i\omega})$.

point ω_j within the frequency grid Ω. Furthermore, most of the regions $\mathcal{P}(\omega_j)$ contain the origin, indicating that the model perturbation $\Delta_R(\mathrm{e}^{\,i\omega})$ is small; the real and imaginary uncertainty bounds of $\mathcal{P}$ at a frequency $\omega_j \in \Omega$ are larger than $\Delta_R(\mathrm{e}^{\,i\omega_j})$ itself.

Once β and $\mathcal{P}(\omega_j)$ for each of the elements of Δ_R are available, a frequency dependent upper bound $\bar{\Delta}_R(\omega)$ for the model perturbation Δ_R in (8.13) can be computed. This is done by evaluating the corner points of the rectangular frequency response regions $\mathcal{P}(\omega_j)$ and computing the largest distance to the origin. This yields an amplitude bound $\bar{\Delta}_R(\omega_j)$ at the frequency points $\omega_j \in \Omega$. Subsequently, the bound β on the first derivative of the frequency response of Δ_R can be used to interpolate to form a continuous (non-parametric) frequency dependent error bound $\bar{\Delta}_R(\omega)$. This gives a clear overview of the frequency dependent character of the model uncertainty or model perturbation Δ_R.

The results on the estimation of a frequency dependent (non-parametric) upper bound $\bar{\Delta}_R(\omega)$ for the unknown, but stable and bounded model perturbation $\Delta_R(\mathrm{e}^{\,i\omega})$ have been depicted in Figure 8.20. For reference purposes, the spectral estimate of Δ_R has also been depicted in this figure.

It should be noted that $\bar{\Delta}_R(\omega)$ is an estimated and guaranteed upper bound for the model perturbation $\Delta_R(\mathrm{e}^{\,i\omega})$ that holds with a certain probability, provided that the *prior* information being introduced is consistent. The probability is determined
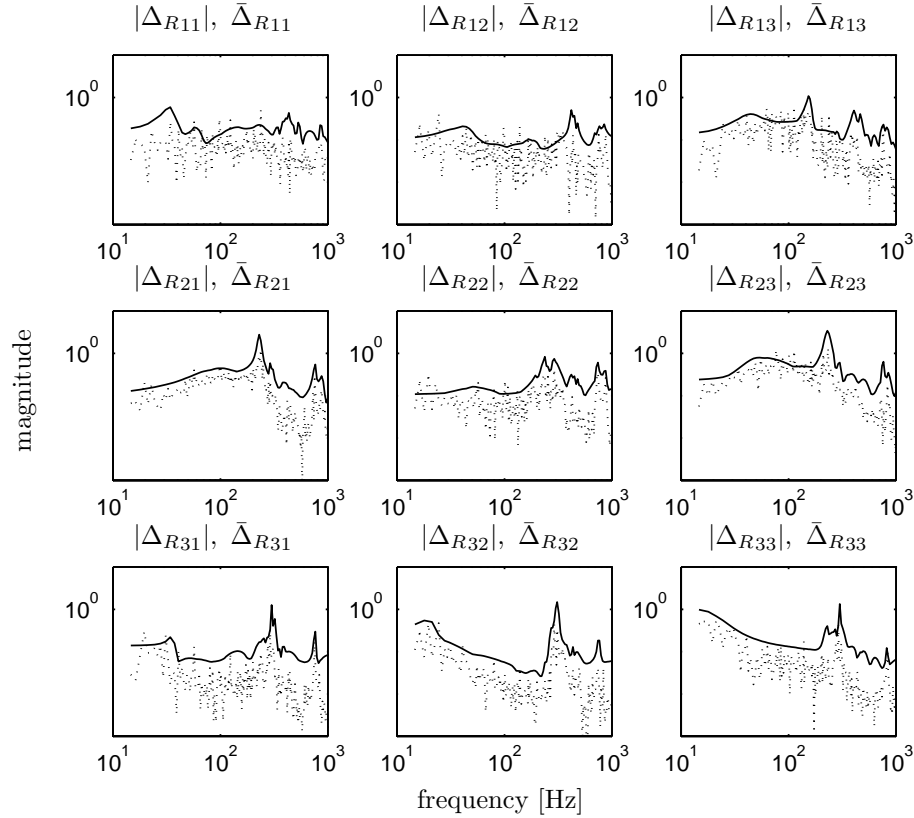
Fig. 8.20: Magnitude Bode plot of spectral estimate of model perturbation Δ_R given
           in (8.13) ($\cdots$) and estimated probabilistic frequency dependent amplitude
           upper bound $\bar{\Delta}_R(\omega)$ (—).

by the probability chosen when computing the convex frequency response regions
$\mathcal{P}(\omega_j)$ as depicted in Figure 8.19. The (validated) prior information has been listed
on Page 227.

Parametric upper bound

It can be seen from Figure 8.20 that the uncertainty estimation routine has found a
frequency dependent upper bound $\bar{\Delta}_R(\omega)$ that gives information on the size and shape
of the model perturbation $\Delta_R(e^{i\omega})$. This information can be used to find *parametric*
frequency dependent weighting functions $(\hat{V}_i, \hat{W}_i)$ that upper bound the estimated
model perturbation $\bar{\Delta}_R(\omega)$. With the estimation of these parametric weighting func-

tion, the set of models $\mathcal{P}_i$ given by

$$\mathcal{P}_i(\hat{N}_i, \hat{D}_i, N_{c,i}, D_{c,i}, \hat{V}_i, \hat{W}_i) = \{P \mid P = (\hat{N}_i + D_{c,i}\Delta_R)(\hat{D}_i - N_{c,i}\Delta_R)^{-1}$$

$$\text{with } \Delta_R \in R\mathcal{H}_\infty \text{ and } \Delta_i := \hat{V}_i\Delta_R\hat{W}_i \text{ satisfies } \|\Delta_i\|_\infty < \gamma_i^{-1}\} \tag{8.14}$$

can be completed to perform performance robustness analysis and the design of an improved robust controller C_{i+1}. For that purpose, the set of models $\mathcal{P}_i$ is rewritten into the standard LFT representation

$$\mathcal{P}_i = \{P \mid P = \mathcal{F}_u(Q, \Delta) \text{ with } \Delta \in R\mathcal{H}_\infty, \ \|\Delta\|_\infty < \gamma^{-1} \text{ and}$$

$$Q = \begin{bmatrix} \hat{W}_i^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix} \begin{bmatrix} \hat{V}_i^{-1} & 0 \\ 0 & I \end{bmatrix} \text{ with} \tag{8.15}$$

$$\begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix} = \begin{bmatrix} \hat{D}_i^{-1}N_{c,i} & \hat{D}_i^{-1} \\ (D_{c,i} + \hat{P}_iN_{c,i}) & \hat{P}_i \end{bmatrix}$$

as mentioned in Corollary 5.2-8. The entries in the coefficient matrix Q in (8.15) are the same entries mentioned in (8.14) to complete the set of models $\mathcal{P}_i$.

It should be noted that the (probabilistic) uncertainty bounding identification has resulted in a frequency dependent upper bound $\bar{\Delta}_{Rij}(\omega)$ for each element (i,j) of the model perturbation Δ_R. It was already observed in Section 5.2.5 that such detailed information can be used to find a scalar stable and stably invertible weighting filter $\hat{V}_{ij}$ that bounds each element (i,j) of $\bar{\Delta}_R(\omega)$ separately via

$$\|\hat{V}_{ij}\bar{\Delta}_{Rij}\|_\infty \leq 1. \tag{8.16}$$

As indicated in Remark 5.2-6, the notation involving the pair $(\hat{V}_i, \hat{W}_i)$ was used to keep track of the (possible) multivariable nature of the model perturbation Δ_R. However, in case detailed information on the upper bound $\bar{\Delta}_{Rij}(\omega)$ of each element of Δ_R is available, it is beneficial to represent the model perturbation Δ_R and the weighting functions in a diagonal form.

As previously indicated in Section 5.2.5, the unweighted coefficient matrix $\bar{Q}$ in (8.15) can be easily modified to account for a diagonal form of the model perturbation Δ_R. This modification is found by multiplying $\bar{Q}_{11}$ with two scaling matrices T_1 and T_2 to obtain

$$\bar{Q} = \begin{bmatrix} T_2\bar{Q}_{11}T_1 & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix}$$

as the unweighted coefficient matrix $\bar{Q}$. Since $\bar{\Delta}_R(\omega)$ consists of 9 scalar elements (3×3), the scaling matrices T_1 and T_2 are given by

$$T_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \ T_2 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}^T$$

to be able to deal with the 9 elements of Δ_R in diagonal form.

Compatible with the diagonal form $\mathrm{diag}\{\bar{\Delta}_{Rij}(\omega)\}$ of the model perturbation $\bar{\Delta}_R(\omega)$, only a stable and stably invertible diagonal weighting filter $\hat{V}_i$ needs to be estimated and $\hat{W}_i$ can be omitted. In this case, the weighting filter $\hat{V}_i$ has a similar diagonal form and is denoted by $\mathrm{diag}\{\hat{V}_{ij}\}$. The diagonal elements $\hat{V}_{ij}$ are the scalar stable and stably invertible weighting filters that bound each element $\bar{\Delta}_{Rij}(\omega)$ separately.

Estimating the scalar stable and stably invertible weighting filters $\hat{V}_{ij}$ is done via a spectral overbounding (the LPSOF algorithm) discussed in Section 6.3.3. The frequency dependent data $\bar{\Delta}_R(\omega)$ obtained from the uncertainty bounding identification is used as an input for the LPSOF algorithm and nine different weighting functions $\hat{V}_{ij}$ are estimated to find a parametric bound for each $\bar{\Delta}_{Rij}(\omega)$ separately.

It is tempting to estimate high order weighting functions $\hat{V}_{ij}$ in order to get a tight parametric upper bound of the frequency domain data $\bar{\Delta}_{Rij}(\omega)$. However, the complexity of the weighting functions $\hat{V}_{ij}$ directly enters into the coefficient matrix

$$
Q = \begin{bmatrix} T_2\bar{Q}_{11}T_1 & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathrm{diag}\{\hat{V}_{ij}\}^{-1} & 0 \\ 0 & I \end{bmatrix} \tag{8.17}
$$

that is used in the LFT representation of the set of models. Computation of an enhanced performance and robust controller is done on the basis of this LFT representation. As motivated before, computational problems and the application of high order controllers can be avoided by limiting the order of the coefficient matrix Q. This has been the motivation for estimating a relatively low complexity factorization $(\hat{N}_i, \hat{D}_i)$. Along the same lines, the order of the weighting functions $\hat{V}_{ij}$ also need to be limited.

With the LPSOF algorithm, relatively low order, stable and stably invertible weighting function $\hat{V}_{ij}$ have been estimated to bound the nine elements of the frequency dependent upper bound $\bar{\Delta}_R(\omega)$. The results are plotted in the amplitude Bode diagram of Figure 8.21. The choice for the order of the different weighting functions $\hat{V}_{ij}$ is determined by inspecting the (minimum) value for $\delta(\theta)$ in (6.65) during the optimization used in the LPSOF algorithm. Additionally, the order of each weighting function $\hat{V}_{ij}$ has been limited to 8 to avoid computational complexities. With these order choices, the total order of the (diagonal) weighting function $\mathrm{diag}\{\hat{V}_{ij}\}$ becomes 48, being the sum of the orders of the scalar weighting functions $\hat{V}_{ij}$ computed with the LPSOF algorithm.

With the knowledge of a *nrcf* $(N_{c,i}, D_{c,i})$ of the controller C_i and the estimates of a nominal factorization $(\hat{N}_i, \hat{D}_i)$ and (diagonal) weighting filter $\mathrm{diag}\{\hat{V}_{ij}\}$, the set of models $\mathcal{P}_i$ has been completed.
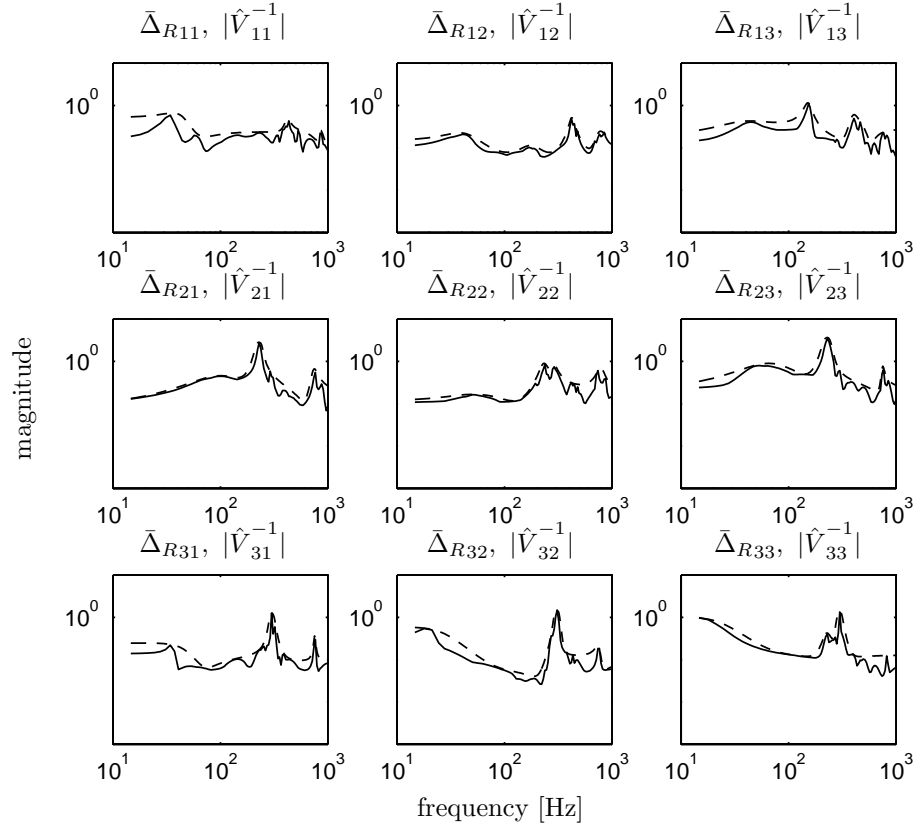
Fig. 8.21: Magnitude Bode plot of stable and stably invertible weighting filters $\hat{V}_{ij}(\mathrm{e}^{\,i\omega})$ (- -) and estimated probabilistic frequency dependent amplitude upper bound $\bar{\Delta}_R(\omega)$ (—).

8.5 Enhancing the Performance Robustly

8.5.1 Robust controller design

Given the estimated set of models $\mathcal{P}_i$, performance robustness can be evaluated and an enhanced robust controller can be designed. As such, the estimated set of models $\mathcal{P}_i$ is used to represent the (limited) knowledge of the plant P_o that is currently available. This knowledge is used to actually evaluate the performance of the controller C_i currently implemented on the plant P_o and to redesign the feedback control.

Performance evaluation

To be able to compare the performance of a newly designed controller C_{i+1} with the performance of the controller C_i, the performance of the controller C_i, implemented

on the plant P_o, should be evaluated *a posteriori*. As the plant P_o is unknown, the estimated set of models $\mathcal{P}_i$ is used to represent the knowledge currently available on the plant P_o to evaluate this performance.

It should be noted that stability robustness does not have to be evaluated for the controller C_i. The knowledge of the controller C_i has been used in the construction of the set of models $\mathcal{P}_i$. As indicated in Remark 5.3-7, the set of models $\mathcal{P}_i$ has been constructed in such a way that stability robustness is satisfied.

To characterize the performance of the controller C_i, the $\mathcal{H}_\infty$ norm

$$\|J(P, C_i)\|_\infty = \|U_2 T(P, C_i) U_1\|_\infty$$

needs to be evaluated for all models $P \in \mathcal{P}_i$. As previously indicated in (5.11), robust performance is satisfied when

$$\|J(P, C_i)\|_\infty = \|U_2 T(P, C_i) U_1\|_\infty \leq \gamma_i \tag{8.18}$$

for all $P \in \mathcal{P}_i$. The value of γ_i indicates the level of (robust) performance of the controller C_i, applied to all models $P \in \mathcal{P}_i$. With the plant $P_o \in \mathcal{P}_i$, the value of γ_i gives an indication of the performance of C_i. This performance level can be compared with the performance γ_{i+1} of a new controller C_{i+1} that will be specifically designed on the basis of the set of models $\mathcal{P}_i$ to achieve

$$\|U_2 T(P, C_{i+1}) U_1\|_\infty \leq \gamma_{i+1} < \gamma_i \ \forall P \in \mathcal{P}_i$$

as mentioned in (2.20).

For the evaluation of (8.18), the result mentioned in Lemma 5.3-2 and the structured singular value $\mu\{M\}$ is being used, see also Corollary 5.3-9. The weighting functions U_2 and U_1 are given in (8.2), (8.3) and (8.8). The set of models $\mathcal{P}_i$ is given in an LFT representation with the coefficient matrix Q given in (8.15) and (8.17). Subsequently, the structured singular value mentioned in Corollary 5.3-9 can be evaluated and the result has been plotted in Figure 8.22.

