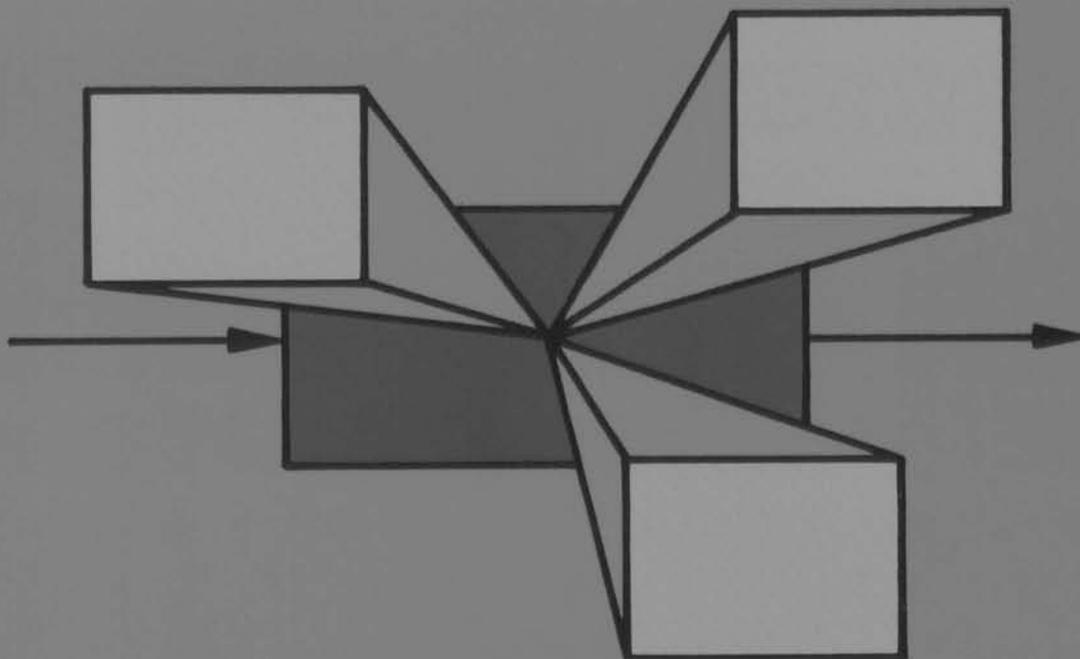


SELECTED
TOPICS
IN

Identification Modelling and Control

Volume 10, 1997

Edited by O.H. Bosgra and P.M.J. Van den Hof
and C.W. Scherer



608

Delft University Press

694145
3058935
9R 2995608

IDENTIFICATION, MODELLING AND CONTROL MODELLING AND CONTROL

Progress Report on Research Activities in the
Mechanical Engineering Systems and Control Group

Edited by O.H. Bosgra, P.M.J. Van den Hof and C.W. Scherer

Volume 10, December 1997



Mechanical Engineering Systems and Control Group
Delft University of Technology

Delft University Press 1997

Contents

Volume 10, December 1997

SELECTED TOPICS IN IDENTIFICATION, MODELLING AND CONTROL

Progress Report on Research Activities in the
Mechanical Engineering Systems and Control Group

Edited by O.H. Bosgra, P.M.J. Van den Hof and C.W. Scherer

Volume 10, December 1997

Mechanical Engineering Systems and Control Group
Delft University of Technology

Delft University Press/1997



SELECTED TOPICS IN IDENTIFICATION,
MODELLING AND CONTROL

Published and Distributed by

Delft University Press
Mekelweg 4
2628 CD Delft
The Netherlands
Tel.: (0)15 - 2783254
Telefax: (0)15 - 2781661

By order of

Mechanical Engineering Systems and Control Group
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
Tel.: +31-15-2786400; Telefax: +31-15-2784717
email: e.m.p.arkesteijn@wbmt.tudelft.nl
www: <http://www-sr.wbmt.tudelft.nl/sr/th>



CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Selected

Selected topics in identification, modelling and control:
progress report on research activities in the mechanical engineering
systems and control group. - Delft:Mechanical Engineering Systems and Control Group,
Delft University of Technology, Vol. 10-ed. by O.H. Bosgra and
P.M.J. Van den Hof and C.W. Scherer.-ill. Met lit.opg.
ISBN 90-407-1636-6
SISO 656 UDC 531.7 + 681.5 NUGI 841

Cover design by Ruud Schrama

© 1997 Copyright Delft University Press. All rights reserved. No part of this journal may be reproduced,
in any form or by any means, without written permission from the publisher.

Contents

Volume 10, December 1997

- Identification of a fluidized catalytic cracking unit: an orthonormal basis function approach
E.T. van Donkelaar, P.S.C. Heuberger and P.M.J. Van den Hof 1
- Closed-loop identification of uncertainty models for robust control design: a set membership approach
M. Milanese, M. Taragna and P.M.J. Van den Hof 9
- Mixed objectives MIMO control design for a Compact Disc player
M. Dettori, V. Prodanovic and C.W. Scherer 19
- Indirect position measurement and singularities in a Stewart platform with an application to model-based control
S.H. Koekebakker 27
- Suppressing non-periodically repeating disturbances in mechanical servo systems
R.L. Tousain, J.C. Boissy, M.L. Norg, M. Steinbuch and O.H. Bosgra 39
- A full block S-procedure with applications
C.W. Scherer 47
- An internal-model-based framework for the analysis and design of repetitive and learning controllers
D. de Roover and O.H. Bosgra 53
- Point-to-point control of a high accuracy positioning mechanism
D. de Roover and F.B. Sperling 61
- Low order control design by feedback relevant identification and closed-loop controller reduction
R.A. de Callafon and P.M.J. Van den Hof 69
- Doctoral dissertations published in 1997 76

Editorial

Since we started with the issuing of this series of research progress reports in April 1990, this is the tenth volume in the series, and we are happy to present this jubilee-edition to our colleagues, friends and contacts around the world. Again, this tenth volume shows a wide range of aspects in both fundamental and applied subjects in systems and control engineering.

Taking a look at the two main application areas of the research in our group, (mechanical) motion control systems and industrial/(petro)chemical production processes, the first branch of research seems to be dominantly present in this issue. Six papers discuss applications of control aspects in mechanical motion control systems, such as CD player, wafer stepper and Stewart platform. The balance between the two application areas will be intended to be restored in forthcoming issues.

Taking a look at 'newcoming' authors, we would like to welcome Mario Milanese and Michele Taragna from the Politecnico di Torino in Italy. The joint paper that is incorporated is the result of a research project supported by the European Community in the scope of the Human Capital and Mobility Program "SIMONET".

Marco Dettori is a Ph.D.-student working in a research project in cooperation with and supported by the Philips Research Laboratories in Eindhoven,

The Netherlands. The paper presented here involves also Vladimir Prodanovic who finished his M.Sc.-studies in the scope of this project.

Rob Tousain is a new Ph.D.-student whose contribution here reflects work that has been done as part of his M.Sc.-studies, performed at the Philips Research Laboratories in Limeil Brevannes, France. We acknowledge also here the contribution of his "industrial" supervisors, Jean-Christophe Boissy, Meindert Norg and Maarten Steinbuch.

The other authors have appeared -regularly- in previous issues of this magazine, and so we assume them to be known to our regular readership.

Additional information on the activities of our group, as well as reprint versions of the papers in this and previous volumes of our progress report, can be found on our WWW-site:

www-sr.wbmt.tudelft.nl/sr

Finally we would like to wish all our colleagues, friends and contacts a happy and prosperous 1998.

Okko Bosgra o.h.bosgra@wbmt.tudelft.nl

Paul Van den Hof p.m.j.vandenhof@wbmt.tudelft.nl

Carsten Scherer c.w.scherer@wbmt.tudelft.nl

Editors

Identification of a fluidized catalytic cracking unit: an orthonormal basis function approach

Edwin T. van Donkelaar[†], Peter S.C. Heuberger[‡] and Paul M.J. Van den Hof

Mechanical Engineering Systems and Control Group

Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

E-mail: e.t.vandonkelaar@wbmt.tudelft.nl

Abstract. Multivariable system identification of a model IV fluidized catalytic cracking unit is performed using a linear time invariant model parametrization based on orthonormal basis functions. This model structure is a linear regression structure which results in a simple convex optimization problem for least squares prediction error identification. Unknown initial conditions are estimated simultaneously with the system dynamics to account for the slow drift of the measured output from the given initial condition to a stationary working point. The model accuracy for low frequencies is improved by a steady-state constraint on the estimated model and incorporation of prior knowledge of the large time constants in the model structure.

The model accuracy is furthermore improved by an iteration over identification of a high order model and model reduction. First a high order model is estimated using an orthonormal basis. This model is reduced and used to generate a new orthonormal basis which is used in the following iteration step for high order estimation.

With the approach followed accurate models are estimated with only a limited amount of data.

Keywords. System identification, orthonormal basis functions, multivariable systems.

1 Introduction

The fluidized catalytic cracking (FCC) process is used to crack crude oil into lighter and more valuable components. The overall economic performance of a refinery largely depends on the economic operation of the FCC unit (Tatrai *et al.*, 1994). Therefore accurate modelling and control of this process is of large importance.

In this report multivariable system identification of a Model IV fluidized catalytic cracking unit is described. The nonlinear simulation model described in McFarlane *et al.* (1993) is used as the process to be identified. The system is multivariable with

large interaction between the several input/output-channels. Characteristic for this system is the combination of fast and slow physical phenomena. Both frequency ranges need to be estimated accurately for high performance control design. This means, however, that long data sequences at a high sampling rate need to be used to capture both slow and fast phenomena in the data.

Also the working point in which open-loop identification is performed is generally not be a stationary point. This causes the measured variables to drift from the working point to the nearest stationary point. These drifts can cause a problem for prediction error identification as these methods assume the signals to be quasi stationary.

To deal with the large dynamic range of the system and the transients in the measured output, an approach is applied which utilizes system-based orthonormal basis functions (Heuberger *et al.* (1995),

[†]The research of Edwin van Donkelaar and Peter Heuberger is supported by the Dutch Technology Foundation (STW) under contract DWT55.3618

[‡]Peter Heuberger is on partial leave from the Dutch National Institute of Public Health and the Environment (RIVM)

Van den Hof *et al.* (1995), Ninness and Gustafson (1994), Ninness *et al.* (1995)). In this approach system poles are chosen on the basis of prior knowledge or prior identification results. With these poles a complete orthonormal basis for stable dynamical systems is generated. The model is parametrized in terms of these basis functions, resulting in a model structure which is linear in the parameters. A least squares identification criterion is used to obtain optimal parameter values which can be calculated efficiently using linear regression techniques.

Initial conditions are estimated simultaneously with the system dynamics without losing the linear regression structure. This is to account for slow drifts of the measured output, which is present because measurements are taken in a nonstationary working point. This improves the estimation of the system dynamics.

To improve the static behaviour of the estimated model, the static gain is fixed. Fixing the static gain of the model amounts to a linear constraint on the parameters. This constraint can be incorporated as a hard constraint or as a soft constraint. In both cases linear regression techniques can be used to calculate the optimal parameter efficiently.

The estimation is further improved by iterating over high order identification with orthonormal basis functions and model reduction. The reduced order model is used to generate a basis for the high order identification in the next iteration step.

The outline of this report is as follows. First, in section 2 the process under consideration is discussed. Next, in section 3 both the preliminary experiments and the experiments for parametric identification are described. In section 4 the parametric identification procedure is described and also the validation results are given. Section 5 concludes this report.

2 The process

The system to be identified is the nonlinear FCCU model described in McFarlane *et al.* (1993). In figure 1 a flow sheet is given of a typical Model IV fluidized catalytic cracking unit is shown. The system consists basically of two subsystems: the riser or reactor and the regenerator.

In the reactor fresh feed of crude oil and hot catalyst coming from the regenerator is mixed which induces the cracking reaction which makes the crude oil to fall down into lighter and more valuable components. These components leave the reactor at the top as gas and are separated in the downstream separators. In the reaction the catalyst is contaminated with carbonaceous material (coke). The spent catalyst is transported to the regenerator to be regenerated.

In the regenerator spent catalyst is regenerated by means of air injection provided by the air blowers (figure 1). The air injection fluidizes the catalyst and removes the coke by an exothermic reaction. The heat induced by this reaction is used to supply the heat for the endothermic reaction in the reactor.

Hence, no additional heat is supplied to the reactor. Because of this, the reactor and the regenerator are highly coupled. The multivariable system shows large interaction between the several input-output channels.

The system also shows both fast and slow dynamic behaviour. The fast behaviour comes from flow and pressure phenomena while the slow behaviour stems from the fact that it takes a long time to reach a thermal equilibrium.

The inputs that can be manipulated for identification purposes are given by

$$u(t) = [F_3(t) T_2(t) F_9(t) p_4(t) \Delta p(t)]^T$$

where the $F_3(t)$ is the fresh feed flow, $T_2(t)$ is the temperature of the fresh feed flow, $F_9(t)$ is the lift air flow, $p_4(t)$ is the pressure in the riser and $\Delta p(t)$ is the pressure difference between regenerator and riser. The measured output vector is given by

$$y(t) = [l_{sp}(t) T_{reg}(t) T_r O_{2sg}(t) V_{11}(t)]^T$$

where $l_{sp}(t)$ is level in the stand pipe of the riser, $T_{reg}(t)$ is the temperature in the regenerator, $T_r(t)$ is the temperature in the reactor, $O_{2sg}(t)$ is the concentration oxygen in the stack gas coming out of the regenerator and $V_{11}(t)$ is the valve position at the point where the wet gas is sent to the main separators. The number of inputs and outputs are denoted as n_u and n_y respectively.

The disturbances acting on the system are the following. A measurable disturbance is the ambient temperature $T_{atm}(t)$ and a disturbance that is not measurable is the changing coking factor $\psi_F(t)$ of the incoming fresh feed. The minimum sample time is $\Delta T = 10$ sec., which is the sample time of the measurement devices.

3 Experiments

First preliminary experiments are performed to assess disturbance dynamics, assessment of linearity of the system and to obtain rough system knowledge. The preliminary experiments that are performed are freeruns and step response experiments.

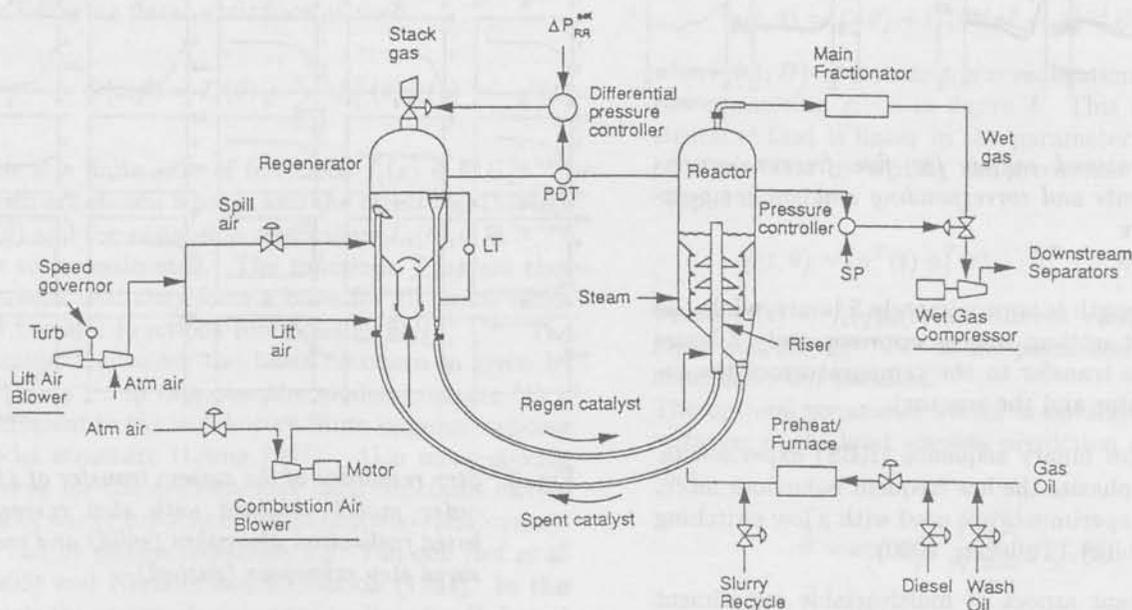


Fig. 1: Schematic of a Model IV fluid catalytic cracking unit (McFarlane et al., 1993)

3.1 Freerun experiments

First several freerun experiments are conducted to assess disturbance dynamics. In figure 2 the outputs of 5 freeruns are given. From these experiments the following observations were made:

- an initial condition disturbance is present. The disturbance is approximately equal for all freeruns that are performed, only a variation of dynamics due to the ambient temperature is observed. This nonlinearity is not accounted for in the identification because rather small experiment lengths are used in parametric identification such that the variation in ambient temperature is limited.
- substantial changes of the coking factor occur once every 5-7 hours which has a large influence on the measured output. The coking factor is a disturbance which is not measurable. If parametric identification is performed on data which is disturbed by a changing coking factor, a considerable bias can be expected. For this reason only the first part of the data will be used in parametric identification.

It is important for the identification approach to account for the transients in the data. Especially the transient in the temperature in the regenerator and the reactor are severe.

3.2 Step response experiments

Step response experiments are performed to assess nonlinearity and obtain a first indication of system dynamics. This knowledge is needed to choose an appropriate sampling time and experiment length. The inputs are successively excited with a step function and the five outputs are measured. The experiments are performed with the amplitudes: $u_{amp} = [2.4 \ 12.5 \ 0.90 \ 0.11 \ 0.10]$, $2u_{amp}$, $-u_{amp}$ and $-2u_{amp}$.

The measured step responses are detrended for the initial condition disturbance with the mean of several freeruns. By comparison of the results with different step sizes, it can be concluded that the system behaves fairly linearly apart from possible activation of valve constraints. It becomes clear that the system has very fast phenomena, therefore decimation is not possible. In figure 3 the measured step responses are given.

3.3 Experiments for parametric identification

The following experiments for parametric identification are performed.

- Pseudo random binary sequence (PRBS) experiments (Ljung, 1987). With this input signal the high frequent behaviour of the system is dominantly present in the data because the

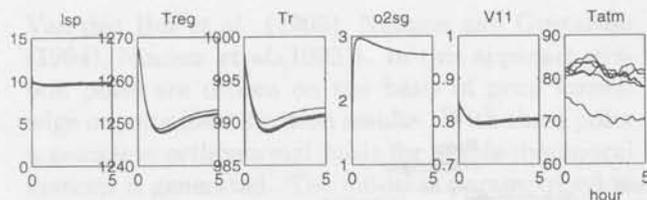


Fig. 2: Measured output for five freerun experiments and corresponding ambient temperature

data length is approximately 5 hours, while the slowest settling time is approximately 2 hours (in the transfer to the temperature of the regenerator and the reactor).

- Random binary sequence (RBS) experiments. To emphasize the low frequent behaviour more, RBS experiments are used with a low switching probability (Tulleken, 1990).

An important aspect of multivariable experiment design is that the inputs are as much uncorrelated as possible to keep the identification problem well conditioned. If different realizations of the signals mentioned above are used for the different input channels, this is approximately satisfied.

4 Parametric identification and validation

The aim of the identification approach is to identify a model which accurately describes all the data that is present: the response to the PRBS signal which contains the high frequent behaviour more than the low frequent, the response to the RBS signal which emphasizes the low frequent behaviour more and the step response data with a major emphasis on low frequent dynamics.

The approach followed here involves basically three steps:

1. a realization algorithm based on step response data is used to obtain a rough parametric model of the system.
2. an orthonormal basis function model is identified using the parametric model obtained in the first step to generate an initial basis. The model is iteratively improved.
3. The previous steps are performed for five multi-input/single-output (MISO) problems. In the last step a full multivariable model is estimated with a basis generated by the identification results of the previous step.

These steps are describe in the sequel of this section.