The structured singular value of $\mu\{M(\mathrm{e}^{i\omega})\}$ has been plotted point wise over the frequency domain range between 10 and 1000Hz. This range incorporates the closed-loop bandwidth and most of the closed-loop phenomena that might be of interest. The maximum value of $\mu\{M(\mathrm{e}^{i\omega})\}$, occuring in this frequency range, indicates the performance robustness margin γ_i. For the controller C_i, applied to the set of models $\mathcal{P}_i$, the value of γ_i is found to be 7.53, as can be seen from the dotted line in Figure 8.22.

As the weighting functions $\hat{V}_{ij}$ in (8.16) have been chosen to normalize the model uncertainty $\bar{\Delta}_{Rij}$, the controller C_i applied to the set of models $\mathcal{P}_i$ does not satisfy performance robustness. In that case, $\mu\{M(\mathrm{e}^{i\omega})\} < 1$ as the weighting functions are chosen to normalize ($\gamma = 1$) the uncertainty. Lowering the value of γ_i to improve performance (robustness) is the task of a newly to be designed controller C_{i+1}.
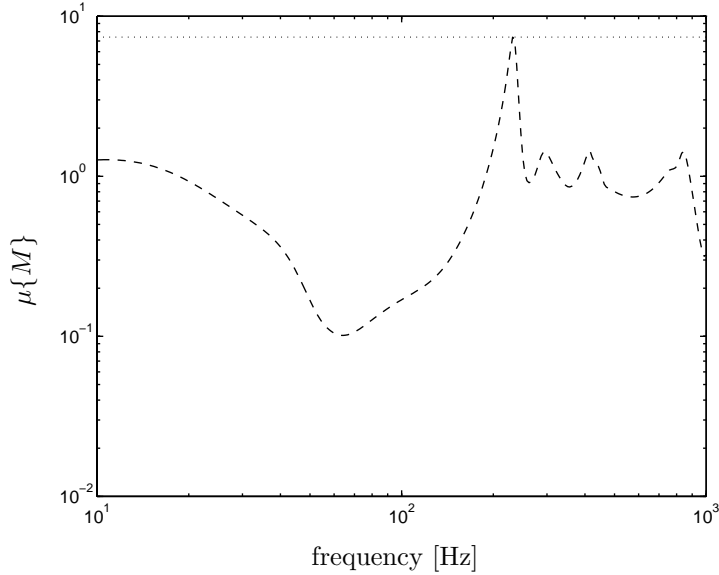
Fig. 8.22: Computation of structured singular value $\mu\{M(e^{i\omega})\}$ for the controller C_i applied to the set of models $\mathcal{P}_i$.

Design of controller

The initial controller C_i consists of three parallel PID controllers. As indicated in the previous section, the controller C_i can and needs to be improved. To redesign the controller, again the estimated set of models $\mathcal{P}_i$ is used.

For the design of the controller C_{i+1}, available software for $\mathcal{H}_\infty$ controller synthesis is being used. For that purpose, the feedback connection $\mathcal{T}(P,C)$ of a controller C with a model $P \in \mathcal{P}_i$ is represented in a lower fractional transformation $\mathcal{F}_l(G,C)$. The entries in the coefficient matrix G of the lower fractional transformation $\mathcal{F}_l(G,C)$ are the same variables needed to construct the set of models $\mathcal{P}_i$ and also contains the weightings U_2 and U_1 associated to the performance characterization. An expression for the transfer function of the coefficient matrix G can be found in Corollary 5.4-1. The controller C_{i+1} is computed via μ-synthesis using a D-K iteration.

In general, the complexity of a controller C_{i+1} generated by a μ-synthesis will be the higher that the complexity of the coefficient matrix G being used. As G contains all the entries of the set of models $\mathcal{P}_i$ and the weighting functions U_2 and U_1, the order of the controller C_{i+1} needs to be reduced significantly in order to be implementable on the DSP. For that purpose, closed-loop reduction tools (Ceton *et al.* 1993, Wortelboer 1993) are used to reduce the order of the designed controller C_{i+1} to an acceptable complexity for implementation, without hardly sacrificing any
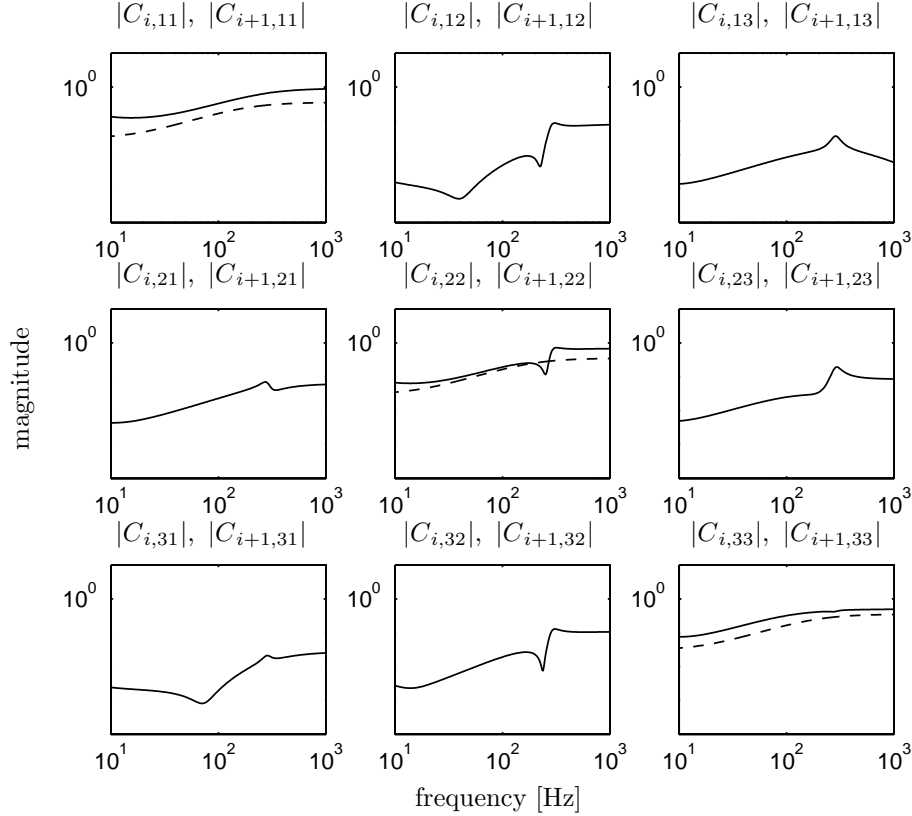
Fig. 8.23: Magnitude Bode plot of newly designed controller C_{i+1} (—) compared
with initial controller C_i(- -).

closed-loop performance. After reduction, the designed controller C_{i+1} is a stable
controller, having a McMillan degree of 11. An amplitude Bode plot of the newly
designed controller C_{i+1} can be found in Figure 8.23.

Compared to the (initial) controller C_i it can be seen that C_{i+1} is a multivari-
able controller. Furthermore, the controller C_{i+1} has additional dynamics to account
for the modelled (uncertain) mechanical resonance modes of the plant P_o. Control-
ling these mechanical resonance modes in the wafer stage will yield better control
performance and enhanced robustness. This will be clarified in the next section.

8.5.2 Evaluation of performance

Before implementing the controller C_{i+1} in a feedback connection with the plant P_o,
the performance and stability of the feedback connection $\mathcal{T}(P_o, C_{i+1})$ needs to be
verified *a priori*. The *a priori* performance evaluation can be carried out with the

estimated set of models $\mathcal{P}_i$, that represent the limited knowledge of the plant P_o.

Similarly to the *a posteriori* performance evaluation, this is done by evaluating

$$\|J(P, C_{i+1})\|_\infty = \|U_2 T(P, C_{i+1}) U_1\|_\infty$$

for all models $P \in \mathcal{P}_i$. In this case, stability robustness for the controller C_{i+1} also has to be checked. Both stability and performance robustness of C_{i+1} can be evaluated with the structured singular value $\mu\{M\}$, see also Corollary 5.3-6 and Corollary 5.3-9.
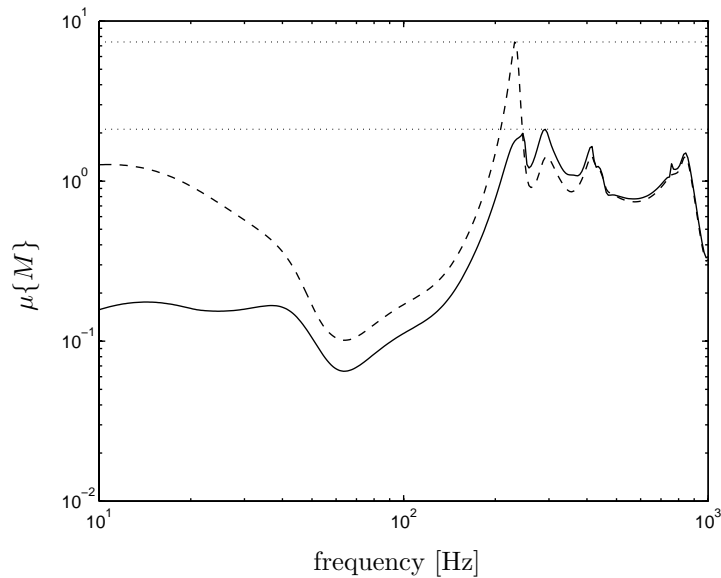
Fig. 8.24: Computation of structured singular value $\mu\{M(e^{i\omega})\}$ for the controller C_i (- -) and the controller C_{i+1} (—) applied to the set of models $\mathcal{P}_i$.

To compare the performance (robustness) of the newly designed controller C_{i+1}, in Figure 8.24 the structured singular value of $\mu\{M(e^{i\omega})\}$ has been plotted point wise over the frequency domain range between 10 and 1000Hz. For both the controller C_i and C_{i+1} the resulting structured singular value $\mu\{M(e^{i\omega})\}$ has been plotted and it can be seen that the controller C_{i+1} has improved the performance robustness and satisfies

$$\|U_2 T(P, C_{i+1}) U_1\|_\infty \leq \gamma_{i+1} < \gamma_i$$

for all models $P \in \mathcal{P}_i$.

For presentation purposes, the weighting functions are scaled to normalize the uncertainty. It can be seen from Figure 8.24 that performance robustness cannot be guaranteed for the controller C_{i+1}, as $\gamma_{i+1} \not< 1$. However, stability robustness is guaranteed and C_{i+1} has a guaranteed improved performance compared to C_i.

Implementation of C_{i+1} on the plant P_o also proves the improved performance of the controller C_{i+1}.

Application of the step reference signals depicted in Figure 8.9 in an x-direction yields a (wafer) step in x-direction. Using the newly designed feedback controller C_{i+1} and the specified reference signals r_1 and r_2, a servo error pattern can be measured. With the controller C_{i+1} implemented on the plant P_o, the servo error $u_{c,\mathsf{x}}$ depicted in Figure 8.25 is obtained.
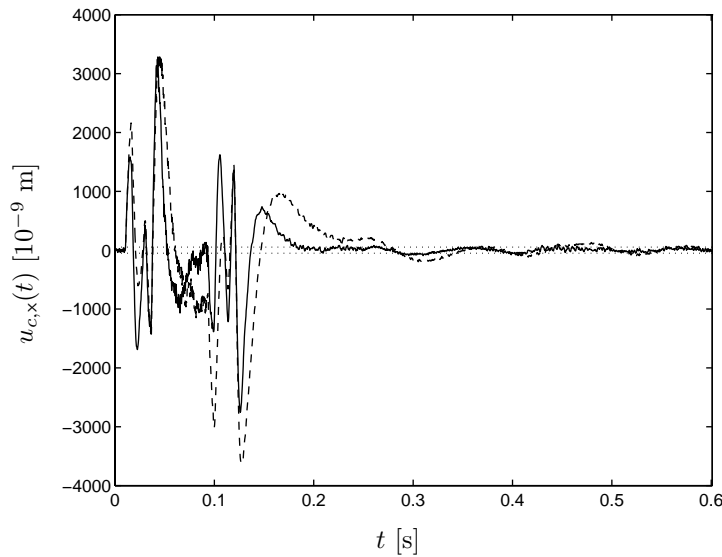
Fig. 8.25: Servo error u_c response to a step in x-direction using controller C_{i+1} (—) and using controller C_i (- -) and required accuracy interval ($\cdots$).

Although there is a slight improvement in the performance of the feedback system due to reduction and increased speed of the servo error, one can aim for additional performance improvement. The newly designed controller C_{i+1} is now implemented on the plant P_o and provides new closed-loop data to model the dynamics of the plant P_o.

8.6 Multiple Iterations of Identification and Control

8.6.1 Iterative approach

As indicated before in Section 7.3, once a newly designed controller C_{i+1} has been designed and implemented, the identification and control design procedure discussed in Section 8.4 and Section 8.5 can be reinvoked. Basically, the same steps of nomi-

nal model identification, model uncertainty estimation and robust control design are performed, whereas the subscript i in C_i and $\mathcal{P}_i$ is updated via $i = i + 1$.

Applying the designed controller to the plant and reinvoking the estimation of a set of models $\mathcal{P}_i$ and redesigning the controller C_i, sets up an iterative approach of model set estimation and (robust) control design. The iterative approach is used to improve the modelling and control of the unknown plant P_o by progressively lowering the (robust) performance criterion.

The new feedback connection $\mathcal{T}(P_o, C_{i+1})$ provides new data for system identification purposes. The information of the new controller C_{i+1} is re-used, as C_{i+1} now becomes the new initial controller C_i used during identification. The information of the (new) initial controller C_i is used to gather data from the closed-loop system and to structure the set of models $\mathcal{P}_i$ to be identified.

The auxiliary model P_x used to construct the filter F in (4.24) can also be updated. The filter F is used to access a rcf $(N_{o,F}, D_{o,F})$ of the plant P_o and can be updated by the nominal model $\hat{P}_i$ that has been estimated.

The method of constructing of a new set of models $\mathcal{P}_i$ remains unaltered. First a new nominal factorization $(\hat{N}_i, \hat{D}_i)$ has to be estimated. Subsequently, the model uncertainty estimation must be used to characterize an upper bound for the model uncertainty associated to the newly identified nominal factorization $(\hat{N}_i, \hat{D}_i)$. Both complete again a newly identified set of models $\mathcal{P}_i$.

Once a set of models $\mathcal{P}_i$ is available, the *a posteriori* performance evaluation discussed in Section 8.5.1 can be performed. Followed by a robust control design, the performance of the feedback connection $\mathcal{T}(P_o, C_i)$ with $i = i + 1$ may be further improved in case the condition

$$\|U_2 T(P, C_i) U_1\|_\infty \leq \gamma_{i+1} < \gamma_i$$

can be met for all models $P \in \mathcal{P}_i$. Once this condition cannot be met, either due to the unavoidable size of the model uncertainty or the inherent performance limitations of the plant P_o, the iteration can be stopped. The results on such an iterative approach to improve the performance of the wafer stage progressively, have been summarized in the following section.

8.6.2   Further improving performance

The procedure of identifying a nominal model, estimating the model uncertainty bounding, constructing a set of models and the design of a robust controller is the same in every step of an iterative approach. To avoid a needless repetition of these steps, in this section only the (robust) performance results are mentioned.

To indicate the progressive performance improvement during the subsequent steps within the iterative procedure, the (robust) performance evaluation for 4 consecutive

steps of subsequent model set estimation and controller design have been summarized
in Figure 8.26.
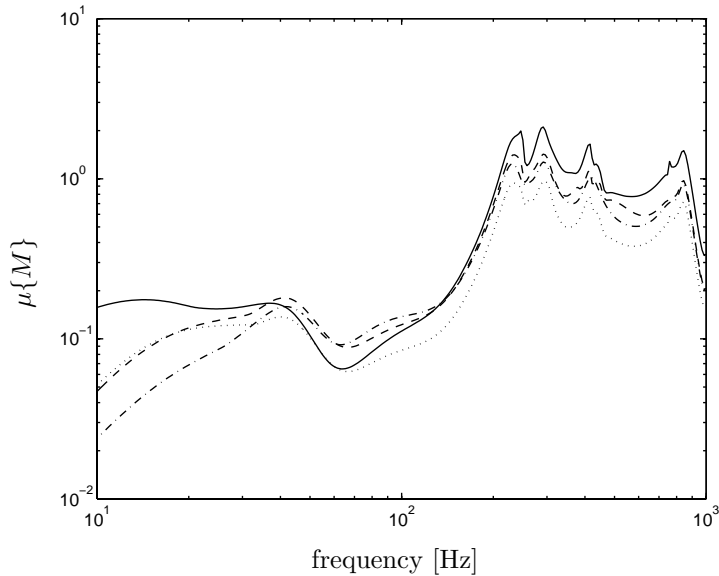
Fig. 8.26: Computation of structured singular value $\mu\{M(e^{i\omega})\}$ for 4 consecutive
controllers C_1 (—), C_2 (- -), C_3 (-·), C_4 (···) found by subsequent model
set $\mathcal{P}_i$ estimation and controller C_{i+1} design.

It can be observed from Figure 8.26 that the (robust) performance has been im-
proved progressively. The controllers C_{i+1} were designed on different sets of models
$\mathcal{P}_i$. The sets of models $\mathcal{P}_i$ are estimated on the basis of closed-loop data coming from
the feedback connection of the plant P_o and the (previous) controller C_i implemented
on the plant P_o.

For reasons of completeness, the amplitude Bode plot of the four consecutive
controllers has been depicted in Figure 8.27. From this figure it can be observed
that the high frequent dynamics and the dynamics of the off-diagonal terms of the
multivariable controller is changing significantly. This is due to the fact that during
the subsequent identification and control design, renewed knowledge of the plant P_o is
obtained. On this renewed knowledge, especially coming from the more high frequent
dynamics and the off-diagonal elements, the model-based controlelr is built and is
able to improve the robust performance.

Finally it can be mentioned here that the fourth controller C_4 being designed (and
reduced) has a McMillan degree of 19 and satisfies performance robustness condition
as $\mu\{M\} < 1$. As the performance robustness could be guaranteed, the iterative
approach of model set estimation and (robust) control design is stopped since desired
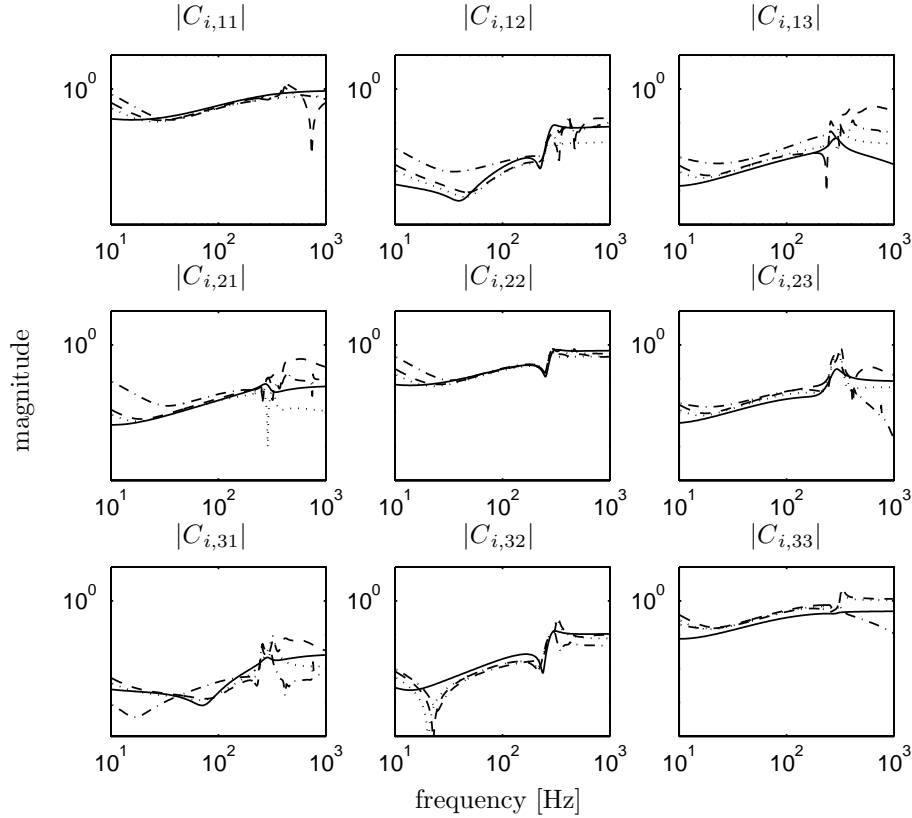
Fig. 8.27: Magnitude Bode plot of 4 consecutive controllers C_1 (—), C_2 (- -), C_3 (-·),
C_4 (···) found by an iterative approach of mode set estimation and robust
control design.

performance is satisfactory.

The satisfactory performance of the feedback connection of the wafer stage and
designed feedback controller is confirmed by applying step reference signals. To illus-
trate the performance improvement, the step reference signals depicted in Figure 8.9
are applied in x-direction to perform a (wafer) step in x-direction. Similar step ref-
erence signals can also be given in y-direction, but for illustrative purposes only the
x-direction will be displayed here. With the designed feedback controller C_4 imple-
mented on the plant, the servo error $u_{c,x}$ depicted in Figure 8.28 is obtained.

It can be seen that the servo error reaches the required accuracy interval in less
than 0.15s, which is a significant improvement compared to the servo error generated
by the initial controller C_i that consists of 3 parallel PID controllers. The first part of
the servo error remains nearly unchanged as the size and shape of this signals highly
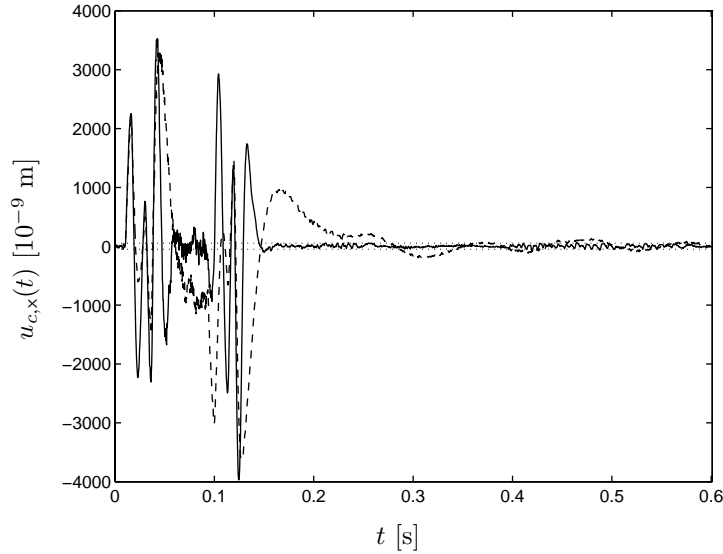
Fig. 8.28: Servo error u_c response to a step in x-direction using controller C_4 (—)
and using (initial) controller C_i (- -) and required accuracy interval ($\cdots$).

depends on the feedforward signal instead of the feedback signal (de Roover 1997).

8.7   Discussion

This chapter has presented the results of the model-based procedure given in
Procedure 2.5-4 applied to the mechanical servo mechanism present in a wafer stage.
The application has showed how a set of models $\mathcal{P}$ can be estimated on the ba-
sis of closed-loop data and used for robust control design. Furthermore, the ap-
plication showed how to improve the (robust) performance of the feedback control
system by systematically monitoring the performance robustness, as indicated in
Procedure 2.5-4.

The wafer stage application included the estimation of multivariable models on the
basis of closed-loop data. The problem of estimating models on the basis of closed-
loop data is addressed in this thesis by the algebraic framework of fractional model
representations that provides a unified approach to handle the identification of stable
and unstable systems on the basis of closed loop data.

A set of models is estimated on the basis of closed-loop data by formulating a
structure for the set of models that is based on a dual-Youla based perturbation of
the coprime factors of the nominal model. This structure enables an uncertainty
bounding identification routine to operate directly on (filtered) closed-loop data to

estimate a bound for the model uncertainty to complete the set of models.

The subsequent identification of a set of models $\mathcal{P}$ on the basis of closed loop data, robust controller design and controller implementation while monitoring the performance robustness has delivered valuable models of the wafer stage (de Roover 1997) and a satisfactory robust controller. The feedback controller produces a fast and accurate servo mechanism and is able to improve the performance of a wafer stepper by increasing the number of processed wafers per hour (throughput) and the number of acceptable IC's per wafer (yield).

Acknowledgements

Part V

Epilogue

9

Conclusions and Recommendations

9.1 Contributions of this Thesis

In this thesis it has been advocated that the design of an enhanced performing and robust controller for an unknown plant can be done on the basis of system identification techniques followed by a subsequent robust control design. Although a model will always be a simplified representation of the unknown plant, robust control can account for the presence of a modelling error being made. As such, the system identification procedure developed in this thesis addresses the problem of approximate identification and the effects of modelling errors in developing models for control applications.

The identification procedure in this thesis focuses on the estimation of both an (approximate) nominal model and an upper bound on the modelling error. In this way, a set of models is constructed on which an enhanced performing and robust controller is designed. The need for (high) performance and the conflicting nature between performance and robustness has motivated the estimation and construction of a set of models that it is suitable for robust control design. Making the set of models suitable for control design, is done in this thesis by keeping track of the intended control application of the set of models. This is done by estimating and constructing a set of models in which the feedback (controller) plays a leading role.

As such, in this thesis the following research areas have been merged to come up with a model based approach that is able to monitor and improve the feedback controlled performance of an unknown plant. thesis include the following topics

- Control relevant identification of nominal models

- Estimation of model uncertainty

- Robust control design

The emphasize of this thesis lies on the field of system identification, for which new

results on closed-loop identification and model uncertainty set estimation for robust control design are presented.

Within the abovementioned areas existing results on model uncertainty estimation and robust controller design techniques have been used. However, to improve the suitability of a model uncertainty set obtained by system identification techniques, the issue of model complexity of both the nominal model and the model uncertainty bound is addressed in this thesis. New results are presented on the estimation of reduced order complexity models to model a possibly unstable plant. Additionally, a structure of the model uncertainty set is proposed that is based on coprime factor perturbations. This structure is shown to be particularly useful for identification purposes and is applicable to data obtained under closed-loop or controlled conditions.

Along the lines of a motivation to merge the research areas, in this thesis the identification and construction of a set of models is done entirely in a closed-loop manner. A set of models is proposed that is based on the estimation of a low complexity nominal model obtained via closed-loop experiments. Furthermore, the modelling errors are described via perturbations on the nominal model in a closed-loop setting and are also estimated on the basis of closed-loop experiments.

Both the identification and the robust controller design have been integrated in this thesis in a scientific, iterative procedure that enables a subsequently improvement of the performance of the feedback controlled plant. The closed-loop approach and the use of corresponding closed-loop experiments has led to the development of so-called (multivariable) closed-loop identification techniques in this thesis. The closed-loop identification techniques are able to deal with data coming from closed-loop experiments. Furthermore, the techniques are mainly based on the framework of stable factorizations to enable a unified approach for the identification of stable and unstable plants. The effectiveness of the procedure has been illustrated on the identification and control of a multivariable servo mechanism present in a wafer stepper.

9.2   Main Conclusions

Designing a controller that is able to improve the performance of a controlled system by means of system identification and robust control design begins by recognizing the interaction and the link between modelling and control. Characterizing a model on the basis of open-loop considerations can be misleading in case the model needs to be used in a closed-loop setting.

To characterize the quality of the model, the errors being made during modelling need to be quantified. Furthermore, robust control can account for the presence of modelling errors being made to design a robust performing controller. Taking into account the link between modelling and control, not only the (nominal) modelling but also the characterization of the modelling errors must be subjected to a description based on closed-loop considerations.

To treat the issue of approximate modelling for control, a set of models built up from a nominal model and an upper bound on the model error can be used. To take into account the link between modelling and control, both the estimation of a nominal model and the estimation of the model error should take into account the intended control application. For that purpose, approximate models should be identified using a closed-loop relevant identification criterium. Additionally, the model error should be described in a closed-loop setting to study the effect of model perturbations in a feedback controlled conditions. Both requirements unleash the need to perform the modelling on the basis of closed-loop experiments.

A unified treatment of stable and unstable systems operating under feedback can be obtained by the framework of stable factorizations. In this framework, the dynamical behaviour of the possibly unstable system is thought to be split up in a map of two stable factors. Next to the unified treatment of stable and unstable systems, the algebraic framework used in this thesis provides the following benefits.

- The algebraic framework of (stable) fractional representations opens the possibility of an open-loop equivalent identification of the stable factors. In this way, the identification of a possible unstable system on the basis of closed-loop data can be handled relatively easily.

- The algebraic framework allows possible model errors to be described in a closed-loop setting, using a perturbation in a dual-Youla parametrization. In this way, a set of models is obtained that is able to study the effect of model perturbations under feedback controlled conditions. Specifically, the set of models describes all models that are stabilized by a given feedback controller. This is done by considering a nominal stable factorization perturbed by an unknown, but bounded stable operator.

Taking into account the intended control application unleashes the need for an iterative scheme, as the control application (the feedback controller) is not know yet. Starting with an initial feedback controller, an iterative scheme of modelling and redesigned feedback controllers will provide information on the dynamics of the plant relevant for feedback control.

Information on the required complexity of the nominal model, shape of allowable model uncertainty and attainable robust performance is gathered while performing such an iteration. Including the characterization of modelling errors enables one to avoid performance deterioration during the iterative scheme. Avoiding performance deterioration during the iterative scheme also opens the possibility to formulate (in)validation criteria to accept or refuse models and/or controllers found during the iterative scheme.

9.3   Retrospect and Recommendations

The interaction between modelling and control has been recognized before, see e.g. Åström and Wittenmark (1971) or Skelton (1989). The problem of designing (robust) controllers on the basis of models that exhibit model errors has been ongoing field of research (Zhou *et al.* 1996). Modelling by system identification and the use of data obtained under closed-loop conditions to address the problem the interaction between modelling and control has also been studied before, see e.g. Gevers and Ljung (1986) or Bitmead *et al.* (1990*a*). The benefits of using the algebraic framework of fractional representations has been motivated in Hansen (1989) or Schrama (1992*b*) both for dealing with closed-loop data and to address the problem of closed-loop relevant identification.

In this perspective, this thesis continues and combines the work that has been done in these areas. As a result, a reasonably complicated iterative procedure has been developed in which the performance of the feedback controlled plant is monitored and improved by subsequent system identification and robust control design. This procedure was presented in Procedure 2.5-4 and depicted schematically in Figure 2.7.

For many applications, the proposed procedure may be too complicated to perform and simplifications may be necessary. After all, the only objective is retune the feedback controller to improve the performance of the feedback controlled plant. As such, the following remarks can be given.

- The proposed procedure delivers more than necessary. Especially, the identification provides information on the required complexity of the nominal model, shape of allowable model uncertainty and attainable robust performance. One may question the benefit of having such a detailed model. Sacrificing the need to guarantee to attain robust performance may lead to a more straightforward method of controller tuning, see e.g. Hjalmarsson *et al.* (1994).

- In many situation, the plant is known to be stable. An identification based on stable factorizations may be an exaggerated approach to handle the modelling of a stable plant. In that case, more simplified versions to deal with closed-loop data, such as presented in (Van den Hof and Schrama 1993), may be advisable. Development and analysis of special and user friendly identification algorithms, that deal with specific closed-loop situations, is advisable (Van den Hof *et al.* 1997).

- The need for an iterative procedure for joint identification and control (Schrama 1992*b*) is also advocated in this thesis. Clearly, the iterative procedure is introduced to replace the far more complicated problem of a joint optimization of modelling and control. Possibilities to perform such a joint optimization still need to be investigated.

- The need for finding low complexity models may be questioned in the near future. With increasing computing power or the possibility to compute low complexity robust controller on the basis of complex models, the need for deliberate undermodelling or low complexity models may become superfluous. In this perspective, research in the area of fixed order controller design is valuable. Still, the need for low complexity models and/or deliberate undermodelling remains valuable to avoid the estimation, simulation and computation of unnecessarily and highly complicated models.

The above mentioned items also indicate possibilities for new research in the field of closed-loop relevant system identification and robust control design. Despite possible requirements on simplification, the procedure is generally applicable to many feedback controlled systems that may require an improved tuning of the feedback connection. The identification requires and can deal with data obtained under closed-loop conditions, while both the identification of stable and unstable systems can be handled.

Part VI

Addenda

A

Matrix & State Space Formulae

A.1 Matrix Inversion Lemma

Let A be a sqare matrix that is partioned as follows

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} and A_{22} are again square matrices. In case A_{11} is non-singular, the matrix A admits the following decomposition

$$A = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}$$

where $S := A_{22} - A_{21}A_{11}^{-1}A_{12}$ is the Schur complement of A_{11} with respect to A. In case A_{22} is non-singular, the matrix A admits the following decomposition

$$A = \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix}$$

where $S := A_{11} - A_{12}A_{22}^{-1}A_{21}$ is the Schur complement of A_{22} with respect to A. It can be seen that A is non-singular if and only if the Schur complement is non-singular.

A.2 State Space Realization of $T(P, C)$

Let the state space realization of the system P be given by

$$P \doteq \left[\begin{array}{c|c} A_p & B_p \\ \hline C_p & D_p \end{array} \right]$$

and let the state space realization of the controller C be given by

$$C \doteq \left[\begin{array}{c|c} \mathsf{A}_c & \mathsf{B}_c \\ \hline \mathsf{C}_c & \mathsf{D}_c \end{array} \right]$$

then the transfer function

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

has a state space realization

$$T(P,C) \doteq \left[\begin{array}{c|c} \mathsf{A} & \mathsf{B} \\ \hline \mathsf{C} & \mathsf{D} \end{array} \right]$$

where the state space matrices are given by

$$\mathsf{A} = \begin{bmatrix} \mathsf{A}_p - \mathsf{B}_p M \mathsf{D}_c \mathsf{C}_p & \mathsf{B} M \mathsf{C}_c \\ -\mathsf{B}_c \mathsf{C} + \mathsf{B}_c \mathsf{D}_p M \mathsf{D}_c \mathsf{C}_p & \mathsf{A}_c - \mathsf{B}_c \mathsf{D}_p M \mathsf{C}_c \end{bmatrix}$$

$$\mathsf{B} = \begin{bmatrix} \mathsf{B}_p M \mathsf{D}_c & \mathsf{B}_p M \\ \mathsf{B}_c - \mathsf{B}_c \mathsf{D}_p M \mathsf{D}_c & -\mathsf{B}_c \mathsf{D}_p M \end{bmatrix}$$

$$\mathsf{C} = \begin{bmatrix} \mathsf{C}_p - \mathsf{D}_p M \mathsf{D}_c \mathsf{C}_p & \mathsf{D}_p M \mathsf{C}_c \\ -M \mathsf{D}_c \mathsf{C}_p & M \mathsf{C}_c \end{bmatrix}$$

$$\mathsf{D} = \begin{bmatrix} \mathsf{D}_p M \mathsf{D}_c & \mathsf{D}_p M \\ M \mathsf{D}_c & M \end{bmatrix}$$

and $M := (I + \mathsf{D}_c \mathsf{D}_p)^{-1}$.