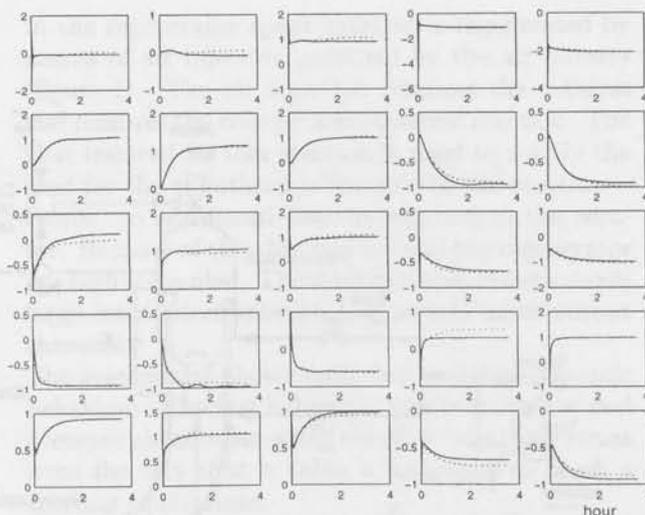


Fig. 3: Step responses of the system transfer of 41th order model obtained with step response based realization algorithm (solid) and measured step responses (dashed).

4.1 Step response realization algorithm

First the realization algorithm described in Van Helmont *et al.* (1990) is used to obtain a state-space description directly from the step response coefficients. The algorithm is similar to the algorithm of Kung (1978) but does not act on the Hankel matrix with pulse response coefficients but with step response coefficients. This has the advantage that no discrete differencing has to be applied to the step response data to obtain impulse response coefficients, which increases the influence of disturbances. The emphasis of the obtained models is more on the low frequent behaviour than with the algorithm of Kung. The identification of the MIMO model is split into 5 separate MISO identification problems. The reason for this is that the identification problem becomes computationally more tractable. Also the input and output weighting and compensation for time delays can be performed on each transfer function separately. This flexibility is necessary to obtain accurate models.

In figure 3 the resulting model is given. The order of the estimated models are: from u to $y_1(t)$ 10th order, 6th to y_2 , 9th to y_3 , 10th to y_4 and 6th order to y_5 . This makes a 41st order MIMO model.

The MISO realization models describe the step response data accurately. However, the models are not capable of predicting the output of the PRBS and RBS data well.

4.2 ORTFIR identification

In identification with orthonormal basis functions the following parametrization is used

$$G(z, \theta) = D(\theta) + \sum_{i=1}^n L_i^T(\theta) f_i(z) \quad (1)$$

This is a finite sum of functions $f_i(z) \in \mathbb{R}H_2^{n_b \times n_u}$ which are chosen a priori and the direct feedthrough $D(\theta)$ and the expansion coefficients $L_i(\theta) \in \mathbb{R}^{n_y \times n_b}$ are to be estimated. The functions $f_i(z)$ are chosen such that they form a basis for all stable rational transfer functions in $\mathbb{R}H_2^{n_u \times n_y}$. The simplest choice for the basis functions is given by $f_i(z) = z^{-i}$. In this case the model structure (1) is equivalent to the well known finite impulse response model structure (Ljung, 1987). Also more specific choices for the orthonormal basis functions can be made, where prior knowledge of the system dynamics can be incorporated; see e.g. Van den Hof *et al.* (1995) and Ninness and Gustafson (1994). In this article the approach presented in Van den Hof *et al.* (1995) will be followed.

In Van den Hof *et al.* (1995) orthonormal basis functions are generated using prior knowledge of the system in terms of rough pole locations or an identified model, of which only the state space matrices $\{A, C\}$ or $\{A, B\}$ are used. From this prior knowledge an inner system $G_b(z)$ is constructed with balanced state space realization $\{A_b, B_b, C_b, D_b\}$. Now, an orthonormal basis is constructed as follows

$$f_i(z) = (zI - A_b)^{-1} B_b G_b^{i-1}(z), \quad i = 1, 2, \dots \quad (2)$$

With this choice, the parametrization (1) coincides with the series connection of filters given in 4. Here $x_i(t)$ denotes the balanced state of the filter.

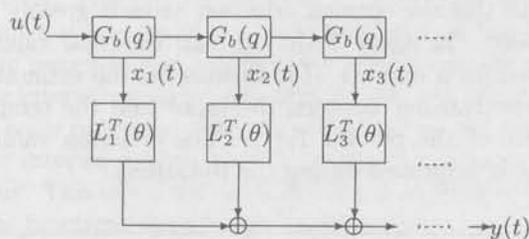


Fig. 4: Model parametrization with generalized orthogonal basis functions

From (2) it can be seen that if the $\{A_b, B_b\}$ is chosen correctly, only the state space matrices $\{C, D\}$ need to be estimated. Hence, if the prior knowledge of the system dynamics is accurate, only a limited number of coefficients needs to be estimated. This results in models with limited bias and variance.

The output prediction with this model structure can be conveniently expressed with

$$y(t, \theta) = D(\theta) + C(\theta)(zI - A)^{-1} B u(t)$$

where $\{A, B\}$ is a state-space realization of the series connection given in figure 4. This is a model structure that is linear in the parameter. This can be made clear by writing the prediction of a single output as

$$\hat{y}(t, \theta) = [u^T(t) \tilde{u}_1^T(t) \cdots \tilde{u}_n^T(t-n)] \theta$$

where $\tilde{u}_i(t) = f_i(q)u(t)$ are filtered versions of the input and $\theta \in \mathbb{R}^{n_b \times n_y}$ is the parameter that is to be identified from the data.

The optimal parameter vector is obtained by minimization of the least squares prediction error criterion

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N \varepsilon^2(t, \theta)$$

with the prediction error defined by $\varepsilon(t, \theta) = y(t) - \hat{y}(t, \theta)$. The optimal parameter estimate is equal to the least-squares optimal solution of the overdetermined set of equations $Y = \phi \theta$, where $Y^T = [y^T(1) \cdots y^T(N)]$ and the rows of ϕ are given by $[u^T(t) \tilde{u}_1^T(t) \cdots \tilde{u}_n^T(t-n)]$. The analytic solution of this optimization problem is given by $\hat{\theta} = (\phi^T \phi)^{-1} \phi^T Y$.

Hence, because the model structure is linear in the parameter, the optimal parameter vector is unique and can be calculated analytically.

Estimation of initial conditions

In the measured data of the FCCU a transient is present due to an initial condition that is not a stationary working point. To account for this, the initial condition is estimated simultaneously with the system dynamics. This can be done without losing the linear regression structure as follows. The model structure is extended to

$$\hat{y}(t, \theta) = (D(\theta) + C(qI - A)^{-1} B(\theta)) u(t) + C A^{t-1} x_0 \quad (3)$$

where $\{A, C\}^1$ are a priori chosen state-space matrices and $D(\theta), B(\theta)$ and x_0 are the parameters that are to be estimated from the data. This boils down to solving the least-squares optimal parameter vector for the overdetermined set of equations

$$Y = [\phi \ \phi_{x_0}] \begin{bmatrix} \theta \\ x_0 \end{bmatrix} := \phi_{ext} \theta_{ext}$$

¹Note, that the $\{A, C\}$ is used as prior information rather than $\{A, B\}$ without losing the linear regression structure.

where the rows of ϕ_{x0} are given by CA^{t-1} . Estimation of initial conditions can be used to reduce the bias due to unknown initial conditions at the expense of an increased variance.

The estimated transient of the initial condition and the measured output for the reactor temperature are given in figure 5. The transient due to the nonstationary initial condition is fitted accurately.

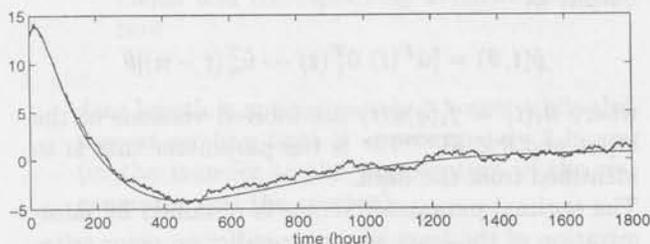


Fig. 5: Measured data $T_r(t)$ with RBS experiment (dotted) and the estimated initial condition contribution (solid).

Enforcement of the static gain

The low frequent and static behaviour of the system is barely present in the RBS data due to the relatively short data length compared to the slowest time constant. Therefore the static gain of the estimated models can be inaccurate. To remedy this, the static gain is enforced on the model by means of a constraint, that is linear in the parameter vector. Therefore the linear regression structure is preserved. The static gain of the model 1 is given by

$$K_{ss}(\theta) = D(\theta) + C(I - A)^{-1}B(\theta) := Q\theta \quad (4)$$

Any static gain K_{ss} can be enforced on the estimated model by using the Lagrangian of the constrained optimization problem. This boils down to solving

$$\begin{bmatrix} \phi_{ext}^T \phi_{ext} & (Q \ 0)^T \\ (Q \ 0) & 0 \end{bmatrix} \begin{bmatrix} \theta_{ext} \\ \lambda \end{bmatrix} = \begin{bmatrix} \theta_{ext}^T Y \\ K_{ss}^T \end{bmatrix}$$

where λ is the Lagrange multiplier. This is uniquely solvable because the matrix on the left hand side is square and invertible.

The constraint is enforced on the model such that the steady-state gain is equal to the specified one. However, the steady-state gain taken from the step response data is not accurate; therefore possibly unnatural behaviour is enforced on the model. To alleviate this, *soft* constraints are used, which are constraints that can be violated. A soft constraint can be implemented by adding one or more equations of the type (4) to the overdetermined set of equations that has to be solved for the unconstrained problem.

Figure 6 shows the measured step responses, together with the step response of the model resulting from applying no static state constraint as well as from using a soft constraint. The model with the soft constraint fits the measured step response well, while the model with no constraint has a considerable deviation in the steady state gain.

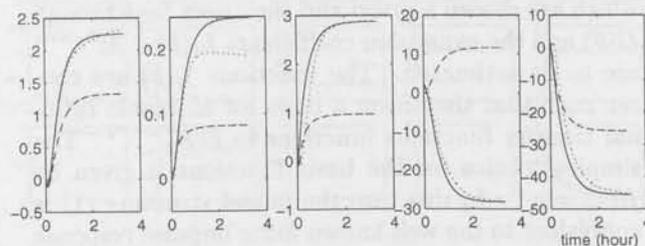


Fig. 6: Measured step response for the input $u(t)$ and the output $T_r(t)$ (dotted). Step responses of the estimated models with no static gain constraint (dashed), and a with a soft constraint (solid).

Iterative model enhancement

For further improvement of the model, an iterative scheme of ORTFIR identification and balanced model reduction (Moore, 1981) is applied. In this iteration the following steps are applied:

Step 1. generate basis functions,

Step 2. estimate a high order model with the ORTFIR model structure (1),

Step 3. reduce the high order model with e.g. balanced reduction, and use the reduced order model to generate a basis in the first step.

With this the optimal criterion value is greatly improved. In figure 7 the optimal criterion value is given for a number of iterations for the estimation of the transfer between the input and the temperature of the reactor $T_r(t)$. The criterion value is clearly improved during the iterations.

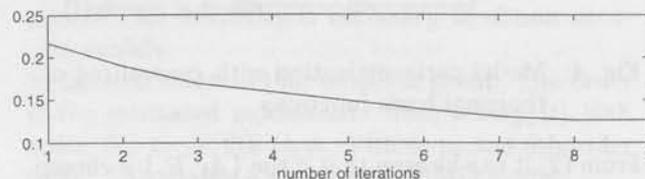


Fig. 7: Optimal value of the criterion function for 9 iterations (low model order: 6, high model order: 20) for the estimation of the transfer between $u(t)$ and $T_r(t)$

The high order is chosen such that all dynamical phenomena are incorporated in the model. This can be assessed by inspection of the estimated expansion coefficients $L_i(\theta)$. Equivalent to the impulse response coefficients, these coefficients go to zero for stable systems for high enough model order n (Van den Hof *et al.*, 1995). Therefore estimated expansion coefficients are denoted as the generalized impulse response coefficients. An example is given in figure 8.



Fig. 8: Generalized impulse responses of the estimated model with soft constraints for the transfer from $u(t)$ to $T_7(t)$.

The aim of the iteration is to concentrate the energy of the estimated model in the first few expansion coefficients such that a low order model can be derived. This trend is indeed observed during the iterations but can in general not be guaranteed.

Conditioning of the optimization problem

To calculate the optimal parameter vector the Toeplitz matrix $\phi^T \phi$ needs to be invertible. This implies that ϕ must have full column rank to obtain an estimate without numerical problems. There are several reasons why this may not be the case.

First of all, this can occur if the dynamics present in the basis functions is slow compared to the data length. In that case the inputs filtered with the basis functions forming the columns of ϕ may not be independent. This can be detected by inspection of the impulse responses of the basis functions. If the impulse response has considerable energy outside the time interval given by the data length, the orthogonal basis functions are not orthogonal on the finite time interval and numerical problems are likely to occur. This effect will be denoted the shifting of the basis functions outside the data window.

Secondly, if a high number of repetitions of the basis dynamics n is used the energy in the impulse response of the basis functions shifts to later time instants. This makes the basis functions to shift out of the data window, resulting in a badly conditioned optimization problem.

To avoid numerical problems, the following strategy is followed. Because slow dynamics is present in the basis functions, the number of repetitions is restricted to $n = 1$. Fast dynamics is added to the

dynamics of the basis functions to add extra flexibility in the model. The added dynamics can be any set of stable poles. In the identification of the FCCU, poles are added in the origin.

The part of the regression matrix regarding the initial conditions ϕ_{x0} has another character than the regression matrix for the dynamic part ϕ : the columns of the first consists of transient responses and the columns of the second of responses to signal with mean value zero. Due to this difference the number of repetitions n that can be used in ϕ_{x0} is larger than that can be used in ϕ before bad conditioning occurs. Therefore, the number of repetitions of the basis functions in ϕ_{x0} is taken to be $n = 3$ to give the estimation of the transient extra flexibility.

4.3 The full MIMO model

The five identified MISO models, of order 5, 7, 7, 6 and 8 respectively, are combined into one MIMO model. The dynamics of this model is used to generate a basis for the full MIMO system. To avoid numerical problems, only one group is used to parametrize the model, i.e. $n=1$.

The optimal value of the identification cost function could be improved from $V_{opt} = 0.11$ for the combination of the five MISO models to $V_{opt} = 0.099$ for the MIMO model. For this an RBS data set is used as identification data and a PRBS data set is used for validation. Similar results are observed if an RBS data set, other than the identification data, is used as validation set.

The step responses of this model are given in figure 10 together with the measured step responses.

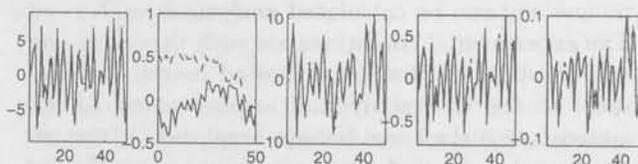


Fig. 9: Measured output (solid) and output prediction (dashed) for the 33rd order MIMO model.

The output prediction of the MIMO model is given in figure 9. This is based on a PRBS validation set. The output prediction of the second output seems inaccurate, however this is mainly due to the initial condition disturbance in the validation set. This is only accounted for by the mean of five freeruns which is rather inaccurate. The other outputs are predicted accurately.

As conclusion, the identified model with the described approach is consistent with the step response data, the RBS data set and the PRBS data set.

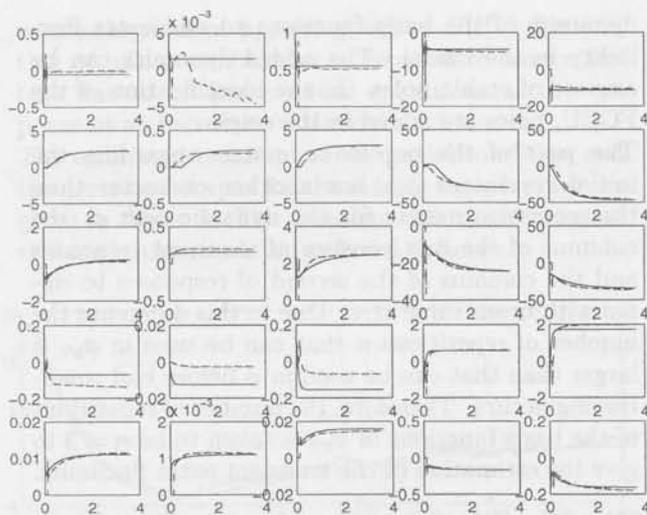


Fig. 10: Step responses of the 33th order multivariable model (solid), (dashed) and the measured step response (dotted).

5 Conclusions

In this paper the identification of a nonlinear simulation model for the Model IV catalytic cracking unit is described. The model structure is based on orthonormal basis functions where the basis functions are chosen using prior knowledge of the system dynamics obtained from identification based on the step response data. This results in a linear regression model structure. To obtain an optimal parameter estimate, a least squares identification criterion is used. Therefore the optimal parameter vector is unique and can be calculated analytically.

The experimental conditions are such that data sets can be obtained that have limited length with respect to the slowest dynamical phenomena of the system. Also the time domain amplitude of the input signal is limited due to possible activation of system constraints.

To account for a slow drift of the measured data due to an initial condition which is not a stationary working point, initial conditions are estimated simultaneously with the system dynamics. The static and low frequent behaviour of the model is hardly present in the data due to the limited data length. To accuracy of the model in this frequency range hard, soft or mixed steady-state constraints are incorporated in the identification procedure. This can be implemented while preserving the linear regression structure.

The resulting model is consistent with both the step response data and the input-output data. Hence, both fast and slow dynamics are estimated accurately. This is obtained with only a limited amount

of data by making fruitful use of prior knowledge of the system.

References

- Golub, G. and C.I. Van Loan (1983). *Matrix Computations*. John Hopkins University Press, Baltimore.
- 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. Autom. Control*, AC-40, 451-465.
- Kung, S. (1978). A new identification and model reduction algorithm via singular value decompositions. *Proc. 12th Asilomar Conf. Circuits Syst. and Computers*, Pacific Grove, CA, Nov. 6-8, 1978, pp. 705-714.
- Ljung, L. (1987). *System Identification - Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.
- McFarlane, R.C., R.C. Reineman, J.F. Bartee and C. Georgakis (1993). Dynamic simulator for a model IV fluid catalytic cracking unit. *Computers Chem. Engineering*, 17, 275-300.
- Moore, B.C. (1981). Principal component analysis in linear systems: controllability, observability and model reduction. *IEEE Trans. Autom. Control*, AC-26, 17-32.
- Ninness, B.M. and F. Gustafson (1994). *A Unifying Construction of Orthonormal Bases for System Identification*. Technical report EE9433, Department of Electrical Engineering, University of Newcastle, Australia.
- Ninness, B., J. Gomez en S. Weller (1995). MIMO systems identification using orthonormal basis functions. *Proc. 34th IEEE Conf. Decision and Control*, New Orleans, pp. 703-708.
- Tatrai, F.Z., P.A. Lant, P.L. Lee, I.T. Cameron and R.B. Newell (1994). Model reduction for regulatory control: an FCCU case study. *Trans. IChemE*, 72, Part A, 402-407.
- Tulleken, H.J.A.F. (1990). Generalized binary noise test-signal concept for improved identification experiment design. *Automatica*, 26, 37-49.
- Van den Hof, P.M.J., P.S.C. Heuberger and J. Bokor (1995). System identification with generalized orthonormal basis functions. *Automatica*, 31, 1821-1834.
- Van Helmont, J.B., A.J.J. Van der Weiden and H. Anneveld (1990). Design of optimal controller for a coal fired Benson boiler based on a modified approximate realization algorithm. In R. Whalley (Ed.), *Application of Multivariable System Techniques (AMST 90)*. Elsevier Publ. Comp., London, pp. 313-320.

Closed-loop identification of uncertainty models for robust control design: a set membership approach †

Mario Milanese[§], Michele Taragna[§] and Paul Van den Hof[#]

[§] Dipartimento di Automatica e Informatica, Politecnico di Torino; Corso Duca degli Abruzzi 24, I-10129 Torino, Italy; E-mail: milanese@polito.it, taragna@polito.it

[#] Mechanical Engineering Systems and Control Group; Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands; E-mail: p.m.j.vandenhof@wbmt.tudelft.nl

Abstract. The paper considers the problem of identifying uncertainty model sets, defined by an approximated model of the plant to be identified and a frequency domain bound on the modeling error. It is supposed that the measurements consist of time domain samples, collected in closed loop operations and corrupted by a power bounded noise. The model is supposed to be used for robust control design, whose performance is measured by a given closed loop H_∞ norm, and the "goodness" of the model is measured by the discrepancy between the closed loop performance predicted by the model and the one actually achieved on the plant. It is shown that identifying a model minimizing this discrepancy is equivalent to finding the best approximated model of the dual Youla parametrization of the plant in a suitably weighted H_∞ norm. Then, an optimal uncertainty model is derived for the dual Youla parametrized plant, from which an uncertainty model for the actual plant is obtained. Such uncertainty model is finally used for designing a robust controller and evaluating the closed loop performance that can be guaranteed when the designed controller is applied to the actual plant.

Keywords. System identification, closed-loop identification, set-membership identification, uncertainty models, identification for control; robust control.

1 Introduction

In the past few years, a growing attention has been devoted to set membership methodologies for system identification (see e.g. Milanese et al. (1989), Kurzhanski and Veliov (1994), Smith and Dahleh (1994), Milanese et al. (1996)), largely motivated by the important progress in robust control design realized in the 80's.

Robust control methodologies aim to design controllers guaranteeing to meet the specifications not

for a single nominal model, but for all models obtained by given perturbations of the nominal model. Such model set, called *uncertainty model*, is introduced to take into account that models derived by any identification method are always affected by uncertainty. A quite popular class of uncertainty models is obtained by considering dynamic perturbations, bounded in the frequency domain. The simplest case is the additive uncertainty model \mathcal{M} defined as the set:

$$\mathcal{M}(M, W_M) = \{M(z) + \Delta(z) : \|W_M^{-1}(z)\Delta(z)\|_\infty < 1\} \quad (1)$$

where $M(z)$ is the transfer function of the nominal model, $\Delta(z)$ is the transfer function of the perturbation and $W_M(z)$ is a known transfer function. A large body of literature is available for designing robust controllers for such uncertainty models. How-

†The original version of this paper was presented at the 36th IEEE Conf. Decision and Control, 10-12 December 1997, San Diego, CA, USA. Copyright of this paper remains with IEEE. This research was supported in part by funds of Ministero dell'Università e della Ricerca Scientifica e Tecnologica, of CNR Project "Algorithms for Identification and Robust Control of Uncertain Systems" and EC Project "SIMONET".

ever, in most practical applications, such models are not directly available to the control designer and have to be identified from actual measurements on the unknown process P_0 to be controlled and from available prior information (or assumptions) on P_0 and on the noise corrupting the measurements.

Since the final goal is to guarantee high performance of the controlled plant, it is of relevance to provide an uncertainty model able to achieve this requirement. In Canale et al. (1998) and Canale et al. (1996) it is shown how to derive tight uncertainty models and evaluate the performance that can be guaranteed in closed loop on the true plant P_0 , using open loop experiments. The methods used in those papers need that the plant P_0 to be identified is asymptotically stable.

In this paper, a method is proposed to achieve the same goals using closed loop experiments, thus allowing to identify uncertainty models for unstable plants. An approach is followed that is closely related to the one in Van den Hof et al. (1996). Here, in particular, the focus is on deriving tight uncertainty models for the case of measurements corrupted by power bounded noise.

Another interesting feature, shared with few others papers (Van den Hof et al., 1996; Hakvoort and Van den Hof, 1995; De Callafon and Van den Hof, 1997), is that the uncertainty models are tuned to the closed loop measure of performance that is underlying the control design.

2 Problem formulation

As a general set-up, the linear time-invariant feedback interconnection of Fig. 1 is considered, where u and y are the measurable input and output of the plant, r_1 and r_2 are reference signals and e is a disturbance signal.

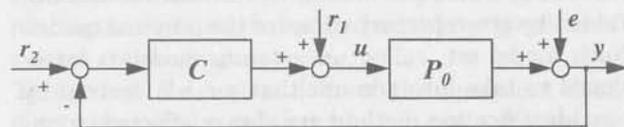


Fig. 1: Feedback configuration.

A performance function of a closed loop configuration composed of plant P_0 and controller C is a system property, such as a sensitivity function, a complementary sensitivity function, etc. This control performance function can be formalized as an element $J(P_0, C)$ in some normed (Banach) space. The control performance cost is then measured by the norm $\|J(P_0, C)\|$, and a corresponding control design method will provide a controller that minimizes this cost. Many control design methods are

based on the minimization of a particular performance cost. In the paper, the following ones are considered in detail (Van den Hof and Schrama, 1995):

- *Mixed sensitivity optimization.* The mixed sensitivity design is reflected by the choice

$$J(P_0, C) = \begin{bmatrix} V_1 (I + P_0 C)^{-1} \\ V_2 P_0 C (I + P_0 C)^{-1} \end{bmatrix} \in \mathcal{RH}_\infty^{2 \times 1} \quad (2)$$

with weighting functions $V_1, V_2 \in \mathcal{RH}_\infty$, and the corresponding control performance cost is $\|J(P_0, C)\|_\infty$. In the sequel, \mathcal{RH}_∞ denotes the set of real rational stable transfer functions.

- *H_∞ design based on robustness optimization.* This control design scheme proposed in McFarlane and Glover (1990) is reflected by the choice

$$J(P_0, C) = \begin{bmatrix} P_0 \\ I \end{bmatrix} (I + C P_0)^{-1} [C \ I] \in \mathcal{RH}_\infty^{2 \times 2} \quad (3)$$

and the corresponding control performance cost is $\|J(P_0, C)\|_\infty$.

For given model M and controller C designed on the basis of M , it holds that:

$$\begin{aligned} | \|J(M, C)\|_\infty - \|J(P_0, C) - J(M, C)\|_\infty | &\leq \\ &\leq \|J(P_0, C)\|_\infty \leq \\ &\leq \|J(M, C)\|_\infty + \|J(P_0, C) - J(M, C)\|_\infty. \end{aligned} \quad (4)$$

The following terms can be distinguished:

$\|J(P_0, C)\|_\infty$ is the *achieved performance* when the compensator C is applied to the true plant P_0 ;

$\|J(M, C)\|_\infty$ is the *designed performance* when the compensator C is applied to the identified model M ;

$\|J(P_0, C) - J(M, C)\|_\infty$ is the *performance degradation*, due to the fact that C has been designed from M rather than from P_0 .

One aims at minimization of the upper bound of the performance cost in (4). However, this simultaneous optimization over both M and C is intractable by common identification and control design techniques, because they can optimize either the model or the controller, each while the other element is fixed. This has led to the introduction of several iterative schemes making use of separate stages of identification and control design, see e.g. (Van den Hof and Schrama, 1995; Zang et al., 1995; Bitmead et al., 1997) and the references therein. In the identification stage of the i -th iteration, a new model M_i

is obtained by minimizing the performance degradation $\|J(P_0, C_{i-1}) - J(M, C_{i-1})\|_\infty$, where C_{i-1} is the controller designed in the previous iteration. In the control design stage, a new controller C_i is designed by minimizing the designed performance $\|J(M_i, C)\|_\infty$.

Indeed, a major motivation for iteration is due to the fact that a caution factor is introduced in control design based on model M_i only. This factor is used in order to prevent that the designed performance is high while the achieved performance may be poor and even the closed loop stability may be not achieved. The caution factor is progressively reduced as iterations go on and, hopefully, modeling error decreases.

In order to have a more systematic approach to deal with modeling errors, in this paper a method is proposed to derive, from measured data and suitable prior information, not only a model \hat{M} but also a tight bounding function $W_{\hat{M}}$ on the modeling error $\Delta = P_0 - \hat{M}$. In this way, an uncertainty model $\mathcal{M}(\hat{M}, W_{\hat{M}})$ is obtained of the form (1) guaranteeing that $P_0 \in \mathcal{M}(\hat{M}, W_{\hat{M}})$. Such uncertainty model is suitable to be used by robust control techniques, giving a controller with guaranteed achieved performance.

3 Dual Youla parametrization approach

A closed loop identification approach is adopted, based on the (dual) Youla parametrization of all plants that are stabilized by a given known controller (Van den Hof and Schrama, 1995). Given the feedback configuration in Fig. 2, it can be shown that, for given C stabilizing P_0 in closed loop, the unique value of R_0 that corresponds to the real plant P_0 is determined by

$$R_0 = D_c^{-1} (I + P_0 C)^{-1} (P_0 - P_x) D_x. \quad (5)$$

In the scheme, C has right coprime factorization (rcf) $C = N_c D_c^{-1}$, P_x is any auxiliary system stabilized by C with rcf $P_x = N_x D_x^{-1}$ and

$$S = D_c^{-1} (I + P_0 C)^{-1} \quad (6)$$

Now, defining the signals v , x as indicated in Fig. 2 and writing the node equations $x = D_x^{-1} (u + N_c v)$, $y = N_x x + D_c v$ and $u = r_1 + C r_2 - C y$, it follows that:

$$\begin{aligned} v &= (D_c + P_x N_c)^{-1} (y - P_x u) = \\ &= (D_c + P_x N_c)^{-1} [y - P_x (r_1 + C r_2 - C y)] \quad (7) \end{aligned}$$

$$\begin{aligned} x &= (D_x + C N_x)^{-1} (u + C y) = \\ &= (D_x + C N_x)^{-1} (r_1 + C r_2) \quad (8) \end{aligned}$$

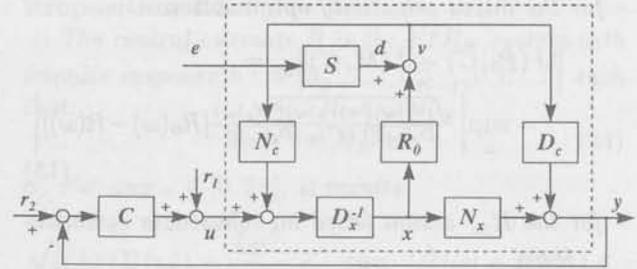


Fig. 2: Dual Youla representation of the data generating system.

and

$$v = R_0 x + S e. \quad (9)$$

Signals v and x can be reconstructed from closed loop data through known filters, provided the controller C is known. Thus, the identification of R_0 forms an open loop identification problem based on reconstructed measurements v , x .

The next propositions show that identifying a model \hat{M} of P_0 minimizing $\|J(P_0, C) - J(\hat{M}, C)\|_\infty$ is equivalent to finding an approximated model \hat{R} of R_0 minimizing a suitably weighted H_∞ norm.

Proposition 3.1 For given M , let

$$R = D_c^{-1} (I + M C)^{-1} (M - P_x) D_x. \quad (10)$$

Then,

- for the mixed sensitivity optimization:

$$\begin{aligned} J(P_0, C) - J(M, C) &= \\ &= \begin{bmatrix} -V_1 \\ V_2 \end{bmatrix} D_c (R_0 - R) D_x^{-1} (I + C P_x)^{-1} C \quad (11) \end{aligned}$$

- for the H_∞ design based on robustness optimization:

$$\begin{aligned} J(P_0, C) - J(M, C) &= \\ &= \begin{bmatrix} I \\ -C \end{bmatrix} D_c (R_0 - R) D_x^{-1} (I + C P_x)^{-1} [C \ I] \quad (12) \end{aligned}$$

Proof: See Appendix. \square

In the sequel of the paper, SISO case is considered and the identification criteria particularize according to the following proposition, where any term of type $F(\omega)$ stands for $F(j\omega)$ or $F(e^{j\omega})$ in continuous or discrete time domain respectively.

Proposition 3.2 (SISO case) For given M , let R as in (10). Then,

- for the mixed sensitivity optimization:

$$\begin{aligned} & \|J(P_0, C) - J(M, C)\|_\infty = \\ & = \sup_\omega \left| \frac{\sqrt{|V_1(\omega)|^2 + |V_2(\omega)|^2} N_c(\omega)}{D_x(\omega)[1+C(\omega)P_x(\omega)]} [R_0(\omega) - R(\omega)] \right| \end{aligned} \quad (13)$$

- for the H_∞ design based on robustness optimization:

$$\begin{aligned} & \|J(P_0, C) - J(M, C)\|_\infty = \\ & = \sup_\omega \left| \frac{D_c(\omega)[1+|C(\omega)|^2]}{D_x(\omega)[1+C(\omega)P_x(\omega)]} [R_0(\omega) - R(\omega)] \right|. \end{aligned} \quad (14)$$

Proof: The H_∞ norm of a stable transfer matrix G is defined as:

$$\|G\|_\infty = \sup_\omega \bar{\sigma}(G(\omega)) \quad (15)$$

where $\bar{\sigma}$ is the greatest singular value of G , i.e. the positive square root of the maximum eigenvalue of $G^H G$.

For the mixed sensitivity optimization, in the SISO case it results that, from (11):

$$J(P_0, C) - J(M, C) = \begin{bmatrix} -V_1 \\ V_2 \end{bmatrix} \frac{N_c}{D_x(1+CP_x)} (R_0 - R)$$

and result (13) directly follows from definition (15), since for a rank-one matrix the H_∞ norm and the Frobenius norm coincide.

Analogously for the H_∞ design based on robustness optimization, in the SISO case it results, from (12):

$$J(P_0, C) - J(M, C) = \begin{bmatrix} C & 1 \\ -C^2 & -C \end{bmatrix} \frac{D_c}{D_x(1+CP_x)} (R_0 - R).$$

Since the singular values of $\begin{bmatrix} C & 1 \\ -C^2 & -C \end{bmatrix}$ are $\sigma_1 = 0$ and $\sigma_2 = 1 + |C|^2$, result (14) follows from definition (15). \square

If the compensator C is stable, a valid choice for the auxiliary system P_x stabilized by C is $N_x = 0$ and $D_x = 1$, and in this case $N_c = C$ and $D_c = 1$ can be chosen as rcf of the compensator. Then the R.H.S. of (13) and (14) can be rewritten as $\|R_0 - R\|_\infty^{W_C} = \sup_\omega W_C^{-1}(\omega) |R_0(\omega) - R(\omega)|$, where

- for the mixed sensitivity optimization:

$$W_C^{-1}(\omega) = \sqrt{|V_1(\omega)|^2 + |V_2(\omega)|^2} |C(\omega)| \quad (16)$$

- for the H_∞ design based on robustness optimization:

$$W_C^{-1}(\omega) = 1 + |C(\omega)|^2. \quad (17)$$

Note that this particular choice of coprime factors leads to the situation that $R_0 = \frac{P_0}{1+P_0C}$, and identification of R_0 actually reduces to the indirect method of closed loop identification.

4 Set membership identification

In the previous section it has been shown that the identification of P_0 is equivalent to the open loop identification of R_0 , which is stable since it is supposed that C stabilizes the closed loop system. The SM approach developed in recent years for robust identification of SISO stable systems from open loop data can be used to identify an uncertainty model $\mathcal{R}(\hat{R}, W_{\hat{R}})$ of R_0 defined as the set:

$$\mathcal{R}(\hat{R}, W_{\hat{R}}) = \left\{ \hat{R} + \Delta : |\Delta(\omega)| \leq W_{\hat{R}}(\omega), \forall \omega \right\}. \quad (18)$$

Then, an uncertainty model $\mathcal{M}(\hat{M}, W_{\hat{M}})$ of the plant P_0 can be determined on the basis of $\mathcal{R}(\hat{R}, W_{\hat{R}})$.

Methods for identifying uncertainty models have been developed for various specific cases, according to the type of experimental information (e.g. time or frequency domain data) and the noise assumptions, see e.g. Milanese et al. (1996), Van den Hof and Schrama (1995) and Fiorio et al. (1997) and the references therein. Here the case of identification of SISO, linear time-invariant, discrete-time systems using time domain data corrupted by power bounded noise is worked out in some detail. It is supposed that R_0 is a causal, BIBO stable, SISO, linear time-invariant, discrete-time system with impulse response $h^{R_0} = \{h_0^{R_0}, h_1^{R_0}, \dots\}$, that controller C is known and stable, that known sequences r_1, r_2 are applied and that N output samples y_0, \dots, y_{N-1} are measured.

In view of (9), the experimental measurements give the following information on the impulse response h^{R_0} of R_0 :

$$v_\ell = \sum_{k=0}^{\ell} h_k^{R_0} x_{\ell-k} + d_\ell, \quad \text{for } \ell = 0, \dots, N-1 \quad (19)$$

where v_ℓ and x_ℓ , for $\ell = 0, \dots, N-1$, are known, derived from measurements $r_{1\ell}, r_{2\ell}$ and y_ℓ , for $\ell = 0, \dots, N-1$, through (7)-(8), and $d_\ell = \sum_{k=0}^{\ell} h_k^S e_{\ell-k}$. For the sake of simplicity, zero initial conditions are considered, but extension to nonzero case is easy. The noise sequence e is supposed unknown but power bounded, that is:

$$e^N = [e_0 \dots e_{N-1}]^T \in \mathcal{B}_e = \left\{ e^N \in \mathbb{R}^N : \frac{1}{\sqrt{N}} \|e^N\|_2 \leq \varepsilon \right\}. \quad (20)$$

Then sequence d is power bounded, since system S is stable and $\frac{1}{\sqrt{N}} \|d^N\|_2 \leq \sup_\omega |S(\omega)| \frac{1}{\sqrt{N}} \|e^N\|_2 \leq$

$\sup_{\omega} |D_c^{-1}(\omega) [1 + P_0(\omega) C(\omega)]^{-1}| \varepsilon$. The transfer function $F(z) = [1 + P_0(z) C(z)]^{-1}$ is not known. However, $F(z)$ is the transfer function from r_2 to $f = r_2 - y$, N samples of which are known from measured data. From these samples some estimate $\hat{F}(z)$ can be derived and $\delta = \sup_{\omega} |D_c^{-1}(\omega) \hat{F}(\omega)| \varepsilon$ can be used as an estimate of the power bound on d .

Now the aim is to derive an uncertainty model for R_0 , consisting of a nominal model and a measure of its modeling error. From equation (19) it follows that the experimental measurements give information on $h_k^{R_0}$ only for $k < N$. Thus, from measurements only it is not possible to derive a finite bound on modeling error. To this end, some prior information on R_0 is needed. To make a minimal use of prior information, a *residual* type is assumed, i.e. giving constraints on the tail of h^{R_0} only. In particular it is assumed that $R_0 \in K_R$ where:

$$K_R = \{R : |h_k^R| \leq L\rho^k, \forall k \geq N\} \quad (21)$$

with known $L \geq 0$ and $0 < \rho < 1$. For a discussion of such type of prior information, see e.g. Giarré et al. (1997).

The Feasible Systems Set, i.e. the set of all systems consistent with prior information and available measurements, is then given by:

$$FSS = \{h^R \in K_R : \frac{1}{\sqrt{N}} \|v^N - X_N T_N h^R\|_2 \leq \delta\} \quad (22)$$

where $v^N = [v_0 \dots v_{N-1}]^T$, T_N is the truncation operator defined as $T_N h^R = [h_0^R \dots h_{N-1}^R]^T$ and X_N is the lower triangular $N \times N$ Toeplitz matrix formed by the sequence $x^N = [x_0 \dots x_{N-1}]^T$.

The *FSS* is the smallest set of systems that, on the basis of assumed prior information and available measurements, is guaranteed to include R_0 , thus representing the "best" possible uncertainty model for R_0 . However, this set is not in a suitable form to be used by available robust design techniques. Then, the smallest uncertainty model of the form (18) is looked for, such that $FSS \subseteq \mathcal{R}(\hat{R}, W_{\hat{R}})$. This is obtained by computing \hat{R} as a central estimate, i.e. the center of the minimal ball in the $\|\cdot\|_{\infty}^{W_C}$ norm including *FSS* with radius:

$$r = \sup_{R \in FSS} \|\hat{R} - R\|_{\infty}^{W_C} = \inf_{\hat{R}} \sup_{R \in FSS} \|\hat{R} - R\|_{\infty}^{W_C}. \quad (23)$$

The quantity r is called (local) radius of information in SM identification literature and represents the minimal error that can be guaranteed on the basis of the given prior information and measurements. For this reason, \hat{R} is called (locally) optimal estimate of R_0 .