A.3   State Space Realization of ND^{-1}

Given a *rcf* (N, D), this section described the state space realization of P computed via $P = ND^{-1}$. Dual state space formulae exist for a *lcf* $(\tilde{N}, \tilde{D})$ with $P = \tilde{D}^{-1}\tilde{N}$.

Let (N, D) be a *rcf* where N and D have a state space realization with a common state matrix A_{ND} and input matrix B_{ND}. Then, the state space realization of the *rcf* (N, D) can be written as

$$\begin{bmatrix} N \\ D \end{bmatrix} \doteq \left[\begin{array}{c|c} \mathsf{A}_{ND} & \mathsf{B}_{ND} \\ \hline \mathsf{C}_N & \mathsf{D}_N \\ \mathsf{C}_D & \mathsf{D}_D \end{array} \right] \tag{A.1}$$

and the state space realization of $P := ND^{-1}$ is given by

$$P \doteq \left[\begin{array}{c|c} \mathsf{A}_p & \mathsf{B}_p \\ \hline \mathsf{C}_p & \mathsf{D}_p \end{array} \right]$$

where the state space matrices are given by

$$
\begin{array}{rcl}
\mathsf{A}_p & = & \mathsf{A}_{ND} - \mathsf{B}_{ND}\mathsf{D}_D^{-1}\mathsf{C}_D \\
\mathsf{B}_p & = & \mathsf{B}_{ND}\mathsf{D}_D^{-1} \\
\mathsf{C}_p & = & \mathsf{C}_N - \mathsf{D}_N\mathsf{D}_D^{-1}\mathsf{C}_D \\
\mathsf{D}_p & = & \mathsf{D}_N\mathsf{D}_D^{-1}
\end{array}
\tag{A.2}
$$

provided that D_D is non-singular. It can be observed that in this case P in general will have a McMillan degree less than or equal to the McMillan degree of (N, D).

A.4 State Space Realization of $\bar{Q}$

This section describes the state space realization of the unweighted coefficient matrix $\bar{Q}$ as used int he LFT respresentation of the set of model $\mathcal{P}$. To simplify notations, the state space realization of $\bar{Q}$ is presented as a series connection of two state space realizations.

Let (N, D) be a *rcf* of the system P and let (N_c, D_c) be a *rcf* of the controller. Then the unweighted coefficient matrix $\bar{Q}$ is given by

$$\bar{Q} = \left[\begin{array}{cc} D^{-1}N_c & D^{-1} \\ (D_c + PN_c) & P \end{array} \right] = \underbrace{\left[\begin{array}{cc} D^{-1} & 0 \\ P & I \end{array} \right]}_{\bar{Q}_1} \underbrace{\left[\begin{array}{cc} N_c & I \\ D_c & 0 \end{array} \right]}_{\bar{Q}_2} \tag{A.3}$$

and can be viewed as a series connection of the two transfer function matrices $\bar{Q}_1$ and $\bar{Q}_2$ on the right hand side of (A.3). Similar to the state space realization of the *rcf* (N, D) in (A.1), the state space realization of the *rcf* (N_c, D_c) can be written as

$$\left[\begin{array}{c} N_c \\ D_c \end{array} \right] \doteq \left[\begin{array}{c|c} \mathsf{A}_{N_cD_c} & \mathsf{B}_{N_cD_c} \\ \hline \mathsf{C}_{N_c} & \mathsf{D}_{N_c} \\ \mathsf{C}_{D_c} & \mathsf{D}_{D_c} \end{array} \right] \tag{A.4}$$

yielding

$$\left[\begin{array}{cc} N_c & I \\ D_c & 0 \end{array} \right] \doteq \left[\begin{array}{c|cc} \mathsf{A}_{N_cD_c} & \mathsf{B}_{N_cD_c} & 0 \\ \hline \mathsf{C}_{N_c} & \mathsf{D}_{N_c} & I \\ \mathsf{C}_{D_c} & \mathsf{D}_{D_c} & 0 \end{array} \right] \tag{A.5}$$

as a state space realization for $\bar{Q}_2$ in (A.3). The state space realization of the *rcf* (N, D) is given in (A.1) and yields

$$D^{-1} \doteq \left[\begin{array}{c|c} \mathsf{A}_{ND} - \mathsf{B}_{ND}\mathsf{D}_D^{-1}\mathsf{C}_D & \mathsf{B}_{ND}\mathsf{D}_D^{-1} \\ \hline \mathsf{D}_D^{-1}\mathsf{C}_D & \mathsf{D}_D^{-1} \end{array} \right] \tag{A.6}$$

as state space realization for D^{-1}. Combining (A.2) and (A.6) yields

$$\left[\begin{array}{cc} D^{-1} & 0 \\ P & I \end{array} \right] \doteq \left[\begin{array}{c|cc} \mathsf{A}_{ND} - \mathsf{B}_{ND}\mathsf{D}_D^{-1}\mathsf{C}_D & \mathsf{B}_{ND}\mathsf{D}_D^{-1} & 0 \\ \hline \mathsf{D}_D^{-1}\mathsf{C}_D & \mathsf{D}_D^{-1} & 0 \\ \mathsf{C}_N - \mathsf{D}_N\mathsf{D}_D^{-1}\mathsf{C}_D & \mathsf{D}_N\mathsf{D}_D^{-1} & I \end{array} \right] \tag{A.7}$$

as a state space realization for $\bar{Q}_1$ in (A.3). Henceforth, the state space realization of $\bar{Q}$ is found by a series connection of (A.7) and (A.5). This leads to a state space realization where the McMillan degree is the sum of the McMillan degree of the system P and the controller C.

B

Least-Squares Frequency Response Curve Fitting

B.1 Problem formulation

To formulate the multivariable frequency domain identification problem, consider the following set $\mathcal{G}$ of noisy complex frequency response data observations $G(\omega_j)$, evaluated at N frequency points ω_j.

$$\mathcal{G} := \{G(\omega_j) \mid G(\omega_j) \in \mathbb{C}^{p \times m}, \text{ for } j \in 1, \dots, N\} \tag{B.1}$$

The aim of the identification problem discussed in this appendix[1] is to find a linear time invariant multivariable model P of *limited complexity*, having m inputs and p outputs, that *approximates* the data $\mathcal{G}$ in (B.1).

To address the limited complexity, the model $P(\theta)$ is parametrized by a either a left or right polynomial MFD that depends on a real valued parameter θ of limited dimension. The specific parametrization of the polynomial MFD of $P(\theta)$ is discussed in the next section. The approximation of the data $\mathcal{G}$ by the model $P(\theta)$ is addressed by considering the following additive error.

$$E_a(\omega_j, \theta) := [G(\omega_j) - P(\xi(\omega_j), \theta)] \text{ for } j \in 1, \dots, N \tag{B.2}$$

The complex variable $\xi(\cdot)$ in (B.2) is used to denote the frequency dependency of the model $P(\theta)$. In this way, $\xi(\omega_j) = i\omega_j$ to represent a continuous time model, whereas $\xi(\omega_j) = e^{i\omega_j T}$ (shift operator) or $\xi(\omega_j) = (e^{i\omega_j} - 1)/T$ (δ operator) to represent a discrete time model with sampling time T.

To tune the additive error E_a in (B.2), both an input-output frequency weighted curve fit error E_w with

$$E_w(\omega_j, \theta) := W_{out}(\omega_j) E_a(\omega_j, \theta) W_{in}(\omega_j) \tag{B.3}$$

---

[1]Parts of this appendix have also been published in de Callafon *et al.* (1996*a*)

and an element-wise frequency weighted curve fit error E_s with

$$E_s(\omega_j, \theta) := S(\omega_j). * E_a(\omega_j, \theta) \tag{B.4}$$

will be considered in this appendix. In (B.4) .∗ is used to denote the Schur product; an element-by-element multiplication.

Using the notation E to denote the frequency weighted curve fit error E_w in (B.3) and E_s in (B.4), the deviation of the data $\mathcal{G}$ is characterized by following the norm function $J(\theta)$.

$$J(\theta) := \sum_{i=1}^{N} \text{tr}\{E(\omega_j, \theta) E^*(\omega_j, \theta)\} = \|E(\theta)\|_F^2 \tag{B.5}$$

In (B.5) * is used to denote the complex conjugate transpose, $\text{tr}\{\cdot\}$ is the trace operator and $\|E(\theta)\|_F$ denotes the Frobenius norm operating on the matrix $E(\theta) = [E(\omega_1, \theta) \ \cdots \ E(\omega_N, \theta)]$. Consequently, the goal of the procedure described in this appendix is to find a real valued parameter $\hat{\theta}$ of limited complexity that can be formulated by the following minimization.

$$\hat{\theta} := \arg \min_{\theta \in \mathbb{R}} J(\theta) \tag{B.6}$$

B.2 Parametrization

The multivariable model is represented by either a left or right polynomial MFD, respectively given by

$$P(\xi, \theta) = A(\xi^{-1}, \theta)^{-1} B(\xi^{-1}, \theta) \tag{B.7}$$
$$P(\xi, \theta) = B(\xi^{-1}, \theta) A(\xi^{-1}, \theta)^{-1} \tag{B.8}$$

where A and B denote parametrized polynomial matrices in the indeterminate ξ^{-1}.

For a model having m inputs and p outputs, the the polynomial matrix $B(\xi^{-1}, \theta)$ is parametrized by

$$B(\xi^{-1}, \theta) = \sum_{k=d}^{d+b-1} B_k \xi^{-k} \tag{B.9}$$

where $B_k \in \mathbb{R}^{p \times m}$, d denotes the number of leading zero matrix coefficients and b the number of non-zero matrix coefficients in $B(\xi^{-1}, \theta)$. For the left MFD in (B.7), $A(\xi^{-1}, \theta)$ is parametrized by

$$A(\xi^{-1}, \theta) = I_{p \times p} + \xi^{-1} \sum_{k=1}^{a} A_k \xi^{-k+1} \tag{B.10}$$

where $A_k \in \mathbb{R}^{p \times p}$ and a denotes the number of non-zero matrix coefficients in the monic polynomial $A(\xi^{-1}, \theta)$. The parameter θ is determined by the corresponding

unknown matrix coefficients in the polynomials. Hence,

$$\theta = \left[\begin{array}{cccccc} B_d & \cdots & B_{d+b-1} & A_1 & \cdots & A_a \end{array} \right] \tag{B.11}$$

and $\theta \in \mathbb{R}^{p \times (mb+pa)}$ for the left MFD in (B.7). Dual results can be formulated for the right MFD in (B.8).

Additionally to the full polynomial parametrization presented here, so-called structural parameters d_{ij}, b_{ij} and a_{ij} with $d := \min\{d_{ij}\}$, $b := \max\{b_{ij}\}$, and $a := \max\{a_{ij}\}$ can be used to specify a non-full polynomial parametrization. In this way, the parameter θ as given in (B.11) may contain prespecified zero entries at specific locations. This may occur in a discrete time model with $\xi^{-1} = z^{-1}$ where the value of d_{ij} has a direct connection with the number of time delays from the jth input to the ith output.

B.3   Computational Procedure

B.3.1   Iterative minimization

In this section, the minimization of (B.6) will be discussed by means of an iterative procedure of convex optimization steps similar to the SK-iteration of Sanathanan and Koerner (1963). The attention will be restricted to a parametrization of $P(\xi, \theta)$ based on the left MFD (B.7) as dual results can be obtained for a right MFD. To extend the SK-iteration to the multivariable case, first consider the (unweighted) additive curve fit error of (B.2).

For a model $P(\xi, \theta)$ parametrized by left MFD, (B.2) can be written as

$$E_a(\omega_j, \theta) = A(\xi(\omega_j)^{-1}, \theta)^{-1} \tilde{E}(\omega_j, \theta) \tag{B.12}$$

where $\tilde{E}(\omega_j, \theta)$ is the equation error defined by

$$\tilde{E}(\omega_j, \theta) := A(\xi(\omega_j)^{-1}, \theta) G(\omega_j) - B(\xi(\omega_j)^{-1}, \theta). \tag{B.13}$$

Substituting the parametrization (B.7) for the polynomials A, B, the equation error in (B.13) can be represented by

$$\tilde{E}(\omega_j, \theta) = G(\omega_j) - \theta \Phi(\omega_j) \tag{B.14}$$

where θ is given in (B.11) and

$$\Phi(\omega_j) = \left[\begin{array}{c} I_{m \times m} \xi(\omega_j)^{-d} \\ \vdots \\ I_{m \times m} \xi(\omega_j)^{-(d+b-1)} \\ G(\omega_j) \xi(\omega_j)^{-1} \\ \vdots \\ G(\omega_j) \xi(\omega_j)^{-a} \end{array} \right] \tag{B.15}$$

with $\Phi(\omega_j) \in \mathbb{C}^{(mb+pa) \times m}$.

A matrix $\tilde{E}(\theta)$ can be formed by stacking $\tilde{E}(\omega_j, \theta)$ column-wise for $j \in 1, \ldots, N$ and this yields

$$\arg \min_{\theta \in \mathbb{R}} \|\tilde{E}(\theta)\|_F^2 = \arg \min_{\theta \in \mathbb{R}} \|\mathsf{G} - \theta\mathsf{P}\|_F^2 \tag{B.16}$$

where G and P are found by stacking the real and imaginary part of respectively $G(\omega_j)$ and $\Phi(\omega_j)$ for $j \in 1, \ldots, N$. Due to the linear appearance of the parameter θ, (B.16) corresponds a standard least squares problem that can be solved by numerical reliable tools as e.g a QR-factorization with (partial) pivoting (Golub and Van Loan, 1989).

Due to the fact that $A(\xi^{-1}, \theta)$ in (B.12) also depends on the parameter θ, the linear appearance of the parameter θ in (B.12) is violated. In order to facilitate the convexity in minimizing the two-norm on the equation error in (B.16), an iterative procedure similar as in Sanathanan and Koerner (1963) can be used. An estimate $\hat{\theta}_t$ in step t is computed by replacing $A(\xi(\omega_j)^{-1}, \theta)$ in (B.12) by a *fixed* $A(\xi(\omega_j)^{-1}, \hat{\theta}_{t-1})$ based on an estimate $\hat{\theta}_{t-1}$ obtained from the previous step $t - 1$. In this way the Frobenius norm of an output weighted equation error $\tilde{E}_w(\omega_j, \hat{\theta}_{t-1}, \theta) = A(\xi(\omega_j)^{-1}, \hat{\theta}_{t-1})^{-1}\tilde{E}(\omega_j, \theta)$ needs to be minimized repeatedly according to

$$\hat{\theta}_t = \arg \min_{\theta \in \mathbb{R}} \|\tilde{E}_w(\hat{\theta}_{t-1}, \theta)\|_F^2.$$

This generalizes the SK-iteration to multivariable models parametrized by a left polynomial MFD. A dual approach can be formulated for a right polynomial MFD.

The estimate obtained from the SK-iteration is not optimal in the sense of (B.6) in presence of noise and/or incorrect model order, but it does provide a tool to find an initial estimate for a GN-optimization (Whitfield, 1987). Furthermore, the convex optimization to be solved in each step of the multivariable SK-iteration supports the estimation of models with many parameters. The computational procedure to obtain the parameter $\hat{\theta}$ in case of the (weighted) curve fit errors of (B.3) and (B.4) is presented in the subsequent sections.

B.3.2   Input-output weighting

The input-output weighted curve fit error of (B.3) can be rewritten into

$$E_w(\omega_j, \theta) = \tilde{W}_{out}(\omega_j, \theta)\tilde{E}(\omega_j, \theta)W_{in}(\omega_j) \tag{B.17}$$

where $\tilde{W}_{out}(\omega_j, \theta) := W_{out}(\omega_j)A(\xi(\omega_j)^{-1}, \theta)^{-1}$ and $\tilde{E}(\omega_j, \theta)$ is given in (B.13).

Using a similar approach of iterative minimization steps as used in Section B.3.1, the parameter θ in $\tilde{W}_{out}(\omega_j, \theta)$ in (B.17) is fixed to an estimate $\hat{\theta}_{t-1}$ obtained from the previous step $t - 1$. Consequently, the weighted equation error $\tilde{E}_w$ defined by

$$\tilde{E}_w(\omega_j, \hat{\theta}_{t-1}, \theta) := \tilde{W}_{out}(\omega_j, \theta_t)\tilde{E}(\omega_j, \theta)W_{in}(\omega_j) \tag{B.18}$$

again indicates that the parameter θ to be estimated appears linearly in (B.18).

Although the free parameter θ appears linearly in (B.18), writing down a matrix representation for the weighted equation error $\tilde{E}_w$ similar to (B.16) would inevitably lead to additional (large) sparse matrices that need to be stored in order to compute the least squares solution. The sparse matrices arise from the frequency dependent output (and input) weighting that need to be incorporated (Bayard, 1994). Furthermore, the parameter θ might have a structure containing zero entries at prespecified locations if a non-full polynomial parametrization is being used.

To avoid the computational and memory storage issues that arise from dealing with (large) sparse matrices and to be able to take into account the specific structure that might be present in the parameter θ, a fairly simple and straightforward computational procedure based on Kronecker calculus is presented here. For this purpose consider the following definition.

Definition B.3-1 *Consider two matrices $X \in \mathbb{C}^{n_1 \times n_2}$ and $Y \in \mathbb{C}^{m_1 \times m_2}$, then the Kronecker vector $\mathrm{vec}(X) \in \mathbb{C}^{n_1 n_2 \times 1}$ and the Kronecker product $X \otimes Y \in \mathbb{C}^{n_1 m_1 \times n_2 m_2}$ are respectively defined by $\mathrm{vec}(X) := [x_1 \; \cdots \; x_{n_2}]^T$ and*

$$
X \otimes Y := \begin{bmatrix} x_{1,1}Y & \cdots & x_{1,n_2}Y \\ \vdots & \ldots & \vdots \\ x_{n_1,1}Y & \cdots & x_{n_1,n_2}Y \end{bmatrix}
$$

where $x_{i,j}$ and x_j for $i \in 1, \ldots, n_1$ and $j \in 1, \ldots, n_2$ are used to denote respectively the (i,j)th entry in X and the jth column in X.

The Kronecker product is a well known concept (Bellman, 1970) and by stacking the columns of a matrix to obtain the corresponding Kronecker vector as mentioned in Definition B.3-1, the following result can be obtained.

Proposition B.3-2 *Consider (complex) matrices X, Y and Z with appropriate dimensions, such that the matrix product $C := XYZ$ is well defined. Then $\mathrm{vec}(C)$ satisfies*

$$
\mathrm{vec}(C) = [Z^T \otimes X]\mathrm{vec}(Y).
$$

Proof: The proof can be found in Bellman (1970). $\qquad\qquad\square$

On the basis of Proposition B.3-2, the Kronecker vector of the input/output weighted equation error $\tilde{E}_w(\omega_j, \hat{\theta}_{t-1}, \theta)$ in (B.18) can be written as

$$
\mathrm{vec}(\tilde{E}_w) = \mathrm{vec}(\tilde{W}_{out}GW_{in}) - [[\Phi W_{in}]^T \otimes \tilde{W}_{out}]\mathrm{vec}(\theta)
$$

wherein the arguments ω_j, $\hat{\theta}_{t-1}$ and θ are left out, to avoid notational issues. As the Frobenius-norm satisfies $\|X\|_F^2 = \|\mathrm{vec}(X)\|_F^2$ for an arbitrary matrix X, the

Frobenius-norm on $\tilde{E}_w$ can be characterized by a matrix representation formed by stacking $\text{vec}(\tilde{E}_w(\omega_j, \hat{\theta}_{t-1}, \theta))$ row-wise for $j \in 1, \ldots, N$. This yields the following estimate

$$
\begin{aligned}
\hat{\theta} &= \arg \min_{\theta \in \mathbb{R}} \; \|\text{vec}(\tilde{E}_w(\hat{\theta}_{t-1}, \theta))\|_F^2 \\
&= \arg \min_{\theta \in \mathbb{R}} \; \|\mathsf{G}_w - \mathsf{P}_w \bar{\theta}\|_F^2
\end{aligned}
\tag{B.19}
$$

where $\bar{\theta} = \text{vec}(\theta) \in \mathbb{R}^{p(mb+pa) \times 1}$ according to (B.11). Furthermore, $\mathsf{G}_w \in \mathbb{R}^{2pmN \times 1}$ and $\mathsf{P}_w \in \mathbb{R}^{2pmN \times p(mb+pa)}$ are matrices that can be found by row-wise stacking of the real and imaginary part of respectively $\text{vec}(\tilde{W}_{out}(\omega_j, \hat{\theta}_{t-1}) G(\omega_j) W_{in}(\omega_j))$ and $\text{vec}([\Phi(\omega_j) W_{in}(\omega_j)]^T \otimes \tilde{W}_{out}(\omega_j, \hat{\theta}_{t-1}))$ for $j \in 1, \ldots, N$.

The regression matrix P_w in (B.19) does not exhibit any sparse matrix structure as occurs e.g. in the method of Bayard (1994). In fact, $2pmN \times p(mb + pa)$ entries is the smallest dimension of the regression matrix P_w in order to compute a least squares parameter $\hat{\theta}$ that has $p(mb+pa)$ unknown entries (for a a left full polynomial parametrization) on the basis of N complex frequency domain points of a $p \times m$ multivariable system. In this way memory storage problems are avoided directly as much as possible.

As the parameter θ is converted into a column parameter $\bar{\theta} = \text{vec}(\theta)$, any prespecified zero entries in $\bar{\theta}$ can be incorporated in the estimation of the parameter relatively easy. This can be done by omitting the columns in the regression matrix P_w that correspond to the zero entries in $\bar{\theta}$ and thereby reducing the size of the parameter to be estimated directly.

B.3.3 Schur weighting

Consider the Schur or element-wise frequency weighted curve fit error in (B.4) and rewrite this into

$$
E_s(\omega_j, \theta) = S(\omega_j). * [A(\xi(\omega_j)^{-1}, \theta)^{-1} \tilde{E}(\omega_j, \theta)]
\tag{B.20}
$$

where the equation error $\tilde{E}(\omega_j, \theta)$ was defined in (B.13). Using a similar approach of iterative minimization steps as used in Section B.3.1, the parameter θ in $A(\xi(\omega_j)^{-1}, \theta)^{-1}$ in (B.20) is fixed to an estimate $\hat{\theta}_{t-1}$ obtained from the previous step $t - 1$. Consequently, the weighted equation error $\tilde{E}_s$ defined by

$$
\begin{aligned}
\tilde{E}_s(\omega_j, \hat{\theta}_{t-1}, \theta) := \\
S(\omega_j). * [A(\xi(\omega_j)^{-1}, \hat{\theta}_{t-1})^{-1} \tilde{E}(\omega_j, \theta)]
\end{aligned}
$$

again indicates that the parameter θ to be estimated appears linearly. Finally, it can be verified (leaving out the arguments ω_j, $\xi(\omega_j)^{-1}$, $\hat{\theta}_{t-1}$ and θ) that $\text{vec}(\tilde{E}_s)$ can be rewritten into

$$
\text{vec}(S. * [A^{-1}G]) - \text{diag}(\text{vec}(S))[\Phi^T \otimes A^{-1}]\text{vec}(\theta)
\tag{B.21}
$$

by using the result of Proposition B.3-2. Hence, stacking $\text{vec}(\tilde{E}_s(\omega_j, \hat{\theta}_{t-1}, \theta))$ row wise for each $j \in 1, \ldots, N$ will yield a similar expression for the minimizing argument $\hat{\theta}$ as given in (B.19). However, the matrix G_w in (B.19) now contains real and imaginary part of $\text{vec}(S(\omega_j) . * [A(\xi(\omega_j)^{-1}, \hat{\theta}_{t-1})G(\omega_j)])$, whereas P_w in (B.19) will consist of the real and imaginary part of $\text{diag}(\text{vec}(S(\omega_j)))[\Phi(\omega_j)^T \otimes A^{-1}(\xi(\omega_j)^{-1}, \hat{\theta}_{t-1})]$ for $j \in 1, \ldots, N$. Hence, the same computational procedure can be used to incorporate an element-by-element weighted curve fit error (B.4) by a slight modification of the matrices in (B.19).

266

List of Symbols

Sets

$\mathbb{C}$	Set of complex numbers
$\mathcal{G}$	Set of parametrized deterministic models; Set of frequency domain measurements
$\mathcal{H}_2$	Set of all complex valued functions which are analytic outside the unit disk and squared integrable over the unit circle
$\mathcal{H}_\infty$	Set of all complex valued functions which are analytic and uniformly bounded outside the unit disc and on the unit circle
l_1	Set of absolute summable sequences
$\mathcal{L}_2$	Set of all complex valued functions which are squared integrable on the unit circle
$\mathcal{L}_\infty$	Set of all complex valued functions which are essentially bounded on the unit circle
$\mathcal{M}$	Set of parametrized deterministic models and noise models
$\mathcal{P}, \mathcal{P}_i, \mathcal{P}_{i+1}$	Set of (linear) dynamical models
$\mathbb{R}$	Set of real numbers
$\mathcal{P}(\omega_j)$	Convex frequency response region
$R\mathcal{H}_\infty$	Set of all real rational functions in $\mathcal{H}_\infty$.
$R\mathcal{H}_2$	Set of all real rational functions in $\mathcal{H}_2$.
$\mathcal{S}$	Set of intermediate models used in uncertainty bounding estimation

Abbreviations

ARX	Auto Regressive with eXogenous input
ARMAX	Auto Regressive Moving Average with eXogenous input
DFT	Discrete Fourier Transform
ETFE	Empirical Transfer Function Estimate
FIR	Finite Impulse Response

lcf	Left Coprime Factorization
LTI	Linear Time Invariant
MIMO	Multi Input Multi Output
$nlcf$	Normalized Left Coprime Factorization
$nrcf$	Normalized Right Coprime Factorization
ORTFIR	ORThogonal Finite Impulse Response
(P)RBS	(Pseudo) Random Binary Sequence
(P)RPS	(Pseudo) Random Phase Sequence
rcf	Right Coprime Factorization
SISO	Single Input Single Output
MISO	Multi Input Single Output
SIMO	Single Input Multi Output
w.p.	With Probability

Plant, Model and Controller

C	Arbitrary feedback controller
C_i	Controller implemented in feedback connection
C_{i+1}	(Re)designed feedback controller
C_x	Auxiliary feedback controller
$G,\hat{G}$	(Estimated) frequency response
H	Arbitrary noise system or noise model
$\hat{H}$	Noise model
H_o	(Unknown) noise filter
P	Arbitrary system or model
$\hat{P},\hat{P}_i,\hat{P}_{i+1}$	Deterministic model or nominal model
P_o	(Unknown) plant
P_x	Auxiliary model
T_o	(Unknown) plant and (unknown) noise model
T	Arbitrary deterministic and noise model
$\hat{T}$	(Nominal) deterministic and noise model

Coprime Factorizations

F	Filter for auxiliary signal x
(N,D)	Right coprime factorization of a system P
$(\tilde{D},\tilde{N})$	Left coprime factorization of a system P
$(\hat{N},\hat{D})$	Right coprime factorization of a (nominal) model $\hat{P}$
(N_x,D_x)	Right coprime factorization of an auxiliary model P_x

(N_o, D_o)	Right coprime factorization of the (unknown) plant P_o
$(N_{o,F}, D_{o,F})$	Accessible right coprime factorization of the (unknown) plant P_o
$(\hat{N}_{o,F}, \hat{D}_{o,F})$	(Estimated) frequency domain of the right coprime factorization $(N_{o,F}, D_{o,F})$
(N_c, D_c)	Right coprime factorization of a controller C
$(\tilde{D}_c, \tilde{N}_c)$	Left coprime factorization of a controller C
(X, Y)	Bezout factor associated to a right coprime factorization
$(\tilde{X}, \tilde{Y})$	Bezout factor associated to a left coprime factorization

Model Perturbations

$\mathcal{P}, \mathcal{P}_i, \mathcal{P}_{i+1}$	Set of (linear) dynamical models
$\mathcal{P}(\omega_j)$	Convex frequency response region
$\mathcal{S}$	Set of intermediate models used in uncertainty bounding estimation
Δ_R	Model perturbation in Youla parametrization
$\hat{\Delta}_R$	(Estimated) frequency domain data of Δ_R
$\bar{\Delta}_R$	(Estimated) upper bound for Δ_R
(V, W)	Arbitrary pair of stable and stable invertible weighting functions
$(\hat{V}, \hat{W})$	Estimated pair of stable and stable invertible weighting functions
$\gamma, \gamma_i, \gamma_{i+1}$	(Robust) performance level; Performance value

Feedback

$\mathcal{T}(P, C)$	Feedback connection of P and C
$T(P, C)$	The map from external reference signals r_2, r_1 to output y and input u in a feedback connection
$T_{ext}(P, C_i, C)$	Extended $T(P, C)$ matrix
$H(P, C)$	The map from external reference signals r_2, r_1 to inputs u_c and u in a feedback connection
$J(P, C)$	Control objective function
S_{in}	The input sensitivity function
$\hat{S}_{in}$	Estimated input sensitivity function
S_{out}	The output sensitivity function

Signals

d	uncertainty output signal
e	(white) noise signal or performance output signal

ε	prediction error (signal)
r, r_1, r_2	reference signal
u	(plant) input signal
u_c	(control) input signal or servo error
v	noise signal
w	performance input signal
x	auxiliary or intermediate (input) signal
y	output signal
y_c	(control) output signal
z	(intermediate) output signal or uncertainty input signal

Miscellaneous symbols and notations

A,B,C,D	State space matrices
θ	Free parameter
$\hat{\theta}$	Estimated parameter
$\theta^\star$	Estimated (asymptotic) parameter
$\hat{\theta}_N$	Estimated (finite time) parameter
U_1,U_2	Performance weighting filters
Q	Coefficient matrix
M	Closed-loop Coefficient matrix
x	x-direction in wafer stage
y	y-direction in wafer stage

List of Figures

Bibliography

Adamjan, V.M., D.Z. Arov and M.G. Krein (1978). Infinite block Hankel matrices and related extension problems. *AMS Transl..* **111**, 133–156.

Aling, H. (1989). *Identification of closed loop systems; identifiability, recursive algorithms and application to power plant.* PhD thesis. Delft University of Technology, Delft, the Netherlands.

Anantharam, V. (1985). On stabilization and the existence of coprime factorizations. *IEEE Trans. on Automatic Control.* **30**, 1030–1031.

Anderson, B.D.O. and J.B. Moore (1990). *Optimal Control: Linear Quadratic Methods.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Anderson, B.D.O. and M.R. Gevers (1982). Identifiability of linear stochastic systems operating under linear feedback. *Automatica.* **18**, 195–213.

Antsaklis, P.J. (1986). Proper stable transfer matrix factorizations and internal system descriptions. *IEEE Trans. on Automatic Control.* **31**, 634–638.

Ariaans, L.. (1997). *Structure in practical model error bounds.* PhD thesis. Eindhoven University of Technology, Eindhoven, the Netherlands.

Åström, K.J. and B. Wittenmark (1971). Problems of identification and control. *Journal of Mathematical Analysis and Applications.* **34**, 90–113.

Åström, K.J. and B. Wittenmark (1989). *Adaptive Control.* Addison-Wesley, New York, USA.

Åström, K.J. and B. Wittenmark (1990). *Computer Controlled Systems: theory and design.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Åström, K.J. and P. Eykhoff (1971). System identification - a survey. *Automatica.* **7**, 123–167.

Åström, K.J. and T. Bohlin (1965). *Numerical identification of linear dynamic systems from normal operating records.* Theory of Self-Adaptive Control Systems, Plenum Press, New York, USA, P.H. Hammond (Ed.).

Åström, K.J. and T. Hägglund (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica.* **20**, 645–650.

Athans, M. and P. Falb (1966). *Optimal Control.* McGraw-Hill, New York, USA.

Bai, E.W. and S. Raman (1994). Robust system identification with noisy experimental data - projection operator and linear algorithms. *Automatica*. **30**, 1203–1206.

Bakri, T.M. (1988). Generalized bezoutians and invariant subspaces. *IEEE Trans. on Automatic Control*. **33**, 880–883.

Bayard, D.S. (1992). Statistical plant set estimation using Schroeder-phased multisinusoidal input design. In *Proc. American Control Conference*. Chicago, IL, USA. pp. 2988–2995.

Bayard, D.S. (1994). High order multivariable transfer function curve fitting: algorithms, sparse matrix methods and experimental results. *Automatica*. **30**, 1439–1444.

Bayard, D.S. and R.Y. Chiang (1993). A frequency domain approach to identification, uncertainty characterization and robust control design. In *Proc. 32nd IEEE Conference on Decision and Control*. San Antonio, USA. pp. 2266–2271.

Bayard, D.S., Y. Yam and E. Mettler (1992). A criterion for joint optimization of identification and robust control. *IEEE Trans. on Automatic Control*. **37**, 986–991.

Bitmead, R.R. (1993). Iterative control design approaches. In *Prepr. 12th IFAC World Congress*. Vol. 9. Sydney, Australia. pp. 381–384.

Bitmead, R.R. and Z. Zang (1991). An iterative identification and control strategy. In *Proc. 1st European Control Conference*. Grenoble, France. pp. 1396–1400.