Then the bounding function $W_{\hat{R}}(\omega)$ is obtained by evaluating $\sup_{R \in FSS} |\hat{R}(\omega) - R(\omega)|$. The next proposition provides the solution to this problem.

Proposition 4.1

i) The central estimate \hat{R} is the *FIR* _{N} system with impulse response $h^{\hat{R}} = [h_0^{\hat{R}}, \dots, h_{N-1}^{\hat{R}}, 0, 0, \dots]$ such that:

$$T_N h^{\hat{R}} = X_N^{-1} v^N \quad (24)$$

ii) For any $\omega \in [0, 2\pi]$, it results:

$$\begin{aligned} \sqrt{N} \delta \bar{\sigma}(\Sigma(\omega)) - \frac{L\rho^N}{1-\rho} &\leq \sup_{R \in FSS} |\hat{R}(\omega) - R(\omega)| \leq \\ &\leq \sqrt{N} \delta \bar{\sigma}(\Sigma(\omega)) + \frac{L\rho^N}{1-\rho} \end{aligned} \quad (25)$$

where $\bar{\sigma}(\Sigma(\omega))$ is the maximal singular value of $\Sigma(\omega) = \Omega_N(\omega) X_N^{-1}$, $\Omega_N(\omega) = [\Re(\Psi_N(\omega))^T \Im(\Psi_N(\omega))^T]^T$ and $\Psi_N(\omega) = [1 e^{-j\omega} e^{-j2\omega} \dots e^{-j(N-1)\omega}]$. \square

Proof: From definition (22) of *FSS*, it follows that $(I - T_N)FSS = (I - T_N)K_R$ and $T_N FSS = \{h^N \in \mathbb{R}^N : (h^N - X_N^{-1} v^N)^T X_N^T X_N (h^N - X_N^{-1} v^N) \leq N\delta^2\}$ which is an ellipsoid with center in $X_N^{-1} v^N$. Then it follows that \hat{R} is a center of symmetry of *FSS*. Thus, i) follows from the well known result that a center of symmetry of a set is its Chebiceff center in any norm, see e.g. Kacwicz et al. (1986).

Let $R^N(\omega)$ be the z -transform of $T_N h^R$. Then:

$$\begin{aligned} \sup_{R \in FSS} |\hat{R}(\omega) - R^N(\omega)| - \frac{L\rho^N}{1-\rho} &\leq \sup_{R \in FSS} |\hat{R}(\omega) - R(\omega)| \leq \\ &\leq \sup_{R \in FSS} |\hat{R}(\omega) - R^N(\omega)| + \frac{L\rho^N}{1-\rho}. \end{aligned}$$

Since $\hat{R}(\omega) - R^N(\omega) = \Psi_N(\omega) T_N (h^{\hat{R}} - h^R)$ it results $|\hat{R}(\omega) - R^N(\omega)| = \|\Omega_N(\omega) T_N (h^{\hat{R}} - h^R)\|_2$. Then:

$$\begin{aligned} \sup_{R \in FSS} |\hat{R}(\omega) - R^N(\omega)| &= \\ &= \sup_{h^R : \|v^N - X_N T_N h^R\|_2 \leq \sqrt{N}\delta} \|\Omega_N(\omega) (X_N^{-1} v^N - T_N h^R)\|_2 = \\ &= \sup_{d^N : \|d^N\|_2 \leq \sqrt{N}\delta} \|\Omega_N(\omega) X_N^{-1} d^N\|_2. \end{aligned} \quad (26)$$

The R.H.S. of (26) is $\sqrt{N}\delta$ times the induced ℓ_2 norm of matrix $\Sigma(\omega) = \Omega_N(\omega) X_N^{-1}$, which is well known to be $\bar{\sigma}(\Sigma(\omega))$, thus proving ii). \square

An uncertainty model $\mathcal{R}(\hat{R}, W_{\hat{R}})$ can be obtained by taking \hat{R} as given by i) of proposition 4.1 and

$$W_{\hat{R}}(\omega) = \sqrt{N} \delta \bar{\sigma}(\Sigma(\omega)) + \frac{L\rho^N}{1-\rho}. \quad (27)$$

Note that L and ρ represent some information about the "memory" of the closed loop system. If the duration of the experiment is not shorter than the "memory" of the closed loop system, as needed for obtaining acceptable identification errors, then

the term $\frac{L\rho^N}{1-\rho}$ is typically negligible with respect to $\sqrt{N}\delta\bar{\sigma}(\Sigma(\omega))$. If this is the case, the derived uncertainty model is close to be the smallest uncertainty model of the form (18) guaranteed to include R_0 . Given the uncertainty model $\mathcal{R}(\hat{R}, W_{\hat{R}})$ of R_0 , the corresponding uncertainty model $\mathcal{M}(\hat{M}, W_{\hat{M}})$ of the plant P_0 is then given by the following proposition.

Proposition 4.2 *If C is stable then, with the choice $N_x = 0$, $D_x = 1$, $N_c = C$ and $D_c = 1$:*

$$R_0 \in \mathcal{R}(\hat{R}, W_{\hat{R}}) \Leftrightarrow P_0 \in \mathcal{M}(\hat{M}, W_{\hat{M}}) \quad (28)$$

where

$$\hat{M}(\omega) = \frac{1}{C} \left(\frac{(1 - C\hat{R})^*}{|1 - C\hat{R}|^2 - |C|^2 W_{\hat{R}}^2} - 1 \right) \quad (29)$$

$$W_{\hat{M}}(\omega) = \frac{W_{\hat{R}}}{|1 - C\hat{R}|^2 - |C|^2 W_{\hat{R}}^2} \quad (30)$$

Proof: See Appendix. \square

Note that $R_0 = \frac{P_0}{1 + P_0 C}$, but $\hat{R} \neq \frac{\hat{M}}{1 + \hat{M} C}$.

Making use of such uncertainty model, a new compensator can be designed, using robust design methods. For example, H_∞ design techniques allow one to compute a controller $C_{\mathcal{M}}$ such that

$$C_{\mathcal{M}} = \arg \min_{C \in \mathcal{C}_{rs}} \|J(\hat{M}, C)\|_\infty \quad (31)$$

where \mathcal{C}_{rs} is the set of all controllers guaranteeing robust stability with respect to any system in the uncertainty model $\mathcal{M}(\hat{M}, W_{\hat{M}})$.

Standard H_∞ design techniques require that model \hat{M} and model perturbation bound $W_{\hat{M}}$ have rational transfer functions. Then, $W_{\hat{R}}$ has to be chosen as a rational transfer function overbounding (27), by using e.g. the method in Scheid et al. (1991). Its order has to be kept low because it affects the order of \hat{M} and of $W_{\hat{M}}$, which in turn affects the order of $C_{\mathcal{M}}$. Indeed, even if the order of $W_{\hat{R}}$ is kept low, the order of \hat{M} is large, greater than N , since \hat{R} has transfer function of order N . If a low order model is desired, order reduction techniques can be used to derive from \hat{R} an approximated model \hat{R}_n of order $n < N$. In particular, the closed loop approximation method (Ceton et al., 1993) may be appropriate here. Estimate \hat{R}_n is no more optimal, giving the identification error:

$$E(\hat{R}_n) = \sup_{R \in FSS} \|\hat{R}_n - R\|_\infty^{W_C} = \alpha r \quad (32)$$

where $\alpha > 1$ measures the degradation in the identification error with respect to the radius of information, which is the minimal guaranteed error. Straightforward computation gives:

$$\alpha \leq 1 + \frac{\|\hat{R}_n - \hat{R}\|_\infty^{W_C}}{\sqrt{N}\delta \sup_\omega W_C^{-1}(\omega) \bar{\sigma}(\Sigma(\omega)) - \frac{L\rho^N}{1-\rho}} \quad (33)$$

As $n \rightarrow N$, \hat{R}_n is close to be optimal, i.e. $\alpha \rightarrow 1$. Indeed, typically it results that yet for moderate values of n , $\|\hat{R}_n - \hat{R}\|_\infty^{W_C}$ is small with respect to r and then $\alpha \approx 1$.

In order to derive an uncertainty model of the form $\mathcal{R}(\hat{R}_n, W_{\hat{R}_n})$, a bound on $|\hat{R}_n(\omega) - R_0(\omega)|$ is needed. The following result directly follows from proposition 4.1.

Proposition 4.3 *For any $\omega \in [0, 2\pi]$, it results:*

$$\begin{aligned} \sqrt{N}\delta\bar{\sigma}(\Sigma(\omega)) - \frac{L\rho^N}{1-\rho} - |\hat{R}_n(\omega) - \hat{R}(\omega)| &\leq \\ &\leq \sup_{R \in FSS} |\hat{R}_n(\omega) - R(\omega)| \leq \\ &\leq \sqrt{N}\delta\bar{\sigma}(\Sigma(\omega)) + \frac{L\rho^N}{1-\rho} + |\hat{R}_n(\omega) - \hat{R}(\omega)|. \end{aligned} \quad (34)$$

Typically, the above bounds are sufficiently tight for practical purposes. If needed, tighter bounds can be derived by use of theorem 2 of Giarré et al. (1997). By use of proposition 4.2, a "reduced order" uncertainty model $\mathcal{M}(\hat{M}_n, W_{\hat{M}_n})$ for P_0 can be derived from the "reduced order" uncertainty model $\mathcal{R}(\hat{R}_n, W_{\hat{R}_n})$ for R_0 , where $W_{\hat{R}_n}(\omega)$ is a rational transfer function overbounding the R.H.S. of (34). A "reduced order" robust controller can be derived using in (31) the uncertainty model $\mathcal{M}(\hat{M}_n, W_{\hat{M}_n})$. Since it may be convenient to choose n so that $\|\hat{R}_n - \hat{R}\|_\infty^{W_C}$ is sufficiently small to ensure that $\alpha \approx 1$ and the bounds of proposition 4.3 are reasonably tight, the complexity of the obtained controller may be not as low as desirable. Then, order reduction techniques can be used to derive a controller $C_{\mathcal{M}}^r$ of further reduced order. The performance degradation due to the use of the reduced order controller instead of the full order one $C_{\mathcal{M}}$ can be evaluated by considering the *robust performance* $\bar{J}(C_{\mathcal{M}}^r)$ achievable by $C_{\mathcal{M}}^r$ defined as:

$$\bar{J}(C_{\mathcal{M}}^r) = \sup_{M \in \mathcal{M}(\hat{M}, W_{\hat{M}})} \|J(M, C_{\mathcal{M}}^r)\|_\infty \quad (35)$$

Robust performance $\bar{J}(C_{\mathcal{M}}^r)$ is the minimal performance that it can be guaranteed, using the available information, when controller $C_{\mathcal{M}}^r$ is applied to the unknown plant P_0 . A method for the computation of $\bar{J}(C_{\mathcal{M}}^r)$ is proposed in (De Callafon and Van den Hof, 1997), requiring a sequence of μ -tests which may be computationally demanding. The following proposition gives bounds on $\bar{J}(C_{\mathcal{M}}^r)$ that can be easily computed. Alternative bounds can be found in Van den Hof et al. (1996).

Proposition 4.4

$$\begin{aligned} \|J(\hat{M}, C_{\mathcal{M}}^r)\|_{\infty} - \sup_{\omega} W_C^{-1}(\omega) \left[\sqrt{N} \delta \bar{\sigma}(\Sigma(\omega)) + \frac{L \rho^N}{1-\rho} \right] &\leq \\ &\leq \bar{J}(C_{\mathcal{M}}^r) \leq \\ \|J(\hat{M}, C_{\mathcal{M}}^r)\|_{\infty} + \sup_{\omega} W_C^{-1}(\omega) \left[\sqrt{N} \delta \bar{\sigma}(\Sigma(\omega)) + \frac{L \rho^N}{1-\rho} \right] &\quad (36) \end{aligned}$$

Proof: From propositions 3.2 and 4.2 and from (27), the next inequalities directly follow:

$$\begin{aligned} \bar{J}(C_{\mathcal{M}}^r) &= \sup_{M \in \mathcal{M}(\hat{M}, W_{\hat{M}})} \|J(M, C_{\mathcal{M}}^r)\|_{\infty} \leq \\ &\leq \|J(\hat{M}, C_{\mathcal{M}}^r)\|_{\infty} + \sup_{M \in \mathcal{M}(\hat{M}, W_{\hat{M}})} \|J(\hat{M}, C_{\mathcal{M}}^r) - J(M, C_{\mathcal{M}}^r)\|_{\infty} \\ &= \|J(\hat{M}, C_{\mathcal{M}}^r)\|_{\infty} + \sup_{R \in \mathcal{R}(\hat{R}, W_{\hat{R}})} \sup_{\omega} W_C^{-1}(\omega) |\hat{R}(\omega) - R(\omega)| \\ &= \|J(\hat{M}, C_{\mathcal{M}}^r)\|_{\infty} + \sup_{\omega} W_C^{-1}(\omega) \sup_{R \in \mathcal{R}(\hat{R}, W_{\hat{R}})} |\hat{R}(\omega) - R(\omega)| \\ &\leq \|J(\hat{M}, C_{\mathcal{M}}^r)\|_{\infty} + \sup_{\omega} W_C^{-1}(\omega) \left[\sqrt{N} \delta \bar{\sigma}(\Sigma(\omega)) + \frac{L \rho^N}{1-\rho} \right]. \end{aligned}$$

Since analogous inequalities hold for the lower bound of $\bar{J}(C_{\mathcal{M}}^r)$, the claim (36) is proved. \square

Appendix

Proof of Proposition 3.1.

For any matrix A , B and I (identity matrix) of compatible dimensions, the following equalities hold:

$$AB(I+AB)^{-1} = (I+AB)^{-1}AB \quad (A.1)$$

$$(I+AB)^{-1}A = A(I+BA)^{-1} \quad (A.2)$$

as it can be easily verified pre- and post-multiplying both members of (A.1) by $I+AB$, and multiplying both members of (A.2) by $I+AB$ on the left side and $I+BA$ on the right side.

For the mixed sensitivity optimization, from definition (2) it follows that:

$$\begin{aligned} J(P_0, C) - J(M, C) &= \\ &= \begin{bmatrix} V_1 \left[(I+P_0C)^{-1} - (I+MC)^{-1} \right] \\ V_2 \left[P_0C(I+P_0C)^{-1} - MC(I+MC)^{-1} \right] \end{bmatrix} \end{aligned}$$

But:

$$\begin{aligned} (I+P_0C)^{-1} - (I+MC)^{-1} &= \\ &= (I+P_0C)^{-1}(I+MC)(I+MC)^{-1} + \\ &\quad - (I+P_0C)^{-1}(I+P_0C)(I+MC)^{-1} = \\ &= -(I+P_0C)^{-1}(P_0-M)C(I+MC)^{-1} \quad (A.3) \end{aligned}$$

and, applying the equality (A.1):

$$\begin{aligned} P_0C(I+P_0C)^{-1} - MC(I+MC)^{-1} &= \\ &= (I+P_0C)^{-1}P_0C - MC(I+MC)^{-1} = \\ &= (I+P_0C)^{-1}P_0C(I+MC)(I+MC)^{-1} + \\ &\quad - (I+P_0C)^{-1}(I+P_0C)MC(I+MC)^{-1} = \\ &= (I+P_0C)^{-1} [P_0C(I+MC) + \\ &\quad - (I+P_0C)MC] (I+MC)^{-1} = \\ &= (I+P_0C)^{-1}(P_0-M)C(I+MC)^{-1} \quad (A.4) \end{aligned}$$

so that:

$$\begin{aligned} J(P_0, C) - J(M, C) &= \\ &= \begin{bmatrix} -V_1 \\ V_2 \end{bmatrix} (I+P_0C)^{-1}(P_0-M)C(I+MC)^{-1}. \end{aligned} \quad (A.5)$$

From definitions (5) and (10), it results that:

$$\begin{aligned} R_0 - R &= \\ &= D_c^{-1} \left[(I+P_0C)^{-1}(P_0-P_x) - (I+MC)^{-1}(M-P_x) \right] D_x = \\ &= D_c^{-1} \left\{ (I+P_0C)^{-1}P_0 - (I+MC)^{-1}M + \right. \\ &\quad \left. - \left[(I+P_0C)^{-1} - (I+MC)^{-1} \right] P_x \right\} D_x \quad (A.6) \end{aligned}$$

Applying equalities (A.1) and (A.4):

$$\begin{aligned} (I+P_0C)^{-1}P_0 - (I+MC)^{-1}M &= \\ &= \left[(I+P_0C)^{-1}P_0C - (I+MC)^{-1}MC \right] C^{-1} = \\ &= \left[P_0C(I+P_0C)^{-1} - MC(I+MC)^{-1} \right] C^{-1} = \\ &= (I+P_0C)^{-1}(P_0-M)C(I+MC)^{-1}C^{-1} \quad (A.7) \end{aligned}$$

and then, by substitution of (A.7) and (A.3) in (A.6):

$$\begin{aligned} R_0 - R &= \\ &= D_c^{-1} (I+P_0C)^{-1}(P_0-M)C \cdot \\ &\quad (I+MC)^{-1}(C^{-1}+P_x)D_x = \\ &= D_c^{-1} (I+P_0C)^{-1}(P_0-M)C \cdot \\ &\quad (I+MC)^{-1}C^{-1}(I+CP_x)D_x \end{aligned}$$

and also:

$$\begin{aligned} (I+P_0C)^{-1}(P_0-M)C(I+MC)^{-1}C^{-1} &= \\ &= D_c(R_0-R)[(I+CP_x)D_x]^{-1} = \\ &= D_c(R_0-R)D_x^{-1}(I+CP_x)^{-1} \quad (A.8) \end{aligned}$$

which, substituted in (A.5), gives the result (11). For the H_{∞} design based on robustness optimization, from definition (3) it follows that:

$$J(P_0, C) - J(M, C) =$$

$$= \begin{bmatrix} P_0(I + CP_0)^{-1} - M(I + CM)^{-1} \\ (I + CP_0)^{-1} - (I + CM)^{-1} \end{bmatrix} [C \ I]. \quad (\text{A.9})$$

Applying equalities (A.2) and (A.4), it results that:

$$\begin{aligned} P_0(I + CP_0)^{-1} - M(I + CM)^{-1} &= \\ &= \left[P_0(I + CP_0)^{-1} C - M(I + CM)^{-1} C \right] C^{-1} = \\ &= \left[P_0 C (I + P_0 C)^{-1} - M C (I + M C)^{-1} \right] C^{-1} = \\ &= (I + P_0 C)^{-1} (P_0 - M) C (I + M C)^{-1} C^{-1} \end{aligned}$$

and, from (A.2) and (A.3):

$$\begin{aligned} (I + CP_0)^{-1} - (I + CM)^{-1} &= \\ &= \left[(I + CP_0)^{-1} C - (I + CM)^{-1} C \right] C^{-1} = \\ &= C \left[(I + P_0 C)^{-1} - (I + M C)^{-1} \right] C^{-1} = \\ &= -C (I + P_0 C)^{-1} (P_0 - M) C (I + M C)^{-1} C^{-1} \end{aligned}$$

which, substituted in (A.9) and exploiting (A.8), give the result (12). \square

Proof of Proposition 4.2.