Bitmead, R.R., M.R. Gevers and V. Wertz (1990*a*). *Adaptive Optimal Control - The Thinkings Man's GPC*. Prentice-Hall, Sydney, Australia.

Bitmead, R.R., M.R. Gevers and Z. Zang (1990*b*). Interplay between control law selection and closed loop adaption. In *Prepr. 11th IFAC World Congress*. Vol. 4. Tallinn, Estonia, USSR. pp. 202–207.

Black, H.S. (1934). Stabilized feedback amplifiers. *Bell System Technical Journal*. **13**, 1–18.

Bode, H.W. (1945). *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, USA.

Böling, J.M. and P.M. Mäkilä (1995). On control relevant criteria in H_∞ identification. In *Proc. American Control Conference*. Seatle, USA. pp. 3076–3080.

Bongers, P.M.M. (1991). On a new robust stablity margin. *Proc. MTNS-91: Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing*. pp. 377–382.

Bongers, P.M.M. (1994). *Modeling and Identification of Flexible Wind Turbines and a Factorizational Approach to Robust Control*. PhD thesis. Delft University of Technology, Delft, the Netherlands.

Bongers, P.M.M. and O.H. Bosgra (1990). Low order robust H_∞ controller synthe-

sis. In *Proc. 29th IEEE Conference on Decision and Control.* Honolulu, Hawaii. pp. 194–199.

Bongers, P.M.M. and O.H. Bosgra (1993). Normalized coprime factorizations for systems in generalized state space form. *IEEE Trans. on Automatic Control.* **38**, 348–350.

Bongers, P.M.M. and P.S.C. Heuberger (1990). Discrete normalized coprime factorization and fractional balanced reduction. *Lecture Notes in Control and Information Sciences.* **144**, 307–313.

Box, G.E.P. and G.M. Jenkins (1970). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco, USA.

Boyd, S., L.E. Ghaoui, E. Feron and V. Balakrishnan (1994). *Linear Matrix Inequalities in Systems and Control Theory.* SIAM Studies in applied mathematics, Philadelphia, Pennsylvania, USA.

Boyd, S.P. and C.H. Barrat (1991). *Linear Controller design - Limits of Performance.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Bretthauer, G. and F. Heckert (1991). A priori identifiability of closed loop subsystems for correlated outer signals. In *Proc. 9th IFAC/IFORS symposium on identification and system parameter estimation.* Budapest, Hungary. pp. 815–820.

Caines, P.E. and C.W. Chan (1975). Feedback between stationary stochastic processes. *IEEE Trans. on Automatic Control.* **20**, 498–508.

Caines, P.E. and C.W. Chan (1976). Estimation, identification and feedback. *System Identification: Advances and Case Studies.* pp. 349–406. R.K. Mehra and D.G. Lainiotis (Eds.).

Caines, P.E. and M. Baykal-Gürsoy (1989). On the L_∞ consistency of L_2 estimators. *Systems & Control Letters.* **12**, 71–76.

Callier, F.M. and C.A. Desoer (1991). *Linear System Theory.* Springer-Verlag, New York, USA.

Ceton, C., P.M.R. Wortelboer and O.H. Bosgra (1993). Frequency weighted closed-loop balanced reduction. In *Proc. 2nd European Control Conference.* Groningen, the Netherlands. pp. 697–701.

Chen, C.T. (1984). *Linear System Theory and Design.* CBS College Publishing, New York, USA.

Chen, C.T. and C.A. Desoer (1982). Algebraic theory for robust stability of interconnected systems: necessary and sufficient conditions. In *Proc. 21st IEEE Conference on Decision and Control.* Orlando, USA. pp. 491–494.

Chen, J. (1995). Sensitivity integral relations and design trade-offs in linear multivariable feedback systems. *IEEE Trans. on Automatic Control.* **40**, 1700–1716.

Chen, J. and C.N. Nett (1995a). The Carathéodory-Fejér problem and H_∞/l_1 identification. *IEEE Trans. on Automatic Control*. **40**, 729–735.

Chen, J. and C.N. Nett (1995b). Sensitivity integrals for multivariable discrete-time systems. *Automatica*. **31**, 1113–1124.

Chen, J., C.N. Nett and M.K.H. Fan (1992). Worst-case identification in H_∞: validation of apriori information, essentially optimal algorithms and error bounds. In *Proc. American Control Conference*. Chicago, IL, USA. pp. 251–257.

Chen, T. and B. Francis (1995). *Optimal Sampled-Data Control Systems*. Springer-Verlag, London, England.

Chou, C.T. (1994). *Geometry of Linear Systems and Identification*. PhD thesis. University of Cambridge, Cambridge, England.

Corrêa, G.O. and K. Glover (1984). Pseudo-canonical forms, identifiable parametrizations and simple parameter estimation for linear multivariable systems. *Automatica*. **20**, 429–452.

Corrêa, G.O. and K. Glover (1986). On the choice of parametrization for identification. *IEEE Trans. on Automatic Control*. **31**, 8–15.

Dahleh, M.A. and B. Boyd Pearson Jr (1987a). L_1-optimal compensators for continues time systems. *IEEE Trans. on Automatic Control*. **32**, 889–895.

Dahleh, M.A. and B. Boyd Pearson Jr (1987b). l_1-optimal feedback controllers for MIMO discrete time systems. *IEEE Trans. on Automatic Control*. **32**, 314–322.

de Bruyne, F. (1996). *Aspects on System Identification for Robust Process Control*. PhD thesis. Université Catholique de Louvain, Louvain-La-Neuve, Belgium.

de Bruyne, F. and M.R. Gevers (1994). Identification for control: can the optimal restricted complexity model always be identified?. In *Proc. 33rd IEEE Conference on Decision and Control*. Lake Buena Vista, USA. pp. 3912–3917.

de Callafon, R.A. and P.M.J. Van den Hof (1995a). Control relevant identification for H_∞-norm based performance specifications. In *Proc. 34th IEEE Conference on Decision and Control*. New Orleans, USA. pp. 3498–3503.

de Callafon, R.A. and P.M.J. Van den Hof (1995b). Filtering and parametrization issues in feedback relevant identification based on fractional model representations. In *Proc. 3rd European Control Conference*. Vol. 1. Rome, Italy. pp. 441–446.

de Callafon, R.A. and P.M.J. Van den Hof (1997). Suboptimal feedback control by a scheme of iterative identification and control design. *Mathematical Modelling of Systems*. **3**(1), 77–101.

de Callafon, R.A., D. Roover and P.M.J. Van den Hof (1996a). Multivariable least squares frequency domain identification using polynomial matrix fraction descriptions. In *Proc. 35th IEEE Conference on Decision and Control*. Kobe, Japan. pp. 2030–2035.

de Callafon, R.A., P.M.J. Van den Hof and D.K. de Vries (1994). Identification and control of a compact disc mechanism using fractional representations. In *Prepr. 10th IFAC Symposium on System Identification*. Vol. 2. Copenhagen, Denmark. pp. 121–126.

de Callafon, R.A., P.M.J. Van den Hof and M. Steinbuch (1993). Control relevant identification of a compact disc pick-up mechanism. In *Proc. 32nd IEEE Conference on Decision and Control*. San Antonio, USA. pp. 2050–2055.

de Callafon, R.A., P.M.J. Van den Hof and P.M.M. Bongers (1996*b*). A unified approach to stability robustness for uncertainty descriptions based on fractional model representations. *IEEE Trans. on Automatic Control*. **41**, 723–727.

de Roover, D. (1997). *Motion Control of a Wafer Stage, a design approach for speeding up IC production*. PhD thesis. Delft University of Technology, Delft, the Netherlands.

de Roover, D. and R.N. van Marrewijk (1995). Systematic analysis of the dynamic behaviour of an electro-mechanical servosystem. *Selected Topics in Identification, Modelling and Control*. **8**, 75–83. O.H. Bosgra, P.M.J. Van den Hof and C.W. Scherer (Eds.).

de Roover, D., O.H. Bosgra, F.B. Sperling and M. Steinbuch (1996). High-performance motion control of a flexible mechanical servomechanism. *Selected Topics in Identification, Modelling and Control*. **9**, 69–77. O.H. Bosgra, P.M.J. Van den Hof and C.W. Scherer (Eds.).

de Vries, D.K. (1994). *Identification of Model Uncertainty for Control Design*. PhD thesis. Delft University of Technology, Delft, the Netherlands.

de Vries, D.K. and P.M.J. Van den Hof (1995). Quantification of uncertainty in transfer function estimation: a mixed probabilistic - worst-case approach. *Automatica*. **31**, 543–558.

Deistler, M. and B.D.O. Anderson (1989). Linear dynamic errors-in-variables models, some structure theory. *Journal of Econometrics*. **41**, 39–63.

Desoer, C.A. and A.N. Gündes (1988). Algebraic theory of linear time-invariant feedback systems with two-input two-output plant and controller. *Int. Journal of Control*. **47**, 33–51.

Desoer, C.A. and M. Vidyasagar (1975). *Feedback Systems: Input-Output Properties*. Academic Press, New York, USA.

Desoer, C.A., R.W. Liu, J. Murray and R. Saeks (1980). Feedback system design: the fractional representation approach to analysis and synthesis. *IEEE Trans. on Automatic Control*. **25**, 399–412.

Donkelaar, E.T. and P.M.J. Van den Hof (1996). Analysis of closed-loop identification with a tailor-made parametrization. *Selected Topics in Identification, Modelling and Control*. **9**, 17–24. O.H. Bosgra and P.M.J. Van den Hof (Eds.).

Dötsch, H.G.M. (1998). *Identification for Control Design with Application to a Compact Disk Mechanism*. PhD thesis. Delft University of Technology, Delft, the Netherlands.

Doyle, J.C. (1978). Guaranteed margins for LQG regulators. *IEEE Trans. on Automatic Control*. **23**, 756–757.

Doyle, J.C. (1979). Robustness of multiloop linear feedback systems. In *Proc. 17th IEEE Conference on Decision and Control*. San Diego, USA. pp. 12–18.

Doyle, J.C. (1982). Analysis of feedback systems with structured uncertainties. *Proc. IEE on Control Theory and Applications*. **129**(D), 242–250.

Doyle, J.C. (1984). *Lecture notes in advances in multivariable control*. Honeywell workshop, Mineapolis, USA.

Doyle, J.C., A. Packard and K. Zhou (1991). Review of LFTs, LMIs and μ. In *Proc. 30th IEEE Conference on Decision and Control*. Brighton, England. pp. 1227–1232.

Doyle, J.C. and G. Stein (1981). Multivariable feedback design: concepts for a classical/modern synthesis. *IEEE Trans. on Automatic Control*. **26**, 4–16.

Doyle, J.C., B.A. Francis and A.R. Tannenbaum (1992). *Feedback Control Theory*. MacMillan, New York, USA.

Doyle, J.C., K. Glover, P.P. Khargonegar and B.A. Francis (1989). State space solutions to standard H_2 and H_∞ control problems. *IEEE Trans. on Automatic Control*. **34**, 831–847.

El-Sakkary, A.K. (1985). The gap metric: Robustness of stabilization of feedback systems. *IEEE Trans. on Automatic Control*. **30**, 240–247.

Evans, W.R. (1950). *Control System Dynamics*. McGraw-Hill, New York, USA.

Eykhoff, P. (1974). *System Identification*. John Wiley & Sons, New York, USA.

Farison, J.B., R.E. Graham and R.C. Shelton (1967). Identification and control of linear discrete time systems. *IEEE Trans. on Automatic Control*. **12**, 438–442.

Francis, B.A. (1987). *A course in H_∞ control theory*. Springer-Verlag, Berlin, Germany.

Francis, B.A. and J.C. Doyle (1987). Linear controller theory with an H_∞ optimality criterion. *SIAM Journal on Control and Optimization*. pp. 815–844.

Franklin, G.F., J.D. Powel and A. Emami-Naeini (1991). *Feedback Control of Dynamic Systems*. Addison-Wesley, New York, USA.

Freudenberg, J.S. and D.P. Looze (1985). Right half plane poles and zeros and design tradeoffs in feedback systems. *IEEE Trans. on Automatic Control*. **30**, 555–565.

Freudenberg, J.S. and D.P. Looze (1987). A sensitivity tradeoff for plants with time delay. *IEEE Trans. on Automatic Control*. **32**, 99–104.

Gauss, K.F. (1901). *Theoria Motus Corporum Coelestium.* Liber II, in Werke, Julius Springer, Berlin, Germany.

Gauss, K.F. (1963). *Theory of the motion of the heavily bodies moving about the sun in conic sections.* Dover, New York, USA.

Georgiou, T.T. (1988). On the computation of the gap metric. *Systems & Control Letters.* **11**, 253–257.

Georgiou, T.T. and M.C. Smith (1990). Optimal robustness in the gap metric. *IEEE Trans. on Automatic Control.* **35**, 673–686.

Gevers, M.R. (1978). On the identification of feedback systems. In *Prepr. 4th IFAC Symp. on System Identification.* Tbilisi, USSR. pp. 1621–1630.

Gevers, M.R. (1986). ARMA models, their Kronecker indices and their McMillan degree. *Int. Journal of Control.* **43**(6), 1745–1761.

Gevers, M.R. (1993). Towards a joint design of identification and control?. In *Essays on Control: Perspectives in the Theory and its Applications.* H.L. Trentelman and J.C. Willems (Eds.). Birkhäuser, Boston. pp. 111–151.

Gevers, M.R. and B.D.O. Anderson (1981). Representations of jointly stationary stochastic feedback processes. *Int. Journal of Control.* **33**(5), 777–809.

Gevers, M.R. and L. Ljung (1986). Optimal experiment design with respect to the intended model application. *Automatica.* **22**, 543–554.

Gevers, M.R. and V. Wertz (1984). Uniquely identifiable state-space and ARMA parametrizations for multivariable systems. *Automatica.* **20**, 333–347.

Glover, K. (1984). All optimal Hankel norm approximations of linear multivariable systems with their L_∞-error bounds. *Int. Journal of Control.* **39**(6), 1115–1193.

Glover, K. and D. McFarlane (1989). Robust stabilization of normalized coprime factor plant descriptions with H_∞-bounded uncertainty. *IEEE Trans. on Automatic Control.* **34**, 821–830.

Glover, K. and J.C. Doyle (1988). State-space formulae for all stabilizing controllers that satisfy an H_∞ norm bound and relations to risk sensitivity. *Systems and Control Letters.* **11**, 167–172.

Glover, K. and J.C. Willems (1974). Parametrizations of linear dynamical systems - canonical forms and identifiability. *IEEE Trans. on Automatic Control.* **19**, 640–646.

Goodwin, G.C., M. Gevers and B.M. Ninness (1992). Quantifying the error in estimated transfer functions with application to model order selection. *IEEE Trans. on Automatic Control.* **37**, 913–928.

Green, M. and D.J.N. Limebeer (1995). *Linear Robust Control.* Prentice-Hall.

Gu, G. and P.P. Khargonekar (1992). A class of algorithms for identification in H_∞. *Automatica.* **28**, 299–312.

Guidorzi, R. (1975). Canonical structures in the identification of multivariable systems. *Automatica.* **11**, 361–374.

Gustavson, I., L. Ljung and T. Söderström (1977). Identification of processes in closed loop - identifiability and accuracy aspects. *Automatica.* **13**, 59–75.

Gustavson, I., L. Ljung and T. Söderström (1981). Choice and effect of different feedback configurations. *Trends and Progress in System Identification.* pp. 367–388. P. Eykhoff (ed.).

Hakvoort, R.G. (1990). Optimal experiment design for prediction error identification in view of feedback design. *Selected topics in identification, modelling and control.* pp. 71–78. O.H. Bosgra and P.M.J. Van den Hof (Eds.).

Hakvoort, R.G. (1994). *System Identification for Robust Process Control: Nominal Models and Error Bounds.* PhD thesis. Delft University of Technology, Delft, the Netherlands.

Hakvoort, R.G. and P.M.J. Van den Hof (1994*a*). Frequency domain curve fitting with maximum amplitude criterion and guaranteed stability. *Int. Journal of Control.* **60**(5), 809–825.

Hakvoort, R.G. and P.M.J. Van den Hof (1994*b*). An instrumental variable procedure for the identification of probabilistic frequency response uncertainty regions. In *Proc. 33rd IEEE Conference on Decision and Control.* Lake Buena Vista, USA. pp. 3596–3601.

Hakvoort, R.G. and P.M.J. Van den Hof (1997). Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors. *IEEE Trans. on Automatic Control.* **42**, 1516–1528.

Hakvoort, R.G., R.J.P. Schrama and P.M.J. Van den Hof (1992). Approximate identification in view of LQG feedback design. In *Proc. American Control Conference.* Chicago, IL, USA. 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.* **30**, 679–690.

Hanselmann, H. (1987). Implementation of digital controllers - a survey. *Automatica.* **23**, 7–32.

Hansen, F.R. (1989). *A fractional representation approach to closed loop system identification and experiment design.* PhD thesis. Stanford University, Stanford, USA.

Hansen, F.R. and G.F. Franklin (1988). On a fractional representation approach to closed-loop experiment design. In *Proc. American Control Conference.* Atlanta, USA. pp. 1319–1320.