From definition (18), the uncertainty model $\mathcal{R}(\hat{R}, W_{\hat{R}})$ is given by:

$$\begin{aligned} \mathcal{R}(\hat{R}, W_{\hat{R}}) &= \left\{ R = \hat{R} + \Delta : |\Delta(\omega)| \leq W_{\hat{R}}(\omega), \forall \omega \right\} = \\ &= \left\{ R : |R(\omega) - \hat{R}(\omega)|^2 \leq W_{\hat{R}}^2(\omega), \forall \omega \right\}. \quad (\text{A.10}) \end{aligned}$$

Since C is supposed to be stable, from definition (10) it results that, with the choice $D_c = 1$, $N_x = 0$ and $D_x = 1$:

$$R = \frac{M}{1 + MC}$$

For any ω , let us consider the inequality in (A.10), where the dependence on ω is omitted:

$$\begin{aligned} |R - \hat{R}|^2 \leq W_{\hat{R}}^2 &\Leftrightarrow \\ \left(\frac{M}{1+MC} - \hat{R} \right) \left(\frac{M}{1+MC} - \hat{R} \right)^* &\leq W_{\hat{R}}^2 \Leftrightarrow \\ [M(1-C\hat{R}) - \hat{R}][M(1-C\hat{R}) - \hat{R}]^* &\leq W_{\hat{R}}^2 [1+MC]^2 \Leftrightarrow \\ MM^* [1-C\hat{R}]^2 - W_{\hat{R}}^2 |C|^2 - M[\hat{R}(1-C\hat{R})^* + W_{\hat{R}}^2 C^*]^* + \\ -M^*[\hat{R}(1-C\hat{R}) + W_{\hat{R}}^2 C] + |\hat{R}|^2 &\leq W_{\hat{R}}^2 \Leftrightarrow \\ MM^* - M\hat{M}^* - M^*\hat{M} - \hat{M}\hat{M}^* &\leq W_{\hat{M}}^2 \Leftrightarrow \\ (M - \hat{M})(M - \hat{M})^* &\leq W_{\hat{M}}^2 \Leftrightarrow \\ |M - \hat{M}|^2 &\leq W_{\hat{M}}^2 \end{aligned}$$

with \hat{M} and $W_{\hat{M}}$ defined by (29) and (30) respectively. Then, there is a one-to-one mapping between

the elements of $\mathcal{R}(\hat{R}, W_{\hat{R}})$ and the elements of the uncertainty model $\mathcal{M}(\hat{M}, W_{\hat{M}})$ defined as the set:

$$\begin{aligned} \mathcal{M}(\hat{M}, W_{\hat{M}}) &= \left\{ M : |M(\omega) - \hat{M}(\omega)|^2 \leq W_{\hat{M}}^2(\omega), \forall \omega \right\} = \\ &= \left\{ M = \hat{M} + \Delta : |\Delta(\omega)| \leq W_{\hat{M}}(\omega), \forall \omega \right\} \end{aligned}$$

thus proving the claim. \square

References

- Bitmead, R.R., M. Gevers and A.G. Partanen (1997). Introducing caution in iterative controller design. *Prepr. 11th IFAC Symp. System Identification*, Kitakyushu, Japan, pp. 1701-1706.
- Canale, M., S.A. Malan and M. Milanese (1998). Model quality evaluation in identification for H_∞ control. *IEEE Trans. Autom. Control*, AC-43, no. 1.
- Canale, M., S. Malan, M. Milanese and M. Taragna (1996). Model structure selection in identification for control. *Proc. 13th IFAC World Congress*, San Francisco, CA, vol. I, pp. 109-114.
- Ceton, C., P.M.R. Wortelboer and O.H. Bosgra (1993). Frequency weighted closed-loop balanced reduction. *Proc. 2nd European Control Conf.*, Groningen, the Netherlands, pp. 697-701.
- De Callafon, R.A. and P.M.J. Van den Hof (1997). Suboptimal feedback control by a scheme of iterative identification and control design. *Mathem. Modelling of Systems*, 3, 77-101.
- Fiorio, G., S. Malan, M. Milanese and M. Taragna (1997). Robust design of low order controllers via uncertainty model identification. *Proc. 2nd IFAC Symp. Robust Control Design*, Budapest, Hungary, pp. 49-54.
- Giarré, L., M. Milanese and M. Taragna (1997). H_∞ identification and model quality evaluation. *IEEE Trans. Autom. Control*, AC-42, pp. 188-199.
- Hakvoort, R.G. and P.M.J. Van den Hof (1995). A practically feasible procedure for identification of H_∞ error bounds. *Proc. 3rd Europ. Control Conf.*, Rome, Italy, pp. 925-930.
- Kaciewicz, B.Z., M. Milanese, R. Tempo and A. Vicino (1986). Optimality of central and projection algorithms for bounded uncertainty. *Systems & Control Lett.*, 8, 161-171.
- Kurzanski, A.B. and V.M. Veliov (Eds.) (1994). *Modeling Techniques for Uncertain Systems*, Boston, Birkhäuser.
- McFarlane, D. and K. Glover (1990). *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions*. Berlin: Springer.
- Milanese, M., R. Tempo and A. Vicino (Eds.) (1989). *Robustness in Identification and Control*, New York, Plenum Press.

Mixed objectives MIMO control design for a Compact Disc player

M. Dettori[†], V. Prodanovic and C. W. Scherer

*Mechanical Engineering Systems and Control Group
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.
E-mail: m.dettori@wbmt.tudelft.nl*

Abstract. This paper investigates the application of LMI-based mixed objectives design techniques to a CD player mechanism. In this control design problem the main goal is to keep the time domain amplitude of a tracking error signal bounded in the presence of disturbances and norm bounded uncertainties. To this end we identified in the \mathcal{H}_∞ norm and in the so-called generalized \mathcal{H}_2 norm suitable measures to represent our specifications. The resulting design is shown to exhibit significant performance improvements if compared to a single-objective \mathcal{H}_∞ design.

Keywords. Mixed and multi-objective control, LMI, \mathcal{H}_∞ , generalized \mathcal{H}_2 , Compact Disc player.

1 Notation

Due to the lack of a standard notation, we use $\|T\|_{2 \rightarrow \infty}$ to represent the generalized \mathcal{H}_2 norm of the system T . This notation reflects the fact that this norm is the induced gain of the system from L_2 to L_∞ . For the system gain from L_2 to L_2 , we prefer the standard notation $\|T\|_\infty$.

2 Introduction

A Compact Disc (CD) player is an optical data storage device that decodes and reproduces binary coded information. The information signal is stored in a spiral shaped track on a reflective disc. Starting from the original audio application, the field of use of such a device has been gradually enlarged to new high-performance applications, like CD-ROM or the recent DVD-ROM. A demand emerging from these new applications is to obtain a faster data readout and a shorter access time, together with a higher density of the data on the disc. The way to achieve these improvements is an increase of the rotational frequency of the disc, which requires a

corresponding increase of the bandwidth of the mechanical servo-systems. The presence of parasitic resonances at high frequencies, together with the variations occurring from player to player (due to the manufacturing tolerances in mass-production) lead to the necessity of designing robust controllers. In this paper we will investigate the possibility of improving the track-following and the focusing behavior, by applying recently developed mixed objectives controller design techniques. The name mixed objectives (Bernstein and Haddad, 1989; Khargonekar and Rotea, 1991) stands for the fact that different performance specifications (either in the frequency or in the time domain) are posed on different channels of the plant, such that the transient and steady state behavior, disturbance rejection, and robustness against structured and unstructured uncertainty can be taken into account. Several techniques have been proposed in the literature to handle this kind of problems. We can distinguish two main approaches. The first one uses the Youla parametrization of the controller to cast the problem in an LMI framework (Sznaier and Sideris, 1991; Scherer, 1995b). This approach allows to solve the mixed problems for independent objectives (in this case the name multi-objective problem is used), but presents

[†]The research of Marco Dettori is sponsored by Philips Research Laboratories, Eindhoven, The Netherlands

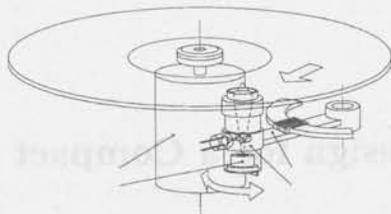


Fig. 1: Schematic view of the CD mechanism

a severe disadvantage. It makes use of certain approximation techniques; improving the accuracy of the approximation lets the order of the controller grow drastically. Often optimal controllers turn out to be infinite dimensional. The second approach overcomes this difficulty at the price of introducing a dependence among the different objectives. We chose to adopt this second approach, which leads to rational controllers of the same order of the generalized plant. As shown in Masubuchi *et al.* (1995), Scherer (1995a), Scherer *et al.* (1997), the formulation of the desired objectives for the closed loop system with analysis LMIs leads directly to synthesis LMIs in terms of transformed versions of the controller parameters that can be easily solved. We are mainly interested in showing that mixed control design is a useful tool for the designer, providing him a framework to express every specification in a "natural way". This avoids the difficult (and often to some extent arbitrary) process of translating all the requirements to the same setting (typically the frequency domain). Being this our purpose, in this paper we will not care too much about other important aspects, like for example representing the model uncertainty in the least conservative way. All these aspects can, in fact, be incorporated in the design, at the price of increasing its complexity.

3 System description

In Fig.1 a schematic view of the CD mechanism is shown. The rotation of the disc is produced by a turn-table DC-motor. The rotational velocity varies according to the position on the disc of the track that is being read. The rotational frequency therefore varies approximatively from 8 Hz (innermost position on the disc) to 4 Hz (outermost position). Track following is performed by a radial arm at the end of which an optical element is mounted. A diode in this element generates a laser beam which is focused, through a system of lenses, in a spot on the information layer of the disc. Focusing action of this spot is performed by an objective lens that can move in a vertical direction. A system of four photodiodes

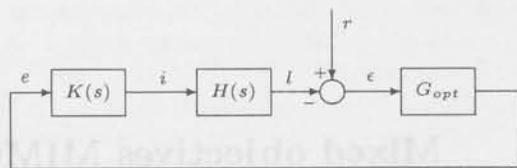


Fig. 2: Block diagram of the CD mechanism

provides position error information, which are the only signals available for control. In fact, neither the track position nor the true spot position can be measured. A controller is needed for accurate radial and focus positioning of the laser spot. In current implementations, the radial and the focus directions are controlled using two independent SISO schemes. This is allowed by the relatively low dynamic interaction between the two loops. MIMO control design has been investigated for the CD player in Steinbuch *et al.* (1994) using μ synthesis.

In Fig.2 a block diagram of the CD mechanism is shown. Each signal is a vector with two components, the radial and the focus one. $H(s)$ is the transfer function of the mechanical actuator which is controlled by the current i and generates the laser spot l on the disc. G_{opt} is the gain of the optical pickup mechanism which converts the displacements ϵ between track and spot in an error signal e . The controller K processes this error signal and generates the current i .

4 Modeling

The only transfer function that can be identified is the one between i and e , that is $P(s) = G_{opt}H(s)$. The gain of this transfer function varies in a non-linear way according to the position on the disc of the track that is being read, due to the movement of the radial arm. In our model we do not consider this effect, but we assume that the laser is reading a track in the middle of the disc, corresponding to a rotational frequency of 6 Hz. As already mentioned, the track position r is not known; we will regard it as a disturbance signal $d = G_{opt}r$ acting at the output of the plant $P(s)$.

In Fig.3 the frequency response (amplitude) of a 12th order model of $P(s)$ is shown. This model has been fitted on a frequency response obtained by spectrum analysis techniques (de Callafon *et al.*, 1996). The MIMO fit has been performed using the toolbox Freqid (de Callafon and Van den Hof, 1996), which implements an iteration based on the Sanathan-Koerner procedure to minimize a least square criterion. The chosen order for the model appeared to be the best trade-off between the con-

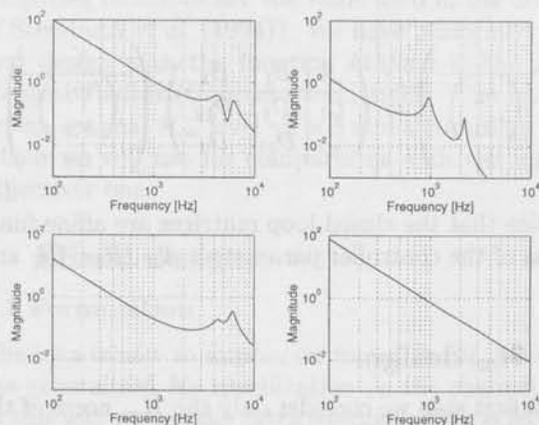


Fig. 3: Frequency response of the 12th order model of $P(s)$

flicting requirements of accuracy in representing the real system and low complexity to make a sensible control design possible. This last issue, as we will see, is critical in LMI-based designs.

At low frequencies, the diagonal elements behave like double integrators, due to the rigid body mode of the radial arm ((1, 1) element) and the optical pick-up unit ((2, 2) element). At higher frequencies, parasitic dynamics appear due to mechanical resonances of the radial arm and the mounting plate, and due to flexible modes of the disc. These resonant modes are especially present in the radial direction, producing two peaks at about 4 and 7 kHz. The main sources of uncertainty we want to account for are the unstructured difference between model and measurements and the variation in the locations of the parasitic resonances. The latter is an effect of manufacturing tolerances in mass production which manifest themselves as variations in the frequency response from player to player. Although this sort of uncertainty can be better described as real parametric, in order to not increase the complexity of the design we will consider it as unstructured norm bounded perturbation.

5 Performance specification

The main issue in the control of a CD player is to guarantee a hard bound on the time domain amplitude of the position error signals. To avoid losing track, the maximum allowable error should be $0.1\mu\text{m}$ in the radial direction and $1\mu\text{m}$ in the focus direction. These bounds should be attained in the presence of disturbances. In our design we will consider two major sources of disturbances, one for each direction: eccentricity of the track and undulation of

the surface of the disc. By standardization of Compact Discs these quantities cannot exceed $100\mu\text{m}$ and 1mm respectively. From these data we derive the crucial specification:

In both directions a time domain attenuation of the disturbances of a factor 1000 should be achieved.

So far, the approach that has been followed to tackle this problem is to translate this requirement into frequency domain specification on the shape of the sensitivity function $S = (I + PK)^{-1}$ (see e.g. Steinbuch *et al.* (1994)). This translation is based on an assumption made on the unmeasurable track signal: its spectrum is assumed to be a series of pulses centered around the rotational frequency and its higher harmonics, with an amplitude that is decaying at a rate -40dB/dec . Experience showed that, in order to meet the time domain specification for the error, the sensitivity should stay below -60dB at the rotational frequency and exhibit sensible attenuation up to 200Hz . Its slope cannot be larger than 40dB/dec in this frequency region. Obviously, this disturbance attenuation requirement puts a lower bound on the achievable closed-loop bandwidth. High bandwidth is undesirable for several reasons: it implies high power consumption (critical especially in portable use), amplification of audible noise and, last but not least, poor robustness against variations in the resonance peaks and unmodeled high frequency dynamics. As a consequence the wish is for the lowest possible bandwidth, compatibly with the required disturbance suppression. This translation of the specifications into the frequency domain is, nevertheless, not completely satisfactory. Being based on a rule of thumb and not on a thorough comprehension of the interaction between time and frequency characteristics, it can in principle lead to conservative designs. The harmonics of the track disturbance spectrum will, in fact, sum up in an unknown way (which will depend on the unknown phase behavior) to the error in the time domain. Our purpose is, therefore, to approach the problem by taking its inherent time-domain nature into account.

6 Problem setting

The first relevant aspect in control design is to choose the criterion which is suitable for the problem at hand. An \mathcal{H}_∞ criterion appears quite a natural choice to take into account robustness aspects and to shape (some of) the relevant closed-loop transfer function(s) in order e.g. to achieve a specified bandwidth. On the other side, the generalized \mathcal{H}_2 norm is convenient to express the disturbance rejection specification. We recall that this norm is the gain

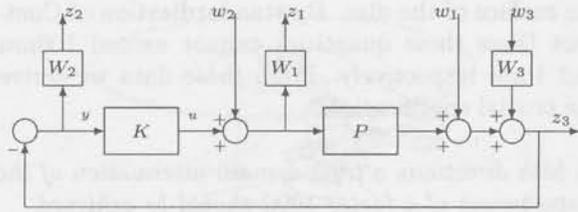


Fig. 4: Control scheme for design

of the system from L_2 to L_∞ . Its value has, therefore, the interpretation of a worst case time domain amplitude amplification for inputs of finite energy. In choosing this norm we are implicitly making a modeling assumption, namely that the disturbances acting on the system are signals of finite energy. An alternative assumption could be modeling the disturbances as bounded amplitude signals and using the peak-to-peak norm. What played an important role in the choice of the first option was the fact that, at present, there are no LMI algorithms available to minimize the peak-to-peak norm. In fact, it is possible only to minimize an upper bound of it (see Scherer *et al.* (1997)) that can be quite loose. As a side remark we like to stress that disturbance modeling emerges as a key issue. Unfortunately, this is not an easy task for the CD player: the track eccentricity is not measurable and it is highly attenuated by the control system (by a factor 1000), what makes it hard to reconstruct it from the error signal.

In Fig.4 the scheme of the generalized plant that we wanted to adopt for this design is depicted. The blocks W_1 , W_2 , and W_3 represent the design weighting functions, w_1 , w_2 and w_3 are the exogenous inputs, z_1 , z_2 and z_3 are the performance outputs, y is the measured output and u is the control input. We consider two performance/uncertainty channels: $T_1 : (w_1^T, w_2^T)^T \rightarrow (z_1^T, z_2^T)^T$ and $T_2 : w_3 \rightarrow z_3$. We have the following state space representations for the generalized plant

$$\begin{pmatrix} \dot{x} \\ z_1 \\ z_2 \\ z_3 \\ y \end{pmatrix} = \begin{bmatrix} A & B_1 & B_2 & B \\ C_2 & D_{21} & D_2 & E_2 \\ C & F_1 & F_2 & 0 \end{bmatrix} \begin{pmatrix} x \\ w_1 \\ w_2 \\ w_3 \\ u \end{pmatrix} \quad (1)$$

the controller K

$$\begin{pmatrix} \dot{x}_K \\ u \end{pmatrix} = \begin{pmatrix} A_K & B_K \\ C_K & D_K \end{pmatrix} \begin{pmatrix} x_K \\ y \end{pmatrix} \quad (2)$$

and the closed loop system

$$\begin{pmatrix} \dot{\zeta} \\ z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} A & B_1 & B_2 \\ C_1 & D_1 & D_{12} \\ C_2 & D_{21} & D_2 \end{pmatrix} \begin{pmatrix} \zeta \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \quad (3)$$

Notice that the closed loop matrices are affine functions of the controller parameters A_K , B_K , C_K and D_K .

7 \mathcal{H}_∞ design

As a first step we consider only the \mathcal{H}_∞ norm of the transfer matrix T_1 , that is

$$\left\| \begin{pmatrix} -W_1 K S & W_1 (I + K P)^{-1} \\ -W_2 S & -W_2 S P \end{pmatrix} \right\|_\infty$$

where $S = (I + P K)^{-1}$ is the sensitivity function, and we design a controller that minimizes this norm. This is a standard mixed KS/SP \mathcal{H}_∞ design which is suitable to achieve robust performance (expressed in terms of the transfer function SP) in presence of unstructured additive uncertainty (see e.g. Skogestad and Postlethwaite (1996)). The weighting matrix W_1 is

$$W_1(s) = \begin{pmatrix} w_u(s) & 0 \\ 0 & w_u(s) \end{pmatrix}$$

where

$$w_u(s) = \frac{5(s+10^3 \cdot 2\pi)(s+1.2 \cdot 10^3 \cdot 2\pi)(s+5 \cdot 10^3 \cdot 2\pi)}{(s+50 \cdot 2\pi)(s+10^4 \cdot 2\pi)(s+1.1 \cdot 10^4 \cdot 2\pi)}$$

has high-pass characteristics (the two high frequent poles are only needed to render it proper). It reflects the size of the additive uncertainty at frequencies above 600Hz and it is also used to force high-frequency roll-off of the controller which limits the bandwidth.

The weighting matrix

$$W_2(s) = \begin{pmatrix} w_{rad}(s) & 0 \\ 0 & w_{foc}(s) \end{pmatrix}$$

where

$$w_{rad}(s) = \frac{0.25(s+75 \cdot 2\pi)(s+100 \cdot 2\pi)}{(s+0.1 \cdot 2\pi)(s+25 \cdot 2\pi)},$$

$$w_{foc}(s) = \frac{0.25(s+70 \cdot 2\pi)(s+80 \cdot 2\pi)}{(s+0.15 \cdot 2\pi)(s+20 \cdot 2\pi)},$$

specifies the performance independently for the radial and the focus directions. Its elements have low-pass behaviors with almost integrating action to achieve disturbance suppression and zeros chosen in order to limit the peaking of the sensitivity (These

weighting functions are the same used in the design of Steinbuch *et al.* (1994)). We have performed this first design with the function *hinfmix* in the *LMI Control Toolbox* (Gahinet *et al.*, 1995). The results are an optimal \mathcal{H}_∞ level $\gamma_1^o = 5$ and a controller K_1 , which we will use for comparisons with the mixed objectives one.

8 Mixed design

8.1 Formulation

The idea is now to impose on top of the \mathcal{H}_∞ design the generalized \mathcal{H}_2 specification on the channel T_2 . In this way we use the extra freedom which is left after the \mathcal{H}_∞ optimization to achieve the desired time domain behavior. On the basis of this observation we can realize that full advantage of mixed design can be taken only in the MIMO case. In fact, in the standard SISO design problems, it is known that the set of optimal \mathcal{H}_∞ controllers is a singleton. Hence, also for suboptimal designs one expects very little freedom to satisfy extra constraints.

We choose the weighting W_3 as:

$$W_3(s) = \begin{pmatrix} w_p(s) & 0 \\ 0 & w_p(s) \end{pmatrix}$$

where

$$w_p(s) = \frac{(2\pi \cdot 6)^2}{s^2 + 1.2 \cdot 2\pi \cdot 8s + (2\pi \cdot 6)^2}$$

Its role is to shape the track disturbance, whose spectrum is assumed to be the one described previously in section 4.

Ideally we would like to solve the real multi-objective problem (i.e. the objectives are mutually independent), that amounts to minimize $\|T_2\|_{2 \rightarrow \infty}$ over the set of all stabilizing controllers that render $\|T_1\|_\infty \leq \gamma_1^o$ satisfied. To explicitly formulate this minimization, we need to recall two analysis results.

- The closed loop system (3) has $\|T_1\|_\infty \leq \gamma_1$ if and only if there exists a matrix $\mathcal{X}_1 > 0$ such that

$$\begin{pmatrix} A^T \mathcal{X}_1 + \mathcal{X}_1 A & \mathcal{X}_1 B_1 & C_1^T \\ B_1^T \mathcal{X}_1 & -\gamma_1 I & D_1^T \\ C_1 & D_1 & -\gamma_1 I \end{pmatrix} < 0 \quad (4)$$

- The closed loop system (3) has $\|T_2\|_{2 \rightarrow \infty} \leq \gamma_2$ if and only if there exists a matrix $\mathcal{X}_2 > 0$ such that

$$\begin{pmatrix} A^T \mathcal{X}_2 + \mathcal{X}_2 A & \mathcal{X}_2 B_2 \\ B_2^T \mathcal{X}_2 & -\gamma_2 I \end{pmatrix} < 0 \quad (5)$$

$$\begin{pmatrix} \mathcal{X}_2 & C_2^T \\ C_2 & \gamma_2 I \end{pmatrix} > 0 \quad (6)$$

$$D_2 = 0. \quad (7)$$

The multi-objective problem amounts to minimize γ_2 under (4), (5), (6) and (7) for $\gamma_1 = \gamma_1^o$. However existing algorithms allow to solve the synthesis problem only after introducing an extra constraint which couples the objectives. We have in fact to try to satisfy the different performance requirements with a common Lyapunov matrix (see Scherer (1995a), Scherer *et al.* (1997)). In this way we are introducing conservatism in our problem, whose amount is hard to quantify. The introduction of this conservatism constitutes the difference between multi-objective and mixed objectives problems. As a practical effect, with this extra constraint the minimization written above turns out to be infeasible. We suggest, instead, to solve the modified problem:

$$\min a\gamma_1 + b\gamma_2 \quad (8)$$

over the constraints (4), (5), (6), (7) and $\mathcal{X}_1 = \mathcal{X}_2 =: \mathcal{X}$. Varying the coefficients a and b we can vary the relative weight of the two criteria in the minimization and see what is the best trade-off between them (measured in terms of the performance achieved by the corresponding controllers). In our design the choice $a = 10$ and $b = 1$ appeared to give the most satisfactory results.

8.2 Synthesis

The minimization in (8) is clearly not convex in \mathcal{X} and the controller parameters. We need therefore to convexify the problem by applying the formal block substitution procedure described in Scherer *et al.* (1997) to get LMI's in the transformed set of variables X, Y, K, L, M and N . The final formulation becomes:

Minimize $10\gamma_1 + \gamma_2$ over the constraints (9)-(12).

This minimization can be coded in one of the available software packages. We used *LMILAB* of the *LMI Control Toolbox* (Gahinet *et al.*, 1995). Notice that the equality constraint (12) can not be directly handled by the solver; in our case, however, due to the structure of D_2, E_2 and F it was automatically satisfied. Once a solution (if any) has been determined, one has only to reconstruct the original controller parameters by applying the inverse variables transformation (Scherer *et al.*, 1997).

However in our design the solver was unable to return a solution, running out of memory at a certain moment of the iteration. The problem has been

$$\begin{pmatrix} AX + XA^T + BM + (BM)^T & K^T + A + BNC & B_1 + BNF_1 & (C_1X + E_1M)^T \\ K + (A + BNC)^T & A^TY + YA + LC + (LC)^T & YB_1 + LF_1 & (C_1 + E_1NC)^T \\ (B_1 + BNF_1)^T & (YB_1 + LF_1)^T & -\gamma_1 I & (D_1 + E_1NF_1)^T \\ C_1X + E_1M & C_1 + E_1NC & D_1 + E_1NF_1 & -\gamma_1 I \end{pmatrix} < 0 \quad (9)$$

$$\begin{pmatrix} AX + XA^T + BM + (BM)^T & K^T + A + BNC & B_2^T + BNF_2 \\ K + (A + BNC)^T & A^TY + YA + LC + (LC)^T & (YB_2)^T + LF_2 \\ B_2 + (BNF_2)^T & YB_2 + (LF_2)^T & -\gamma_2 I \end{pmatrix} < 0 \quad (10)$$

$$\begin{pmatrix} X & I & (C_2X + E_2M)^T \\ I & Y & (C_2 + E_2NC)^T \\ C_2X + E_2M & C_2 + E_2NC & \gamma_2 I \end{pmatrix} > 0 \quad (11)$$

$$D_2 + E_2NF = 0 \quad (12)$$

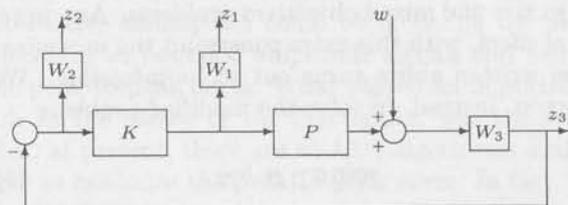


Fig. 5: Modified control scheme

identified in the too high number of decision variables involved. Considering that the generalized plant (combination of $P(s)$ and the three weightings) has order 26, the number of decision variables involved in the optimization procedure was 1488 (351 independent entries each for X and Y that are symmetric 26×26 , 676 for K , 52 for L , 52 for M , 4 for N plus γ_1 and γ_2). There are two possible ways to overcome this obstacle. The first one is trying to further reduce the order of the system and/of the weightings, at the price of losing realism in the design. The second one, that we adopted, is applying techniques based on the so-called Projection Lemma to eliminate variables from the LMI's. In particular, the elimination of the matrix K yields a saving of 676 decision variable. But in order to do that, we have to change the structure of the control scheme. Elimination of K is, in fact, possible only if the exogenous input is the same for every channel (i.e. $w_1 = w_2 = w_3$). The modified control scheme is shown in Fig.5. The \mathcal{H}_∞ channel $w \rightarrow (z_1^T, z_2^T)^T$ has now an S/KS structure and the weight W_2 should be modified to express performances in terms of S and not of SP anymore. Regarding the generalized \mathcal{H}_2 channel, the weight W_3 is moved from the input to the output (losing its interpretation of disturbance shaping filter). With this scheme the compu-

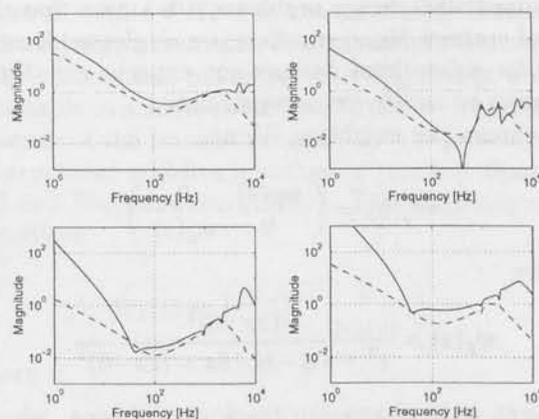


Fig. 6: Frequency response of K_1 (dashed line) and K_2 (solid line)

tation was successful, giving a controller K_2 which ensures an \mathcal{H}_∞ level $\gamma'_1 = 8.7$ for T_1 and a generalized \mathcal{H}_2 level $\gamma'_2 = 3.4$ for T_2 . In Fig.6 the amplitude plots for the two controllers K_1 (dashed line) and K_2 (solid line) are shown. We see that K_2 has higher gain at low frequencies, achieving a better disturbance suppression, and more aggressive behavior at high frequencies, resulting in a higher closed-loop bandwidth. These observations are confirmed by the analysis of the sensitivity transfer function in Fig.7. The mixed design exhibits a higher disturbance rejection and a lower peaking, managing to spread the area of S above 0 dB (whose amount increases with the disturbance attenuation, according to the Bode sensitivity integral relation) over a wider frequency region. This second feature implies less perceptible amplification of audible noise. But our main interest is in the comparison of the

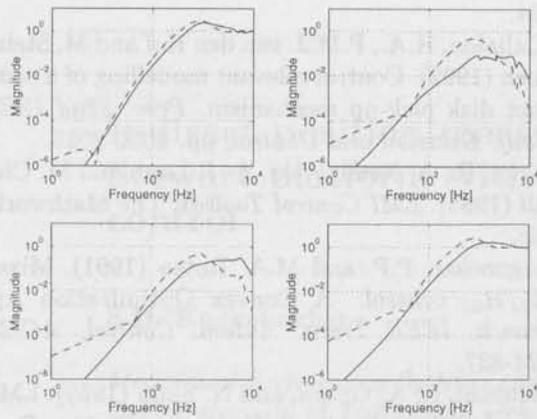


Fig. 7: Sensitivity function for the \mathcal{H}_∞ (dashed line) and the mixed (solid line) designs

time domain properties of the two designs. In Fig.8 is represented the response to a step on the disturbance track. This signal is relevant in the CD player,

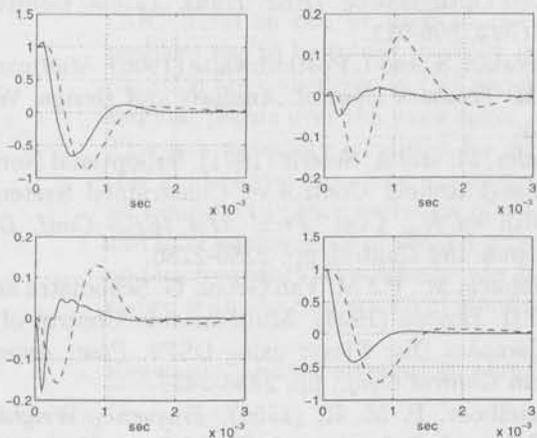


Fig. 8: Response to a disturbance step for the \mathcal{H}_∞ (dashed line) and the mixed (solid line) designs

occurring when there is a command from the user to jump to another track; obviously this should be done in the shortest possible time. The mixed controller performs sensibly better both in terms of the overshoot and the settling time. In the radial direction the overshoot decreases by a factor of 25% and the settling time by a factor of 40%. In the focus direction the improvements are respectively 45% and 50%. Also for the off-diagonal elements the situation is better (the exception being a higher negative peak in the (2,1) element, but the relative value is quite small). As a last comparison we show in Fig.9

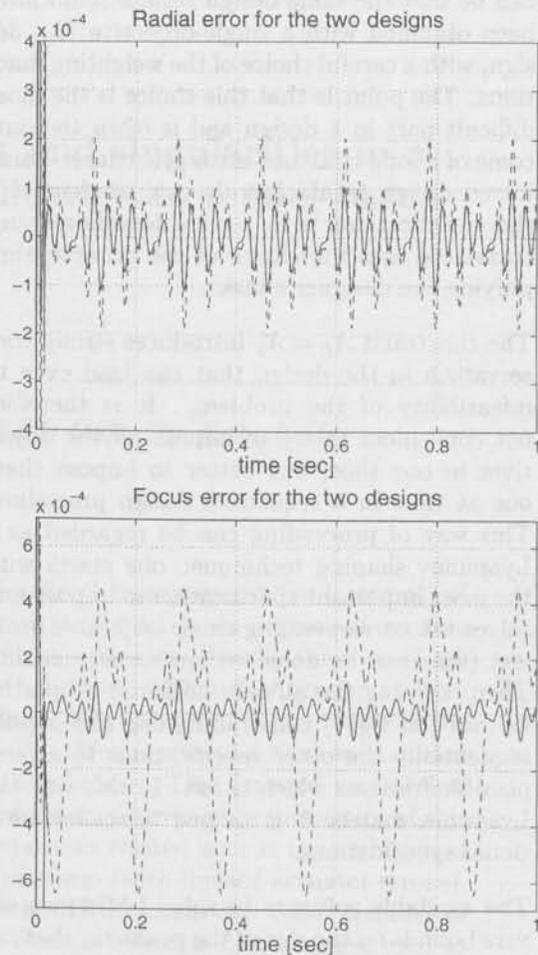


Fig. 9: Radial and focus error for the test disturbance for the \mathcal{H}_∞ (dashed line) and the mixed (solid line) designs

the radial and the focus responses to the test disturbance $w(t) = \sum_{n=1}^7 \frac{1}{n^2} \sin(2n\pi \cdot 6t + \phi_n)$, which represents the model hypothesized for the track eccentricity and the undulation of the disc (ϕ_n is a random initial phase uniformly distributed between 0 and 2π). The plots show that the peak of the error achieved by the mixed controller is about 2 times smaller in the radial direction and about 6 times smaller in the focus.

9 Conclusions

The most important conclusions with respect to the use of mixed objectives LMI-based techniques to this design problems are:

- Mixed design techniques give a powerful tool to express "in a natural way" a wide range of performance specifications. A possible objection

can be that the same design results could have been obtained with a single-objective \mathcal{H}_∞ design, with a careful choice of the weighting functions. The point is that this choice is the most difficult part in a design and is often the outcome of a long trial-and-error procedure. Using mixed design techniques the way to represent the specifications is in a considerable amount translated into the choice of the criteria, simplifying the designer's task.

- The constraint $\mathcal{X}_1 = \mathcal{X}_2$ introduces strong conservatism in the design that can lead even to infeasibility of the problem. It is therefore not convenient to try to impose all the objectives in one shot, but better to impose them one at time in a sequential design procedure. This way of proceeding can be regarded as a Lyapunov shaping technique: one starts with the most important specification to impose and solves the corresponding single-objective problem (that can be done without conservatism). Then, keeping the already achieved properties as (not too tight) constraints, one can impose sequentially the other specifications to exploit possible freedom which is left. In this way the Lyapunov matrix \mathcal{X} is shaped to realize additional specifications.
- The available software to solve LMIs puts severe bounds on the size of the problems that are tractable. This results in a difficulty in doing designs for realistic (not too low order) models and/or in a limited flexibility in choosing the design structure if we need to eliminate LMI variables.
- The software appeared also to be extremely sensitive to numerical conditioning of the data (much more, for example, than the \mathcal{H}_∞ solver of the μ Toolbox). Particular precautions have then to be taken, like moving the poles in 0 a bit on the left (e.g. -10^{-3}), time-scaling of the data to restrict the numerical range, performing closed-loop frequency-weighted model reduction (see Wortelboer (1994)).

References

- Bernstein, D.S. and W.M. Haddad (1989). LQG control with an \mathcal{H}_∞ performance bound: a Riccati equation approach. *IEEE Trans. Autom. Control*, AC-34, 293-305.
- De Callafon, R.A. and P.M.J. Van den Hof (1996). FREQID - Frequency domain identification toolbox for use with Matlab. In: O.H. Bosgra et al. (Eds.), *Selected Topics in Identif., Modelling and Control*, Vol. 9, Delft University Press, pp. 129-134.
- De Callafon, R.A., P.M.J. van den Hof and M. Steinbuch (1992). Control relevant modelling of a compact disk pick-up mechanism. *Proc. 32nd IEEE Conf. Decision and Control*, pp. 2050-2055.
- Gahinet, P., A. Nemirovski, A. J. Laub and M. Chilali (1995). *LMI Control Toolbox*. The Mathworks Inc.
- Khargonekar, P.P. and M.A. Rotea (1991). Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control: A Convex Optimization Approach. *IEEE Trans. Autom. Control*, AC-36, 824-837.
- Masubuchi, I., A. Ohara, and N. Suda (1995). LMI-Based Output Feedback Controller Design. *Proc. American Control Conf.*, pp. 3473-3477.
- Scherer, C.W. (1995a). Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control. In: A. Isidori (Ed.), *Trends in Control, A European Perspective*. Springer Verlag, Berlin, pp. 173-216.
- Scherer, C.W. (1995b). Multiobjective $\mathcal{H}_2/\mathcal{H}_\infty$ control. *IEEE Trans. Autom. Control*, AC-40, 1054-1062.
- Scherer, C.W., P. Gahinet and M. Chilali (1997). Multi-Objective Output-Feedback Control via LMI Optimization. *IEEE Trans. Autom. Control*, AC-42, 896-911.
- Skogestad, S. and I. Postlethwaite (1996). *Multivariable Feedback Control. Analysis and Design*. Wiley.
- Sznaier, M. and A. Sideris (1991). Suboptimal Norm Based Robust Control of Constrained Systems with an H_∞ Cost. *Proc. 31st IEEE Conf. Decision and Control*, pp. 2280-2286.
- Steinbuch, M., P.J.M. Van Groos, G. Schootstra and O.H. Bosgra (1994). Multivariable Control of a Compact Disc Player using DSPs. *Proc. American Control Conf.*, pp. 2434-2438.
- Wortelboer, P. M. R. (1994). *Frequency-Weighted Balanced Reduction of Closed-Loop Mechanical Servo-Systems: Theory and Tools*. Ph.D. Thesis, Mech. Eng. Sys. and Contr. Group, Delft Univ. of Techn., The Netherlands.

Indirect position measurement and singularities in a Stewart platform with an application to model-based control

S.H. Koekebakker

*Mechanical Engineering Systems and Control Group
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
E-mail: s.h.koekebakker@wbmt.tudelft.nl*

Abstract. This paper considers the problem of indirect measurement (and control) of the coordinates of a Stewart platform. The Stewart platform is a six-degrees-of-freedom robot with (6) parallel actuators to be used as a flight simulator. Usually the pose of the system is indirectly measured by the lengths of the actuators. Although the mapping from platform coordinates to actuator length is well defined, it is not injective and therefore the inverse transformation has several solutions globally. Over a subset, a Newton-Raphson (NR) iteration can be used to calculate the local solution. This forward kinematical problem has to be solved to apply model based feedback. Convergence properties of this iteration are considered in this paper. Another important related issue is the exclusion of singular points over the work space of a Stewart platform (with limited actuator stroke). This is a necessary condition for convergence of the NR-iteration and controllability of this kind of systems. The parameters of the new Simona research simulator are used as an example to prove exclusion of singular points from the work space of a parallel robot and infer convergence of the NR-scheme. With guaranteed fast convergence at a sufficient update frequency, this scheme can be used in a model based feedback loop. This will be shown with the Simona flight simulator motion system.