Hansen, F.R., G.F. Franklin and R. Kosut (1989). Closed-loop identification via the fractional representation: Experiment design. In *Proc. American Control Conference.* Pittsburgh, USA. pp. 1422–1427.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1989). H_∞ identification of stable LSI systems: A scheme with direct application to controller design. In *Proc. American Control Conference*. Pittsburgh, USA. pp. 1428–1434.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1990). Identification in H_∞: linear algorithms. In *Proc. American Control Conference*. San Diego, USA. pp. 2418–2423.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1991). Control oriented system identification: a worst-case/deterministic approach in H_∞. *IEEE Trans. on Automatic Control*. **36**, 1163–1176.

Helmicki, A.J., C.A. Jacobson and C.N. Nett (1993). Least squares methods for H_∞ control-oriented system identification. *IEEE Trans. on Automatic Control*. **38**, 819–826.

Heuberger, P.S.C. (1991). *On Approximate System Identification using System Based Orthonormal Functions*. PhD thesis. Delft University of Technology, Delft, the Netherlands.

Heuberger, P.S.C., P.M.J. Van den Hof and O.H. Bosgra (1995). A generalized orthonormal basis for linear dynamical systems. *IEEE Trans. on Automatic Control*. **40**, 451–465.

Hewlett-Packard (1985). *Laser Transducer System 5501A - System Operating and Service Manual*. Hewlett-Packard Company, Everett, WA, USA.

Hjalmarsson, H. (1993). *Aspects on Incomplete Modelling in System Identification*. PhD thesis. Linköping University, Linköping, Sweden.

Hjalmarsson, H., M.R. Gevers and F. de Bruyne (1996). For model based control design, closed loop identification gives better performance. *Automatica*. **12**, 1659–1673.

Hjalmarsson, H., S. Gunnarson and M.R. Gevers (1995). Optimality and suboptimality of iterative identification and control design schemes. In *Proc. American Control Conference*. Seattle, USA. pp. 2559–2563.

Hjalmarsson, H., S. Gunnarsson and M.R. Gevers (1994). A convergent iterative restricted complexity control design scheme. In *Proc. 33rd IEEE Conference on Decision and Control*. Lake Buena Vista, USA. pp. 1735–1740.

Ho, B.L. and R.E. Kalman (1966). Effective construction of linear state-variable models from input/output functions. *Regelungstechnik*. **14**, 545–548.

Horowitz, I.M. (1959). Fundamental theory of automatic linear feedback control systems. *IRE Trans. on Automatic Control*. **4**, 5–19.

Horowitz, I.M. (1963). *Synthesis of Feedback Systems*. Academic Press, New York, USA.

Iglesias, P.A. (1990). On the use of robust regulators in adaptive control. In *Proc. 11th IFAC World Congress*. Vol. 4. Tallinn, Estonia. pp. 214–219.

Jackson, L.B. (1991). *Signals, Systems and Transforms*. Addison-Wesley, New York, USA.

Janssen, P.H.M. (1988). *On Model Parametrizations and Model Structure Selection for Identification of MIMO Systems*. PhD thesis. Eindhoven University of Technology, Eindhoven, The Netherlands.

Johansson, R. (1993). *System Modelling and Identification*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Jørgensen, S.B. and K.M. Hangos (1995). Grey box modelling for control: qualitive models as a unifying framework. *Int. Journal of Adaptive Control and Signal Processing*. **9**(6), 547–562.

Khargonekar, P.P., G. Gu and J. Friedman (1996). Identification in H_∞: theory and applications. In *Identification, Adaptation and Learning, the science of learning models from data*. NATO ASI series, Vol. 153, Springer-Verlag, Berlin, S. Bittanti, G. Picci (Eds.). Como, Italy. pp. 266–288.

Khargonekar, P.P., I.R. Petersen and M.A. Rotea (1988). H_∞-optimal control with state feedback. *IEEE Trans. on Automatic Control*. **33**, 786–788.

Kollár, I. (1994). *Frequency domain system identification toolbox*. The Mathworks, Inc.

Kung, S.Y. (1978). A new identification and model reduction algorithm via singular value decomposition. In *Proc. 12th Asilomar Conference on Circuits, Systems and Computers*. Pacific Grove, USA. pp. 705–714.

Kwakernaak, H. (1985). Minimax frequency domain performance and robustness optimization of linear feedback systems. *IEEE Trans. on Automatic Control*. **30**, 994–1004.

Kwakernaak, H. (1993). Robust control and H_∞ optimization - tutorial paper. *Automatica*. **29**, 255–273.

Kwakernaak, H. and R. Sivan (1972). *Linear Optimal Control Systems*. John Wiley & Sons, New York, USA.

Kwakernaak, H. and R. Sivan (1991). *Modern Signals and Systems*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Landau, I.D. and K. Boumaïza (1996). An output error recursive algorithm for unbiased identification in closed loop. In *Prepr. 13th IFAC World Congress*. San Francisco, USA. pp. 215–220.

Larimore, W.E. (1990). System identification, reduced-order filtering and modeling via canonical variate analysis. In *Proc. American Control Conference*. San Diego, USA. pp. 445–451.

Lauer, H., R.N. Lesnick and L.E. Matson (1960). *Servomechanism Fundamentals*. McGraw-Hill, New York, USA.

Lee, W.S., B.D.O. Anderson, I.M.Y. Mareels and R.L. Kosut (1995). On some key issues in the windsurfer approach to adaptive robust control. *Automatica*. **31**, 1619–1636.

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. In *Proc. American Control Conference*. Chicago, IL, USA. pp. 2834–2841.

Lee, W.S., B.D.O. Anderson, R.L. Kosut and I.M.Y. Mareels (1993*a*). A new approach to adaptive robust control. *Int. Journal of Adaptive Control and Signal Processing*. **7**(3), 183–211.

Lee, W.S., B.D.O. Anderson, R.L. Kosut and I.M.Y. Mareels (1993*b*). On robust performance improvement through the windsurfer approach to adaptive robust control. In *Proc. 32nd IEEE Conference on Decision and Control*. San Antonio, USA. pp. 2821–2826.

Liu, K. and R.E. Skelton (1990). Closed loop identification and iterative control design. In *Proc. 29th IEEE Conference on Decision and Control*. Honolulu, Hawaii. pp. 482–487.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Ljung, L. (1992). A discussion of model accuracy in system identification. *Int. Journal of Adaptive Control and Signal Processing*. **6**(3), 161–171.

Ljung, L. (1993*a*). Information content in identification data from closed-loop operation. In *Proc. 32nd IEEE Conference on Decision and Control*. San Antonio, USA. pp. 2248–2252.

Ljung, L. (1993*b*). Some results on identifying linear systems using frequency domain data. In *Proc. 32nd IEEE Conference on Decision and Control*. San Antonio, USA. pp. 3534–3538.

Ljung, L. and T. Glad (1994). *Modeling of Dynamic Systems*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Ljung, L., I. Gustavsson and T. Söderström (1974). Identification of linear multivariable systems operating under linear feedback control. *IEEE Trans. on Automatic Control*. **19**, 836–840.

Looze, P.L. and J.S. Freudenberg (1991). Limitations of feedback properties imposed by open loop right half plane poles. *IEEE Trans. on Automatic Control*. **36**, 736–739.

Luenberger, D.G. (1967). Canonical forms for linear multivariable systems. *IEEE Trans. on Automatic Control*. **12**, 290.

Luenberger, D.G. (1969). *Optimization by Vector Space Methods*. John Wiley & Sons, New York, USA.

Lunze, J. (1989). *Robust Multivariable Feedback Control*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Maciejowski, J.M. (1989). *Multivariable Feedback Design*. Addison-Wesley, Wokingham, England.

Mäkilä, P.M. and J.R. Partington (1992). Worst-case identification from closed loop time series. In *Proc. American Control Conference*. Chicago, IL, USA. pp. 301–306.

Mäkilä, P.M. and J.R. Partington (1995). On boundend-error identification of feedback systems. *Int. Journal of Adaptive Control and Signal Processing*. **9**(1), 47–61.

Mäkilä, P.M., J.R. Partington and T.K. Gustafson (1995). Worst-case control-relevant identification. *Automatica*. **31**, 1799–1819.

MatLab (1994). *Matlab User's Guide for the Optimization Toolbox, version 1.02*. The Mathworks Inc., South Natick, USA.

McFarlane, D. and K. Glover (1990). *Robust Controller Design using Normalized Coprime Factor Plant Descriptions*. Springer Verlag, Berlin, Germany.

McFarlane, D. and K. Glover (1992). A loop shaping design procedure using H_∞ synthesis. *IEEE Trans. on Automatic Control*. **37**, 759–769.

McKelvey, T. (1995). *Identification of State-Space Models from Time and Frequency Data*. PhD thesis. Linköping University, Linköping, Sweden.

McKelvey, T. (1996). State-space parametrizations of multivariable linear systems using tridiagonal matrix forms. In *Proc. 35th IEEE Conference on Decision and Control*. Kobe, Japan. pp. 6354–3659.

Meyer, D.G. (1990). The construction of special coprime factorizations in discrete time. *IEEE Trans. on Automatic Control*. **35**, 588–590.

Meyer, G.M. and G.F. Franklin (1987). A connection between normalized coprime factorizations and linear quadratic regulator theory. *IEEE Trans. on Automatic Control*. **32**, 227–228.

Middleton, R.H. (1991). Trade-offs in linear control system design. *Automatica*. **27**, 281–292.

Middleton, R.H. and G.C. Goodwin (1986). Improved finite word length characteristics in digital control using delta operators. *IEEE Trans. on Automatic Control*. **31**, 1015–1021.

Milanese, M. and A. Vicino (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica*. **27**, 997–1009.

Milanese, M. and R. Tempo (1985). Optimal algorithms theory for robust estimation and prediction. *IEEE Trans. on Automatic Control*. **30**, 730–738.

Mohtadi, C. (1990). Bode's integral theorem for discrete time systems. *IEE Proc. on Control Theory and Applications.* **137**(D), 57.

Nesterov, Y. and A. Nemirovsky (1994). *Interior-point polynomial methods in convex programming.* SIAM Studies in applied mathematics, Philadelphia, Pennsylvania, USA.

Nett, C.N. (1986). Algebraic aspects of linear control systems stability. *IEEE Trans. on Automatic Control.* **31**, 941–949.

Nett, C.N., C.A. Jacobson and M.J. Balas (1984). A connection between state-space and doubly coprime fractional representations. *IEEE Trans. on Automatic Control.* **29**, 831–832.

Ninness, B.M. (1993). *Stochastic and Deterministic Modelling.* PhD thesis. University of Newcastle, Newcastle, Australia.

Ninness, B.M. and G.C. Goodwin (1995*a*). Estimation of model quality. *Automatica.* **31**, 1771–1797.

Ninness, B.M. and G.C. Goodwin (1995*b*). Rapprochement between bounded-error and stochastic estimation theory. *Int. Journal of Adaptive Control and Signal Processing.* **9**(1), 107–132.

Ninness, B.M. and J.C. Gómez (1995). Mimo system identification using orthonormal basis functions. In *Proc. 35th IEEE Conference on Decision and Control.* New Orleans, USA. pp. 703–708.

Norton, J.P. (1986). *An Introduction to Identification.* Academic Press, London, England.

Nyquist, H. (1932). Regeneration theory. *Bell System Technical Journal.* **11**, 126–147.

Ober, R.J. (1987). Balanced realizations: canonical forms, parametrization, model reduction. *Int. Journal of Control.* **46**, 643–670.

Ober, R.J. (1991). Balanced parametrizations of classes of linear systems. *SIAM Journal on Control and Optimization.* **29**, 1251–1287.

Obinata, G. and J.B. Moore (1988). Characterization of controllers in simultaneous stabilization. *Systems & Control Letters.* **10**, 333–340.

Ogata, K. (1987). *Discrete-Time Control Systems.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Ogata, K. (1990). *Modern Control Engineering.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

O'Young, S.D. and B.A. Francis (1985). Sensitivity tradeoffs for multivariable systems. *IEEE Trans. on Automatic Control.* **30**, 625–632.

Packard, A. and J.C. Doyle (1993). The complex structured singular value. *Automatica.* **29**, 71–109.

Packard, A. and M. Helwig (1989). Relating the gap and graph metrics via the triangle inequality. *IEEE Trans. on Automatic Control.* **34**, 1296–1297.

Partanen, A.G. (1995). *Controller refinement with application to a sugar cane crushing mill.* PhD thesis. The Australian National University, Canberra, Australia. San Antonio, USA.

Partanen, A.G. and R.R. Bitmead (1993). Two stage iterative identification/controller design and direct experiment controller refinement. In *Proc. 32nd IEEE Conference on Decision and Control.* San Antonio, USA. pp. 2833–2838.

Partanen, A.G. and R.R. Bitmead (1995). The application of an iterative identification and controller design to a sugar cane crushing mill. *Automatica.* **31**, 1547–1563.

Partanen, A.G., R.D. Pierce and R.R. Bitmead (1994a). Crushing mill control for sugar cane - a robust nonadaptive LQG/LTR strategy. *Control Engineering Practice.* **2**, 3–16.

Partanen, A.G., Z. Zang, R.R. Bitmead and M. Gevers (1994b). Experimental restricted complexity controller design. In *Prepr. 10th IFAC Symp. on System Identification.* Vol. 1. Copenhagen, Denmark. pp. 177–182.

Partington, J.R. (1991). Robust identification and interpolation in H_∞. *Int. Journal of Control.* **54**, 1281–1290.

Pintelon, R., P. Guillaume, Y. Rolain, J. Schoukens and H. Van hamme (1994). Parametric identification of transfer functions in the frequency domain - a survey. *IEEE Trans. on Automatic Control.* **39**, 2245–2260.

Priestley, M.B. (1981). *Spectral Analysis and Time Series.* Academic Press, London, England.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley & Sons, New York, USA.

Rivera, D.E. and M. Morari (1987). Control-relevant model reduction problems for SISO H_2, H_∞ and μ-controller synthesis. *Int. Journal of Control.* **46**, 505–527.

Rivera, D.E. and S.V. Gaikwad (1992). Modeling for control design in combined feedback/feedforward control. *Proc. American Control Conference.* pp. 1445–1446.

Rivera, D.E., J.F. Pollard and C.E. Garcia (1992). Control-relevant prefiltering: a systematic design approach and case study. *IEEE Trans. on Automatic Control.* **37**, 964–974.

Rivera, D.E., X. Chen and D.S. Bayard (1993). Experimental design for robust process control using Schroeder-phased input signals. *Proc. American Control Conference.* pp. 895–899.

Saeks, R. and J. Murray (1982). Fractional representations, algebraic geometry, and the simultaneous stabilization problem. *IEEE Trans. on Automatic Control.* **27**, 895–903.

Savant, C.J. (1958). *Basic Feedback Control System Design*. McGraw-Hill, New York, USA.

Scheid, R.E., D.S. Bayard and Y. Yam (1991). A linear programming approach to characterizing norm bounded uncertainty from experimental data. In *Proc. American Control Conference*. Boston, USA. pp. 1956–1958.

Scherrer, W., M. Deistler, M. Kopel and W. Reitgruber (1991). Solution sets for linear dynamic errors-in-variables models. *Statistica Neerlandica*. **45**(4), 391–404.

Schrama, R.J.P. (1991). An open loop solution to the approximate closed loop identification problem. *Proc. 9th IFAC/IFORS Symp. on Identification and System Parameter Estimation*. pp. 1602–1607.

Schrama, R.J.P. (1992*a*). Accurate identification for control design: the necessity of an iterative scheme. *IEEE Trans. on Automatic Control*. **37**, 991–994.

Schrama, R.J.P. (1992*b*). *Approximate Identification and Control Design with Application to a Mechanical System*. PhD thesis. Delft University of Technology, Delft, the Netherlands.

Schrama, R.J.P. and O.H. Bosgra (1993). Adaptive performance enhancement by iterative identification and control design. *Int. Journal of Adaptive Control and Signal Processing*. **7**(5), 475–487.

Schrama, R.J.P. and P.M.J. Van den Hof (1992). An iterative scheme for identification and control design based on coprime factorizations. *Proc. American Control Conference*. pp. 2842–2846.

Sefton, J.A. and R.J. Ober (1993). On the gap metric and coprime factor perturbations. *Automatica*. **29**, 723–734.

Sefton, J.A., R.J. Ober and K. Glover (1990). Robust stabilization in the presence of coprime factor perturbations. In *Proc. 29th IEEE Conference on Decision and Control*. Honolulu, Hawaii. pp. 1197–1198.

Skelton, R.E. (1985). On the structure of modelling errors and the inseparability of the modelling and control problems. In *Proc. IFAC Workshop on Model Error Concepts and Compensation*. R.E. Skelton, D.H. Owens (Eds.). Boston, USA. pp. 13–20.

Skelton, R.E. (1989). Model error concepts in control design. *Int. Journal of Control*. **49**(5), 1725–1753.

Skogestad, S. and I. Postlethwaite (1996). *Multivariable Feedback Control*. John Wiley & Sons, Chichester, England.

Smith, M.C. (1989). On stabilization and the existence of coprime factorizations. *IEEE Trans. on Automatic Control*. **34**, 1005–1007.

Smith, R., G. Dullerud, S. Rangan and K. Poola (1997). Model validation for dynamically uncertain systems. *Mathematical Modelling of Systems*. **3**, 43–58.

Smith, R.S. (1995). Model validation for robust control: an experimental process

control application. *Automatica*. **31**, 1637–1648. (reviewed for Automatica, TS-072).

Smith, R.S. and J.C. Doyle (1989). Model invalidation: a connection between robust control and identification. In *Proc. American Control Conference*. Pittsburgh, USA. pp. 1435–1440.

Smith, R.S. and J.C. Doyle (1992). Model validation: a connection between robust control and identification. *IEEE Trans. on Automatic Control*. **37**, 942–952.

Söderström, T. and P. Stoica (1975). Identifiability conditions for linear systems operating in closed loop. *Int. Journal of Control*. **21**(2), 243–255.

Söderström, T. and P. Stoica (1981). Comparison of some instrumental variable methods - consistency and accuracy aspects. *Automatica*. **17**, 101–115.

Söderström, T. and P. Stoica (1983). *Instrumental Variable Methods for System Identification*. Springer Verlag, Berlin, Germany.

Söderström, T. and P. Stoica (1989). *System identification*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Söderström, T., L. Ljung and I. Gustavsson (1976). Identifiability conditions for linear multivariable systems operating under feedback. *IEEE Trans. on Automatic Control*. **21**, 837–840.

Steiglitz, K. (1974). *An Introduction to Discrete Systems*. John Wiley & Sons, New York, USA.

Stein, G. and M. Athans (1987). The LQG/LTR procedure for multivariable feedback control design. *IEEE Trans. on Automatic Control*. **32**, 105–114.

Stix, G. (1995). Trends in semiconductor manufacturing - toward "point one". *Scientific American*. **272**, 72–77.

Sworder, D.D. (1966). A study of the relationship between identification and optimization in adaptive control problems. *Journal of the Franklin Institute*. **281**, 198–213.

Tay, T.T., J.B. Moore and R. Horowitz (1989). Indirect adaptive techniques for fixed controller performance enhancement. *Int. Journal of Control*. **50**, 1941–1959.

Tempo, R. (1988). Robust estimation and filtering in the presence of bounded noise. *IEEE Trans. on Automatic Control*. **33**, 864–867.

Tsai, M.C., E.J.M. Geddes and I. Postlethwaite (1992). Pole-zero cancelations and closed loop properties of an H_∞ mixed sensitivity design problem. *Automatica*. **28**, 519–530.

van de Vegte, J. (1990). *Feedback Control Systems*. Prentice-Hall, New York, USA.

van den Boom, T.J.J. (1992). System identification for coprime factor plant description with bounded error. In *Proc. American Control Conference*. Chicago, IL, USA. pp. 1248–1252.

Van den Hof, P.M.J. (1989). *On residual-based parametrization and identification of*

multivariable systems. PhD thesis. Eindhoven University of Technology, Eindhoven, the Netherlands.

Van den Hof, P.M.J. (1992). System order and structure indices of linear systems in polynomial form. *Int. Journal of Control.* **55**(6), 1471–1490.

Van den Hof, P.M.J. and R.A. de Callafon (1996). Multivariable closed-loop identification: from indirect identification to dual-youla parametrization. In *Proc. 35th IEEE Conference on Decision and Control.* Kobe, Japan. pp. 1397–1402.

Van den Hof, P.M.J. and R.J.P. Schrama (1993). An indirect method for transfer function estimation from closed loop data. *Automatica.* **29**, 1523–1527.

Van den Hof, P.M.J. and R.J.P. Schrama (1995). Identification and control - closed loop issues. *Automatica.* **31**, 1751–1770.

Van den Hof, P.M.J., R.A. de Callafon and E.T. van Donkelaar (1997). CLOSID - a closed loop system identification toolbox for Matlab. In *Prepr. 11th IFAC Symposium on System Identification.* Kitakyushu, Japan. pp. 983–988.

Van den Hof, P.M.J., R.J.P. Schrama and P.M.M. Bongers (1994). On nominal models, model uncertainty and iterative methods in identification and control design. *Lecture Notes in Control and Information Sciences.* **192**, 39–50.

Van den Hof, P.M.J., R.J.P. Schrama, R.A. de Callafon and O.H. Bosgra (1995). Identification of normalised coprime plant factors from closed-loop experimental data. *European Journal of Control.* **1**(1), 62–74.

van Overbeek, A.J.M. and L. Ljung (1982). On-line structure selection for multivariable state space models. *Automatica.* **18**, 529–543.

van Overschee, P. (1995). *Subspace identification, Theory - Implementation - Application.* PhD thesis. Katholieke Universiteit Leuven, Leuven, Belgium.

Verhaegen, M. (1993). Application of a subspace model identification technique to identify LTI systems operating in closed-loop. *Automatica.* **29**, 1027–1040.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica.* **30**, 61–74.

Verma, M. and E. Jonckheere (1984). L_∞-compensation with mixed sensitivity as a broadband matching problem. *Systems & Control Letters.* **4**, 125–129.

Viberg, M. (1994). Subspace methods in system identification. In *Prepr. 10th IFAC Symp. on System Identification.* Vol. 1. Copenhagen, Denmark. pp. 1–12.

Vidyasagar, M. (1984). The graph metric for unstable plants and robustness estimates for feedback stability. *IEEE Trans. on Automatic Control.* **29**, 403–418.

Vidyasagar, M. (1985). *Control System Synthesis: A Factorization Approach.* MIT Press, Cambridge, Massachusetts, USA.

Vidyasagar, M. (1988). Normalized coprime factorizations for nonstrictly proper systems. *IEEE Trans. on Automatic Control.* **33**, 300–301.

Vidyasagar, M., H. Schneider and B.A. Francis (1982). Algebraic and topological aspects of feedback stabilization. *IEEE Trans. on Automatic Control.* pp. 880–894.

Voda, A. and I.D. Landau (1995). The autocalibration of PI controllers based on two frequency measurements. *Int. Journal of Adaptive Control and Signal Processing.* **9**, 395–421.

Wahlberg, B. (1991). System identification using Laguerre models. *IEEE Trans. on Automatic Control.* **36**, 551–562.

Wahlberg, B. (1994). System identification using Kautz models. *IEEE Trans. on Automatic Control.* **39**, 1276–1282.

Wahlberg, B. and L. Ljung (1986). Design variables for bias distribution in transfer function estimation. *IEEE Trans. on Automatic Control.* **31**, 134–144.

Wahlberg, B. and L. Ljung (1991). On estimation of transfer function error bounds. In *Proc. 1st European Control Conference.* Grenoble, France. pp. 1378–1383.

Wahlberg, B. and L. Ljung (1992). Hard frequency-domain error bounds from least-squares like identification techniques. *IEEE Trans. on Automatic Control.* **37**, 900–912.

Willems, J.C. (1971). Least squares stationary optimal control and the algebraic riccati equation. *IEEE Trans. on Automatic Control.* **16**, 621–634.

Willems, J.C. (1986). From time-series to linear systems - part II: Exact modelling. *Automatica.* **22**, 675–694.

Willems, J.C. (1987). From time-series to linear systems - part III: Approximate modelling. *Automatica.* **23**, 87–115.

Willems, J.C. (1991). Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. on Automatic Control.* **36**, 259–294.

Willems, J.C. (1992). Feedback in a behavioral setting. In *Systems, models and feedback: Theory and applications.* A. Isidori and T.J. Tarn (Eds.). Birkhäuser, Boston. pp. 179–191.

Willems, J.C. (1995). The behavioral approach to systems and control. *Trans. Japanese Society of Instrumentation and Control Engineers.* **34**, 603–612.

Wortelboer, P.M.R. (1993). *Frequency Weighted Balanced Reduction of Closed-Loop Mechanical Servo-Systems: Theory and Tools.* PhD thesis. Delft University of Technology, Delft, the Netherlands.

Youla, D.C., H.A. Jabr and J.J. Bongiorno (1976*a*). Modern wiener-hopf design of optimal controllers - part ii the multivariable case. *IEEE Trans. on Automatic Control.* **21**, 319–338.

Youla, D.C., J.J. Bongiorno and H.A. Jabr (1976*b*). Modern wiener-hopf design of optimal controllers - part I: The single-input-output case. *IEEE Trans. on Automatic Control.* **21**, 3–13.

Zames, G. (1963). Functional analysis applied to nonlinear feedback systems. *IEEE Trans. on Circuit Theory*. **10**, 392–404.

Zames, G. (1981). Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses. *IEEE Trans. on Automatic Control*. **26**, 310–320.

Zang, Z., R.R. Bitmead and M.R. Gevers (1992). Disturbance rejection: on-line refinement of controllers by closed loop modelling. *Proc. American Control Conference*. pp. 2829–2833.

Zang, Z., R.R. Bitmead and M.R. Gevers (1995). Iterative weighted least-squares identification and weighted LQG control design. *Automatica*. **31**, 1577–1594.

Zeiger, H.P. and J. McEwen (1974). Approximate linear realizations of given dimension via Ho's algorithm. *IEEE Trans. on Automatic Control*. **19**, 153.

Zheng, W.X. and C.B. Feng (1995). A bias correction method for indirect identification of closed loop systems. *Automatica*. **31**, 1019–1024.

Zhou, K., J.C. Doyle and K. Glover (1996). *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, New Jersey, USA.

Zhu, S.Q. (1989). Graph topology and gap topology for unstable systems. *IEEE Trans. on Automatic Control*. **34**, 848–855.

Zhu, Y.C. (1990). *Identification and Control of MIMO Industrial Processes: An Integration Approach*. PhD thesis. Eindhoven University of Technology, Eindhoven, the Netherlands.

Zhu, Y.C. and A.A. Stoorvogel (1992). Closed loop identification of coprime factors. *Proc. 31st IEEE Conference on Decision and Control*. pp. 453–455.

Zhu, Y.C. and T. Backx (1993). *Identification of Multivariable Industrial Processes for Simulation, Diagnosis and Control*. Springer Verlag, London.

Zhu, Y.C., M.H. Driessen, A.A.H. Damen and P. Eykhoff (1988). *A new scheme for identification and control*. EUT Report 88-E-213, Eindhoven Univ. of Technology, the Netherlands.

Ziegler, J.G. and N.B. Nichols (1942). Optimum settings for automatic controllers. *Trans. American Society Mechanical Eng.*. **64**, 759–768.

Index

Samenvatting

Terugkoppeling Gebaseerde Identificatie voor Verbeterde en Robuuste Regelingen

een fractionele aanpak toegepast op een wafer stage

R.A. de Callafon, Oktober 1998

Teruggekoppelde regelingen en systeemidentificatie zijn beide betrokken bij de regeling en prediktie of modelvorming van de dynamische aspecten van een systeem, en toch worden ze slechts zelden samengevoegd. Het gebruik van experimentele gegevens voor de modellering van de dynamica van een systeem is een krachtig gereedschap voor het aanleveren van modellen om regelaars te ontwerpen voor dat systeem. Echter, meer verfijnde en verbeterde robuuste regelaars kunnen ontwikkeld worden wanneer de systeemidentificatie en het ontwerp van de terugkoppelingsregeling simultaan uitgevoerd worden. Dit proefschrift draagt bij in de ontwikkeling van een dergelijke geïntegreerde aanpak van zowel de terugkoppelingsregeling als de systeemidentificatie, met de bedoeling om een systematische procedure te ontwikkelen voor het ontwerp van een geavanceerde en robuust presterende terugkoppelingsregeling voor een dynamisch systeem.

De nadruk van dit proefschrift ligt op het gebied van de systeemidentificatie en verschaft resultaten en hulpmiddelen voor een zogeheten terugkoppeling gebaseerde identificatie van systemen. Nieuwe resultaten voor de verbetering van de integratie van systeemidentificatie met robuust regelaarontwerp kunnen in dit proefschrift gevonden worden. De geïntegreerde aanpak wordt geïllustreerd en succesvol toegepast op een industriële, hoge nauwkeurigheid, multivariabel mechanisch positioneer systeem, bekend onder de naam wafer stage. Een dergelijke wafer stage wordt gebruikt in wafer steppers voor de fabricage van geïntegreerde circuits.

Om de gesloten-lus prestaties van het terugkoppelingsgeregelde systeem te volgen en te garanderen, wordt een model-gebaseerde procedure voorgedragen. De model-gebaseerde procedure omvat de schatting van een verzameling van modellen, opgebouwd via een nominaal model voorzien van een karakterisering van de modelonzekerheid. De geïdentificeerde verzameling van modellen wordt vervolgens gebruikt in een robuuste regelaarontwerp methodiek om een verbeterde en robuuste terugkop-

pelingsregeling te verkrijgen. Om de verbetering van de gesloten-lus prestaties te garanderen tijdens de uitvoering van de opeenvolgende stappen van modelverzamelingsschatting en robuust regelaarontwerp, zijn er gesloten-lus validatie testen geformuleerd voor zowel de terugkoppeling gebaseerde identificatie en het robuuste regelaarontwerp. De gesloten-lus validatie testen garanderen dat een bovengrens op de gesloten-lus prestaties monotonisch verbeterd kan worden.

Het gebruik van systeemidentificatie voor het vinden van modellen voor regelaarontwerp wordt gekenmerkt door de identificatie van een systeem dat opereert onder gesloten-lus geregelde condities. Bovendien, benaderende modellen of modellen met een lage complexiteit zijn nodig voor het opzetten van een hanteerbaar lage orde regelaarontwerpprobleem. Met dit doel voor ogen, bevat dit proefschrift een kritische evaluatie van zogeheten gesloten-lus benaderende identificatie technieken die gebruikt worden om het probleem van het vinden van benaderende modellen van een (mogelijk instabiel) systeem aan te pakken. Een fractionele model aanpak, waarbij het mogelijk instabiele systeem wordt gerepresenteerd en geïdentificeerd via stabiele coprieme factoren, wordt gemotiveerd en gebruikt om het gesloten-lus benaderende identificatie probleem op te lossen.

Met de fractionele model aanpak wordt het nominale model geïdentificeerd via de schatting van stabiele coprieme factoren. Frequentie domein technieken worden gebruikt om een dergelijke stabiele coprieme factorizatie van het systeem te schatten. De modelonzekerheid wordt gekarakteriseerd via het in beschouwing nemen van een perturbatie in een zogenoemde duale-Youla parametrizatie. De modelonzekerheid wordt geschat via beschikbare technieken voor probabilistische onzekerheidsbegrenzende identificatie. Het wordt aangetoond dat, met behulp van de fractionele model aanpak en de gekozen structuur van de verzameling van modellen, de benaderende en terugkoppeling gebaseerde identificatie van de verzameling van modellen volledig kan en moet worden gedaan, op basis van gesloten-lus experimenten. Daarnaast wordt aangetoond dat de voorgestelde structuur van de verzameling van modellen uitermate geschikt is voor het evalueren en bijhouden van de gesloten-lus prestaties van het teruggekoppelde geregelde systeem.

De voorgestelde model-gebaseerde procedure verschaft een methodiek om systeemidentificatie en robuust regelaarontwerp te integreren. Het levert een systematische procedure op voor het ontwerp van een verbeterde en robuuste terugkoppelingsregeling voor een mogelijk instabiel, multivariabel, dynamisch systeem. Elke opeenvolgende stap van terugkoppeling gebaseerde identificatie en robuust regelaarontwerp binnen de procedure kan gebruikt worden om de gesloten-lus prestaties van het dynamisch systeem stapsgewijs te verbeteren.

Curriculum Vitae

Raymond Arnoud de Callafon was born on March 11, 1968 in Rotterdam, the Netherlands.

1980–1986	Pre-university education (VWO) at O.S.G. J.C. de Glopper, Capelle a/d IJssel, the Netherlands. Majors: Dutch, English, Math A & B, Physics, Chemistry and Biology.
1986-1988	M.Sc. student with Electrical Engineering at Delft University of Technology.
1988-1992	M.Sc. student with Mechanical Engineering at Delft University of Technology. Graduated (cum laude) in the Mechanical Engineering Systems and Control Group at Delft University of Technology on the multivariable control-relevant identification of a Philips compact disc mechanism.
1992-1997	Ph.D. student with the Mechanical Engineering Systems and Control Group at Delft University of Technology, sponsored by the Dutch Systems and Control Theory Network. Application of developed techniques on a Philips compact disc mechanism and the motion control of a wafer stage in cooperation with Philips Research Laboratories.
1997-1998	Research assistant with the Structural Systems and Control Laboratory of Prof. R.E. Skelton at University of California, San Diego.
July 1998	(Acting) tenure-track assistant professor with the Dynamics and Control Group at the University of California, San Diego.

I made a big decision a little while ago.
I don't remember what it was, which prob'ly goes to show
That many times a simple choice can prove to be essential
Even though it often might appear inconsequential.

I must have been distracted when I left my home because
Left or right I'm sure I went. (I wonder which it was!)
Anyway, I never veered: I walked in that direction
Utterly absorbed, it seems, in quiet introspection.

For no reason I can think of, I've wandered far astray.
And that is how I got to where I find myself today.

Bill Watterson in *The Indispensable Calvin and Hobbes.*