Keywords. Flight simulation, parallel motion systems, robotics, forward kinematics.

1 Introduction

In parallel robots such as most flight simulation motion platforms, the position of the system is usually indirectly measured by the length of the actuators. The forward kinematical problem of calculating the platform coordinates given the actuator lengths of a fully parallel six-degrees of freedom system, a so-called Stewart platform (Stewart, 1965), is seen to be solved in roughly two ways in literature.

Using analytic techniques the problem can be transformed to a set of combined polynomial equations whose roots have to be found to solve the forward kinematics (e.g. Husty, 1996). Although these equations can provide insight into the structure of the problem, no closed form solution is known to be presented yet.

Secondly, for long, the forward kinematics has been tackled numerically by performing a Newton-Raphson (NR) iteration scheme (Dieudonne *et al.*, 1972).

$$\bar{x}_{k+1} = \bar{x}_k - J^{-1}(\bar{x}_k)(\bar{l}^* - \bar{l}_k) \quad (1)$$

As the actuator lengths, \bar{l} , are explicit functions of the platform coordinates, \bar{x} , the jacobian, J , is a function of platform coordinates.

$$J_{ij}(\bar{x}) = \frac{\partial \bar{l}_i(\bar{x})}{\partial \bar{x}_j} \quad (2)$$

This latter method is preferred here as it is less involved to be implemented in a real-time model based controller where inversion of the jacobian is already part of the control structure. But although this has

not seen to be considered in literature w.r.t Stewart platforms, convergence or convergence to the actual physical pose in this scheme is not guaranteed in general. As the forward kinematics of a Stewart platform have more than one solution (Husty, 1996), and since singular points of the jacobian for unconstrained actuator lengths exist (Ma and Angeles, 1991), the iteration scheme does not converge globally to the right solution.

A dynamic model of a parallel robot is described as a function of the platform position and its derivatives (Koekebakker *et al.*, 1996). To apply model based feedback (e.g. computed torque, Nijmeijer and Van der Schaft (1990)) based on the actual platform state instead of the desired state (Liu *et al.*, 1991) one would like to guarantee both convergence of the NR-scheme and exclusion of the singular points (singular J -matrix) in the work space. In this paper an algorithm is presented with which this can be guaranteed for general but known (inverse) kinematics of the Stewart platform at hand. Practical relevance is shown by application to the Simona flight simulator motion system (Advani *et al.*, 1997).

First a general theorem on convergence of the Newton-Raphson iteration is considered. After introducing the kinematics of the Stewart platform, the jacobian, J , is shown to be Lipschitz i.e. its change w.r.t. two poses is bounded by a constant times the difference of the respective platform coordinates. With this condition it is possible to derive a radius in which the exclusion of singular points of a Stewart platform is guaranteed. Another singular point exclusion algorithm has also recently been presented by (Merlet, 1997) using the determinant of the jacobian. By gridding the work space with points from which a radius can be calculated, one can preclude a larger working volume up to the whole work space (for limited stroke actuators) from singularities.

Convergence of the NR-iteration scheme can also be guaranteed in a neighbourhood of the solution if some conditions are satisfied. These conditions deal with the maximum and minimum gain of J and again the Lipschitz condition. These can be calculated for Stewart platforms. Exclusion of singular points of J is necessary to calculate a radius of the neighbourhood in which convergence is guaranteed. From this radius, the maximum gain of J and the maximum speed of the actuator, a sufficient update frequency of the iteration can be calculated above which (quadratic) convergence is guaranteed.

The parameters of the new Simona research simulator (Advani *et al.*, 1997) are used as an example which shows that reasonable results can be obtained although the conditions derived are rather conservative.

2 Notation

Capital symbols, X are used for matrices, \bar{x} for vectors, x for scalars. $\bar{x} \times \bar{y}$ denotes the vector product which can also be written as $\tilde{X}\bar{y} = (\tilde{Y})^T \bar{x}$ where \tilde{X} is a skew symmetric ($\tilde{X} = -\tilde{X}^T$) matrix parametrized by the vector, $\bar{x}^T = [x_1 \ x_2 \ x_3]$, such that the result of $\tilde{X}\bar{y}$ is the vector product:

$$\tilde{X} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \quad (3)$$

$X \times Y$ denotes vector wise product of the columns stacked in the matrices.

The index \bar{x}_n is used for the normalizing operation $\bar{x}_n = \bar{x} / |\bar{x}|$ with $|\bar{x}| = \sqrt{\bar{x}^T \bar{x}}$. $P_{\bar{x}_n}$ denotes the orthogonal projector to the (hyper)plane with normal vector \bar{x}_n and can be constructed from the vector product matrix $P_{\bar{x}_n} = (I - \bar{x}_n \bar{x}_n^T) = (\tilde{X}_n)^4 = \tilde{X}_n \tilde{X}_n^T = -(\tilde{X}_n)^2$. Projection matrices have some nice properties like $P = P^T = P^n$.

Motion can be described w.r.t. various frames. A matrix or vector described in some frame is, where appropriate, given a superscript referring to this frame. For the inertial frame or ground coordinates the index \bar{x}^g will be used. As a function of the moving end-effector or platform, vectors will be denoted \bar{x}^m . If a (rotation) matrix maps a vector into another frame it will be denoted as ${}^B R^A$ if R maps from A to B .

The subscript index like \bar{a}_i will be used to refer to the i^{th} -actuator if non actuator dependent variables also appear in the equation.

3 A general convergence theorem on NR-iteration

A weak version of the Newton-Kantorovich theorem (see e.g. Ortega and Rheinboldt, 1970) given in Stoer (1983) will be used. From this theorem convergence of the NR-iteration can be inferred.

It is stated as follows.

Theorem 3.1 *Given: a set $D \subseteq \mathbb{R}^n$, a convex set D_o with $\bar{D}_o \subseteq D$ and a function $\bar{f}: D \rightarrow \mathbb{R}^n$ which is continuous on D and differentiable with derivative $D\bar{f}(\bar{x})$ on D_o .*

If positive constants r, α, β, γ and h can be found for $\bar{x}_o \in D_o$ such that

$$S_r(\bar{x}_o) = \{\bar{x} \mid \|\bar{x} - \bar{x}_o\| < r\} \subseteq D_o \quad (4)$$

$$h = \frac{\alpha\beta\gamma}{2} < 1 \quad (5)$$

$$r = \frac{\alpha}{(1-h)} \quad (6)$$

and if \bar{f} has the following properties:

a)

$$\|D\bar{f}(\bar{x}) - D\bar{f}(\bar{y})\| \leq \gamma \|\bar{x} - \bar{y}\| \quad \forall \bar{x}, \bar{y} \in D_o \quad (7)$$

(This is called the Lipschitz condition)

b)

$$(D\bar{f}(\bar{x}))^{-1} \text{ exists and } \|D\bar{f}(\bar{x})^{-1}\| \leq \beta \quad \forall \bar{x} \in D_o \quad (8)$$

c)

$$\|D\bar{f}(\bar{x}_o)\bar{f}(\bar{x}_o)\| \leq \alpha \quad (9)$$

then

A) Starting at \bar{x}_o the sequence

$$\bar{x}_{k+1} = \bar{x}_k - (D\bar{f}(\bar{x}_k))^{-1}\bar{f}(\bar{x}_k) \quad k = 0, 1, \dots \text{ is well defined and } \bar{x}_k \in S_r(\bar{x}_o) \quad \forall k > 0$$

B) $\lim_{k \rightarrow \infty} \bar{x}_k = \bar{\xi}$ exists, $\bar{\xi} \in S_r(\bar{x}_o)$, and $\bar{f}(\bar{\xi}) = \bar{0}$

C)

$$\forall k \geq 0, \|\bar{x}_k - \bar{\xi}\| \leq \alpha \frac{h^{2k-1}}{1-h^{2k}} \quad (10)$$

With $0 < h < 1$ the iterates converge at least quadratically.

The proof of this theorem is given in Stoer (1983). Roughly speaking, this theorem states that a solution $\bar{\xi}$ can be found in the NR-iteration (B) if the differential $D\bar{f}(\bar{x})$ does not vary too much (a), is far enough from singularities (b) and eventually does not jump too close to the boundary of the defined neighbourhood, at the first iteration (c).

It also states that the iteration will not go out of a specified neighbourhood (A) and converges at a certain speed (C). To derive these conditions for a NR-iteration towards the Stewart platform coordinates, first its kinematics has to be specified.

4 Kinematics

First some general kinematics will be given. Then the kinematics of the Stewart platform will be stated.

4.1 Fundamental kinematics

The motion of a point (mass particle, joint, etc.) is usually most conveniently and invariantly defined w.r.t. the body frame to whom it is connected. The motion of a frame put in another frame generally consists of translation \bar{t} and rotation. The orientation of a frame can be described by a rotation matrix. A rotation matrix consists of perpendicular unit vectors which describe an orthogonal basis of

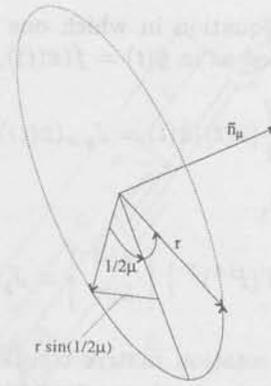


Fig. 1: Rotation parametrized by euler parameters

the frame in the other frame. As a result a rotation matrix, R has the following property:

$$R^T R = I \quad (11)$$

Any 3×3 -matrix with this property and $\det(R) = 1$ is a rotation matrix. With $\det(R) = -1$ also the mirror operation is included (transformation of right hand frames to left hand frames and vice versa). Given $T = {}^B R^A$, the position of a point \bar{p}^A in frame A can now be described in frame B by:

$$\bar{p}^B = \bar{t}^B + T\bar{p}^A \quad (12)$$

To describe the velocity of this point in the other frame one can simply differentiate this equation. Some properties of the time derivative of the rotation matrix can be derived by differentiating (11). This results in skew symmetric matrices which can be parametrized by the vector product matrix of the (thereby defined) angular velocity $\bar{\omega}$.

$$T^T \dot{T} = -\dot{T}^T T = \tilde{\Omega}^A \quad (13)$$

with

$$\tilde{\Omega}^A = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \quad (14)$$

Now with a vector which is rigidly attached to the frame A, $\dot{\bar{p}}^A = \bar{0}$,

$$\dot{\bar{p}}^B = \dot{\bar{t}}^B + T\tilde{\Omega}^A\bar{p}^A = \dot{\bar{t}}^B + \tilde{\Omega}^B\bar{p}^B \quad (15)$$

Where the change of frame for the matrix, $\tilde{\Omega}$, is given by $\tilde{\Omega}^B = T\tilde{\Omega}^A T^T$.

If some variations or velocities can be described as product of a (position-dependent) matrix and vector of other variations this matrix will be called a jacobian matrix.

The jacobian matrix between two sets of variables usually comes out naturally by a time differentiated

version of an equation in which one of the sets is explicitly stated as in $\bar{y}(t) = \bar{f}(\bar{x}(t))$,

$$\dot{\bar{y}}(t) = \frac{\partial \bar{f}}{\partial \bar{x}}(\bar{x}(t))\dot{\bar{x}}(t) = J_{y,x}(\bar{x}(t))\dot{\bar{x}}(t) \quad (16)$$

E.g. given (15),

$$\delta \bar{p}^B = [I \quad T(\hat{P}^A)^T] \begin{bmatrix} \delta \bar{t}^B \\ \delta \bar{\omega}^A \end{bmatrix} = J_{p^B, x} \begin{bmatrix} \delta \bar{t}^B \\ \delta \bar{\omega}^A \end{bmatrix} \quad (17)$$

Although the rotation matrix consists of nine entries, its properties put constraints on these entries. Different parametrizations such as euler angles (three subsequent planar rotations) or euler parameters (four parameters with one normalizing constraint to describe one axis of rotation and the angle of rotation) are possible.

The three euler angles have the disadvantage of a highly non-linear appearance in both the rotation matrix and the euler angle velocity to angular velocity transformation. The latter can even become singular.

It is possible to parametrize the rotation by the unit vector pointing along the axis of rotation \bar{n}_μ , and the angle μ of rotation. (See Fig. 1.) Parametrization by the four euler parameters $\bar{\epsilon} = [\epsilon_0 \quad \bar{\epsilon}_{13}^T]^T$, given by $\epsilon_0 = \cos(1/2 \mu)$ and $\bar{\epsilon}_{13} = \sin(1/2 \mu) \bar{n}_\mu$, results in very convenient (simple to calculate) relations of the rotation matrix and the angular velocity in which the euler parameters and its derivatives play an intermediate role. These relations are extensively dealt with in Nikravesh *et al.* (1985). Without further derivation they will be given here.

The rotation matrix can be calculated by taking a product of two matrices which are linear in $\bar{\epsilon}$:

$$R(\bar{\epsilon}) = G(\bar{\epsilon})L(\bar{\epsilon})^T \quad (18)$$

With

$$G(\bar{\epsilon}) = [-\bar{\epsilon}_{13} \quad \epsilon_0 I + \bar{\epsilon}_{13}] \quad (19)$$

and

$$L(\bar{\epsilon}) = [-\bar{\epsilon}_{13} \quad \epsilon_0 I + (\bar{\epsilon}_{13})^T] \quad (20)$$

$\dot{\bar{\epsilon}}$ can be described as a product of $\dot{\bar{\omega}}$ and a matrix which is a linear function of $\bar{\epsilon}$.

$$\dot{\bar{\epsilon}} = \frac{1}{2} G^T(\bar{\epsilon}) \dot{\bar{\omega}} \quad (21)$$

With angles $-\pi < 1/2 \mu < \pi$, $\bar{\epsilon}_{13}$ can be used as the (orientation) state from which $\epsilon_0 = \sqrt{1 - \bar{\epsilon}_{13}^T \bar{\epsilon}_{13}}$ is obtained.

4.2 Stewart platform kinematics

The Stewart platform (Fig. 2.) consists of an end-effector body whose coordinates can be described

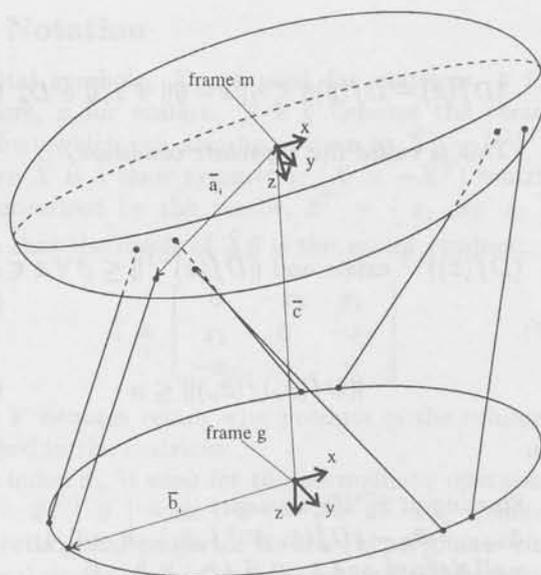


Fig. 2: A schematic example of the Stewart platform

by a body fixed point (e.g. the centre of gravity) which has varying coordinates \bar{c} in the inertial frame and the orientation given by a rotation matrix $T(\bar{\epsilon})$ which can be parametrized by euler parameters.

The end-effector body or platform is connected by six parallel actuators at \bar{a}_i to \bar{b}_i to the inertial frame. The length of the six actuators can be varied. In describing a specific actuator, the superscript i for the i^{th} actuator will be left away. An actuator (Fig. 3.) can be modelled as 2 bodies. A rotating body, b , with a constant distance of r_b of the c.o.g., bc , to the connection of a 2-d.o.f.-rotational gimbal joint to the inertial frame at \bar{b} . The moving actuator body, a , with a constant distance of r_a of the c.o.g., ac , is connected with a 3-d.o.f.-rotational gimbal joint to the platform at \bar{a} . With a 1-d.o.f. controlled sliding joint between these two bodies, the length of the actuator can be varied.

With this assumption also the case (often seen in practise) in which the moving part of the actuator both rotates and slides at the connection with the rotating part and has only 2-d.o.f. rotation w.r.t. the platform, results in the same dynamics.

The kinematics of the Stewart platform will be described by first defining the transformation of the platform pose to actuator coordinates. Then by differentiation also the velocity of all relevant points can be calculated as a function of the platform pose and its time derivatives.

Almost all vectors can be conveniently described in the inertial frame. Apart from \bar{a}_i^m whose time derivative in the moving frame is $\bar{0}$.

The vector, \bar{l}_i , between the two attachment points

of an actuator can be described by

$$\bar{l}_i = \bar{c} + T\bar{a}_i^m - \bar{b}_i \quad (22)$$

Now the length of the actuator, $|\bar{l}_i|^2 = \bar{l}_i^T \bar{l}_i$, and the unit vector in direction of the actuator, $\bar{l}_{n,i} = \frac{\bar{l}_i}{|\bar{l}_i|}$ can be calculated from the platform variables \bar{c}^g and the orientation matrix $T = {}^gR^m$ which will be the only rotation matrix used.

The velocity of the length of the actuators can be calculated by projection of the velocity of the upper gimbal attachment point, \bar{v}_a , in the direction of the actuator. Since $\frac{d}{dt} |\bar{l}| = \frac{d}{dt} \sqrt{\bar{l}^T \bar{l}} = \frac{\bar{l}^T \dot{\bar{l}}}{|\bar{l}|} = \bar{l}_{n,i}^T \dot{\bar{v}}_a$, and the velocity of the upper gimbal points is given by

$$\bar{v}_{a_i} = \dot{\bar{c}} + \bar{\omega} \times T\bar{a}_i^m, \quad (23)$$

the velocity of the actuator is given by

$$\dot{l}_i = \bar{l}_{n,i}^T \bar{v}_{a_i} = \bar{l}_{n,i}^T \dot{\bar{c}} + \bar{l}_{n,i}^T (\bar{\omega} \times T\bar{a}_i^m), \quad (24)$$

i.e. projecting upper gimbal velocity along the actuator direction. With some reordering and written as matrix equation (e.g. \bar{v}_{a_i} stacked in V_a) for all the actuators the jacobian between the actuator and platform velocities comes out.

$$\dot{\bar{l}} = L_n^T \dot{\bar{c}} + (TA^m \times L_n)^T \bar{\omega} = J_{l,x} \dot{\bar{x}} = L_n^T V_a \quad (25)$$

This jacobian matrix, $J_{l,x}(\bar{x})$, is one of the most important variables in kinematics and dynamics of the Stewart platform. The jacobian between platform and gimbal point velocity is defined by

$$\bar{v}_{a_i} = [I \quad T(\bar{A}_i^m)^T] \dot{\bar{x}} = J_{a_i,x} \dot{\bar{x}} \quad (26)$$

The derivative of the unit vectors $\bar{l}_{n,i}$ in the direction of each actuator can be calculated with:

$$\begin{aligned} \dot{\bar{l}}_n &= \frac{d}{dt} \frac{\bar{l}}{|\bar{l}|} = \frac{\dot{\bar{l}} |\bar{l}| - \bar{l} \frac{d}{dt} |\bar{l}|}{|\bar{l}|^2} \\ &= \frac{(I - \bar{l}_n \bar{l}_n^T)}{|\bar{l}|} \bar{v}_a = \frac{1}{|\bar{l}|} P_{l_n} \bar{v}_a \end{aligned} \quad (27)$$

5 NR-convergence to Stewart platform pose

Having stated the kinematical structure of the Stewart platform and given the general theorem on convergence of a NR-iteration, Theorem 3.1, it is now possible to investigate under what conditions the specific NR-iteration of (1) will converge to the physical platform pose.

To derive conditions for the NR-convergence to the right Stewart platform coordinates given the length of the actuators, first an appropriate definition of

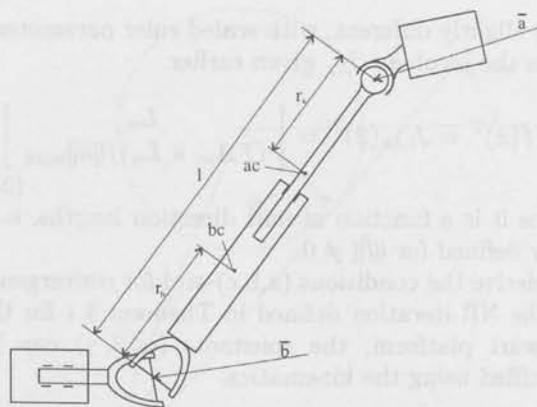


Fig. 3: Stewart platform actuator link construction

the coordinates has to be given. The defined jacobian $J_{l,x}$, is the description of platform translational and angular speed/variation to actuator length variations. To go from orientational parameter variations to angular speed is dependent on the parametrization used.

In this case the first three euler parameters are used to parametrize orientation.

$$\bar{x} = \begin{bmatrix} \bar{c} \\ s\bar{e} \end{bmatrix} = \begin{bmatrix} \bar{c} \\ 2\|\bar{a}\|_{max} \bar{e} \end{bmatrix} \quad (28)$$

where s is a scaling factor which can be used to get less conservative results in specifying a variation of \bar{x} since \bar{c} and \bar{e} have different dimension (m and rad). The scaling factor $s = 2\|\bar{a}\|_{max}$ will be shown to be appropriate in the sequel.

$$\bar{e} = \sin\left(\frac{1}{2}\mu\right) \bar{n}_\mu \quad (29)$$

It will be assumed that after each iteration \bar{e} will be reset to zero. In that case the relation between the angular velocity and the euler parameters is very simple.

$$\bar{\omega} |_{\bar{e}=0} = 2I\dot{\bar{e}} \quad (30)$$

Of course the rotation matrix in the iteration is now the multiplication of all the rotation matrices calculated.

$$T_{k+1} = T(\bar{e}_{k+1})T_k \quad (31)$$

The continuous function, \bar{f} in Theorem 3.1, from which the platform coordinates have to be found can now be given by

$$\|\bar{l}_i^*\| - \|\bar{l}_i\| = \bar{f}_i(\bar{x}) \quad (32)$$

where $\|\bar{l}_i^*\|$ is the measured length of the i^{th} actuator (fixed value for each iteration) and $\|\bar{l}_i\|$ is the length of the i^{th} actuator given \bar{x} (given by (22)). Since the measured length is fixed the derivative function is

only slightly different, with scaled euler parameters, from the jacobian $J_{l,x}$ given earlier.

$$D\bar{f}(\bar{x})^T = J_{l,sx}(\bar{x})^T = \left[\begin{array}{c} L_n \\ (TA_m \times L_n)/\|\bar{a}\|_{max} \end{array} \right] \quad (33)$$

Since it is a function of unit direction lengths, it is only defined for $\|\bar{l}\| \neq 0$.

To derive the conditions (a,b,c)-stpl for convergence of the NR-iteration defined in Theorem 3.1 for the Stewart platform, the constants (α, β, γ) can be specified using the kinematics.

a-stpl)

$$\|J_{l,sx}(\bar{x}_1) - J_{l,sx}(\bar{x}_2)\| < \gamma_{stpl} \|\bar{x}_1 - \bar{x}_2\| \quad (34)$$

$$\|J_{l,x}(\bar{x}_1) - J_{l,x}(\bar{x}_2)\| = \left\| \frac{dL_n}{d(TA_m \times L_n)/\|\bar{a}\|_{max}} \right\| = \bar{\sigma}(dJ) \quad (35)$$

Where in this case the 2-norm of the matrix is taken which is equal to the largest singular value, $\bar{\sigma}$. The Frobenius (semi)-norm is an upper bound on this norm and is advantageous in case of the Stewart platform since specific bounds can be derived on the matrix elements as will be shown later on.

$$\bar{\sigma}(dJ) \leq \|dJ\|_F \quad (36)$$

$$\|dJ\|_F = \sqrt{\sum_{i=1}^6 (\|d\bar{l}_{n,i}\|^2 + \|d(T\bar{a}_i \times \bar{l}_{n,i})/\|\bar{a}\|_{max}\|^2)}$$

To derive a constant γ_{stpl} , the separate elements of the last equation will be described as a function of \bar{x} . This will be done in the next section.

b-stpl) To state the second condition from which a constant number β_{stpl} has to be calculated also the 2-norm is used which can be upper bounded by one over the minimal singular value, σ_{min} , of the jacobian at some pose, \bar{x}_o . By using the maximal variation of the jacobian which has been calculated for the previous condition, the constant, β , becomes an upper bound for the maximum gain of the inverse jacobian over a volume of poses, \bar{x} .

$$\begin{aligned} \|J_{l,sx}^{-1}(\bar{x})\| &= \frac{1}{\sigma_{min}(J_{l,sx}(\bar{x}))} \\ &\leq \frac{1}{\sigma_{min}(J_{l,sx}(\bar{x}_o)) - \|dJ\|_F} \\ &= \beta_{stpl} \end{aligned} \quad (37)$$

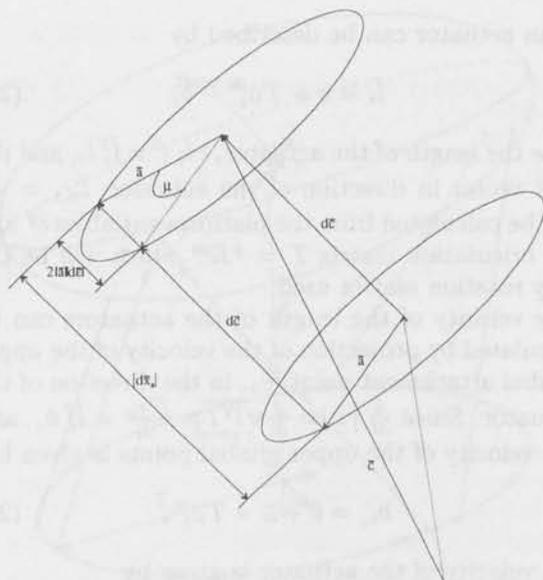


Fig. 4: Motion of connection \bar{x}_a as a function of change in \bar{x} ($d\bar{c}$ and $d\bar{e}$)

c-stpl) To calculate, α_{stpl} also use can be made of the singular values. Since any point in the work space is a possible initial start of the iteration, the maximum condition number of the scaled jacobian over the work space can be taken as the constant, α_{stpl} .

$$\begin{aligned} \|J_{l,sx}^{-1}(\bar{x})\bar{f}(\bar{x})\| &\leq \\ &\leq \frac{1}{\sigma_{min}(J_{l,sx}(\bar{x}))} \bar{\sigma}(J_{l,sx}(\bar{x})) \|d\bar{x}\| \\ &= \alpha_{stpl} \end{aligned} \quad (38)$$

In the next section it will be shown that indeed the jacobian of a Stewart platform is Lipschitz, as **a-stpl)** requires, as long as the actuators of the platform have minimal stroke strictly larger than 0.

6 Lipschitz condition on the Stewart platform jacobian

By constructively analyzing the kinematics of the Stewart platform the following lemma can be derived.

Lemma 6.1 The Stewart platform jacobian, $J_{l,sx}$, is lipschitz i.e. a γ can be found such that

$$\|J(\bar{x}_1) - J(\bar{x}_2)\| \leq \gamma \|\bar{x}_1 - \bar{x}_2\| \quad \forall \bar{x}_1, \bar{x}_2 \in D$$

iff

$$\|\bar{l}_i\| = \|\bar{f}_i(\bar{x})\| > \epsilon > 0 \quad \forall \bar{x} \in D, i \in \{1, \dots, 6\}$$

Note that the requirement of an actuator length larger than zero is an implicit constraint on the set of platform poses.

As the jacobian only consists of unit direction actuator vectors and a rotation which are a function of the platform pose, the maximum gain of the matrix is easily bounded. Remains to be shown that small variations of the platform pose do not result in relative large variations of the jacobian matrix gain. First the change of the unity actuator direction vector, $d\bar{l}_n$, as a function of the change in coordinates $d\bar{x}$ will be considered. Then also the change of the vector product $(\bar{l}_n \times T\bar{a})/\|\bar{a}\|_{max}$ can be bounded given a change $d\bar{x}$.

In two steps $d\bar{l}_n$ will be bounded.

1. Change of the upper gimbal connection $d\bar{x}_a$, given a change of platform coordinates $d\bar{x}$. See Fig. 4. in which the motion of the platform and connected gimbal point, \bar{x}_a , in Fig. 1. is schematically depicted.
2. Change of the unity direction, $d\bar{l}_n$, given a change the upper gimbal. See Fig. 5. in which the motion of actuator as a function of the moving gimbal point \bar{x}_a is schematically depicted.

The total motion of \bar{x}_a is maximal if both rotation and translation move this point in the same direction. One of the advantages of the euler parameter description now becomes apparent. Maximal motion as a function of the rotation is equal to two times the gain of the first three euler parameters times the distance from which point the rotation is considered. It is possible to bound the motion of the gimbal by

$$\|d\bar{x}_a\| \leq \|d\bar{c}\| + 2\|\bar{a}\|\|d\bar{\epsilon}\| \leq \sqrt{2}\|d\bar{x}\| \quad (39)$$

The second step takes into account that as the upper gimbal moves within a ball (Fig. 5.) the maximum change of the actuator direction is achieved if this line just touches the ball. In that case $(\bar{l}_n + d\bar{l}_n) \perp d\bar{x}_a$. With some geometry

$$\sin(\phi) = \frac{\|d\bar{x}_a\|}{\|\bar{l}\|} \quad (40)$$

$$\cos(\phi) = \sqrt{1 - \left(\frac{\|d\bar{x}_a\|}{\|\bar{l}\|}\right)^2} \quad (41)$$

the following monotonous upper bounding function can be derived for change of $d\bar{l}_n$.

$$\begin{aligned} \|d\bar{l}_n\| &\leq 2 \sin(1/2\phi) = \\ 2\sqrt{\frac{1 - \cos(\phi)}{2}} &= \sqrt{2} \sqrt{1 - \sqrt{1 - \left(\frac{\|d\bar{x}_a\|}{\|\bar{l}\|}\right)^2}} \end{aligned}$$

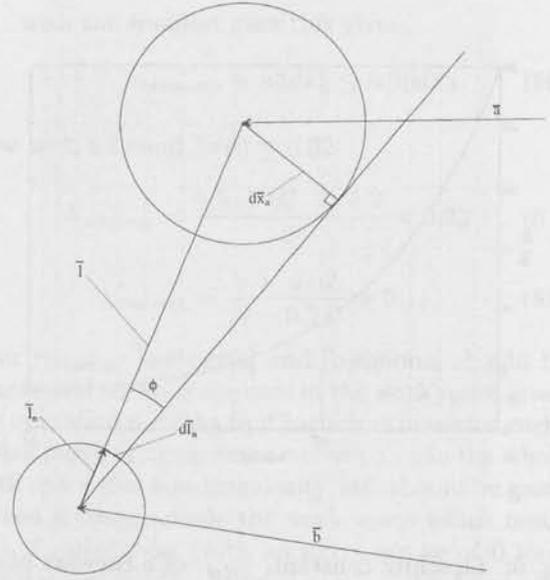


Fig. 5: Change of the vector \bar{l}_n as a function of motion in \bar{x}_a

$$\begin{aligned} &= \frac{\|d\bar{x}_a\|}{\|\bar{l}\|} + \mathcal{O}\left(\frac{\|d\bar{x}_a\|}{\|\bar{l}\|}\right)^3 \\ &\leq \sqrt{2} \frac{\|d\bar{x}_a\|}{\|\bar{l}\|} \end{aligned} \quad (42)$$

Taking into account (39) gives

$$\|d\bar{l}_n\| \leq \frac{2}{\|\bar{l}\|} \|d\bar{x}\| \quad (43)$$

So the length of the actuator should remain strictly larger than zero. It is easy showing that if this is not the case arbitrary small variations of the pose can result in nondecreasing variations of $d\bar{l}_n$ i.e. not satisfying the Lipschitz condition.

Now bounding the vector product is also possible. In general

$$\|\bar{a} \times \bar{b}\| \leq \|\bar{a}\|\|\bar{b}\| \quad (44)$$

$$(\bar{a} + \bar{c}) \times (\bar{b} + \bar{d}) = (\bar{a} \times \bar{b}) + (\bar{a} \times \bar{d}) + (\bar{c} \times \bar{b}) + (\bar{c} \times \bar{d}) \quad (45)$$

and rotation does not change the 2-norm.

$$\|T\bar{a}\| = \|\bar{a}\| \quad (46)$$

Change of the moving gimbal due to rotation is bounded by

$$\|d(T\bar{a})\| = 2\|\bar{a}\|\|d\bar{\epsilon}\| = \|ds\bar{\epsilon}\| \quad (47)$$

Now

$$\begin{aligned} \|d(\bar{l}_n \times T\bar{a})/\|\bar{a}\|_{max}\| &\leq \\ &\leq \|d\bar{l}_n \times T\bar{a}_n\| + \|\bar{l}_n \times d(T\bar{a}_n)\| \\ &\quad + \|d\bar{l}_n \times d(T\bar{a}_n)\| \\ &\leq \frac{\sqrt{2}}{\|\bar{l}\|} \|d\bar{x}_a\| + \|ds\bar{\epsilon}\| + \frac{2\|ds\bar{\epsilon}\|\|d\bar{x}\|}{\|\bar{l}\|} \end{aligned}$$

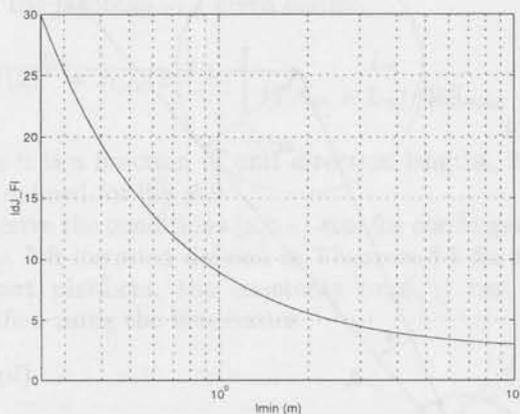


Fig. 6: Lipschitz constant, γ_{stpl} , of a Stewart platform as a function of the minimal actuator length

So in this way every element of the jacobian matrix is explicitly bounded by the variation of the platform pose and the (implicitly platform pose dependent) actuator length. If a minimal actuator length is given a specific γ_{stpl} can be calculated.

Adding the bounds for each of the actuators in the F-norm of the jacobian results in an explicit number for γ_{stpl} . Assuming small platform pose variations i.e. second order effects in the last equation are relatively small e.g. $\|d\bar{x}\| < .25$, a minimal actuator length of $2m$ of the Simona platform gives

$$\|dJ\|_F \leq \sqrt{\sum_{i=1}^6 1^2 + \sum_{i=1}^6 (1.82 + .07)^2} \|d\bar{x}\| \leq 5.2 \|d\bar{x}\|$$

Note that this γ_{stpl} is valid for any Stewart platform having minimal actuator length of $2m$ independent of gimbal point coordinates. As a function of the minimal actuator length, γ_{stpl} is depicted in Fig. 6. The limited difference of two jacobians given limited difference of the platform pose can be used to derive guaranteed convergence of the NR-iteration but also to exclude singular points of the jacobian from part of the platform working space as will be shown in the next Section.

7 Exclusion of singular points

Singular points of a Stewart platform are those platform poses at which the jacobian, $J_{l,sx}$ becomes singular. At these points at least one platform pose variation will not result in actuator length variation and is therefore not supported by the actuators or

uncontrollable from forces along the actuator directions. These points exist in the usual Stewart platform configurations if the actuators would not have length constraints (Ma and Angeles, 1991).

Given a point \bar{x}_o it is possible to calculate the minimum gain (singular value) of the jacobian. This value does not change more than the maximal variation of the jacobian. To be singular this value should be zero. So with

$$\begin{aligned} \sigma_{min}(J_{l,sx}(\bar{x})) &\leq \\ &\leq \max(0, \sigma_{min}(J_{l,sx}(\bar{x}_o)) - \bar{\sigma}(dJ(\bar{x}, \bar{x}_o)), \end{aligned}$$

and

$$\|dJ\| < \gamma \|\bar{x} - \bar{x}_o\|, \quad (48)$$

$J_{l,sx}^{-1}(\bar{x})$ exists in the ball around \bar{x}_o in platform coordinate space with radius

$$r_s \geq \frac{\sigma_{min}(J(\bar{x}_o))}{\gamma} \quad (49)$$

By calculating r_s over "boxes" $d\bar{x} = \sqrt{2}r_{s,min}$ (see Fig. 7) a grid is taken which precludes the whole work space from singular points of the jacobian.

Algorithm 7.1

- Choose an expected minimal singular value of the jacobians, over the grid points.
- Take a grid such that the boxes with radius $r_{s,min}$ around the grid points fill the whole work space.
- Calculate the minimal singular value of the jacobian at every grid point.
- If the minimal singular value is larger than the expected value, there are no singular points in the work space, the algorithm finishes. If not, start another iteration choosing a smaller expected minimal singular value.

Lemma 7.1 *If the work space of a robotic manipulator having a bounded Lipschitz constant does not have any singular point this will be detected by Algorithm 7.1 in a finite number of iterations.*

Off course the boundary of the work space (in six dimensions!) should be known which is a stand alone problem (treated in Luh *et al.*, 1996). To calculate an upper bound for the gain of the inverse jacobian (β_{stpl}) a finer grid should be taken. (This will increase calculation time tremendously, e.g. n^6 grid points extra.)

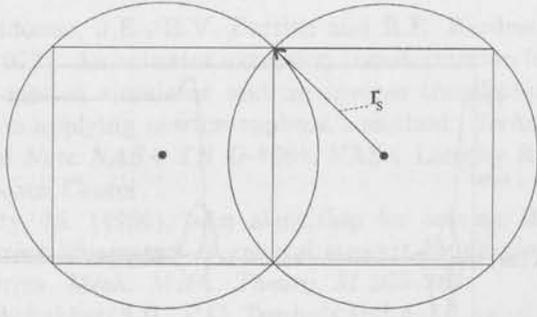


Fig. 7: By gridding the six dimensional space it is possible to exclude the whole work space from singular points

variable	value	description
r_a	1.60 m	upper gimbal radius
r_b	1.65 m	lower gimbal radius
d_a	0.20 m	upper gimbal spacing
d_b	0.60 m	lower gimbal spacing
l_{min}	2.08 m	minimal actuator length
s	1.25 m	actuator stroke
s_{op}	1.15 m	operational actuator stroke
\dot{l}_{max}	1 m/s	maximum actuator speed

Table 1: The Simona motion system parameters

8 Sufficient update frequency of the Simona simulator

The bounds derived are rather conservative in most cases. By calculating a convergence radius for the NR-iteration in the practical example of the new Simona research simulator shows that it is possible to guarantee that this iteration can be used if a reasonable frequency is used to update the platform pose. Further the singular points can really be precluded from the work space in this case.

Some platform parameters can be found in Table 1.

To satisfy NR-iteration convergence the following constants were obtained.

a-simona With the minimal length of the actuator $l_{min} = 2.08m$, $\gamma_{simona} = 5.2$

b-simona With gridding a minimal radius of $r_{s,min} = .09$ is found. (About 200000 grid points had to be calculated to exclude the work space). The smallest singular value found is 0.75 and with finer gridding a $\beta_{simona} = 2$ can be guaranteed.

c-simona The first value of $\bar{f}(\bar{x}_o)$ in any point can not be larger than $\bar{f}(\bar{x}_o) \leq \bar{\sigma}(J)\|d\bar{x}\|$ Together

with the smallest gain this gives

$$\alpha_{simona} = \kappa\|d\bar{x}\| \leq 3.5\|d\bar{x}\| \quad (50)$$

Now with a bound $\|d\bar{x}\| \leq 0.02$

$$h_{simona} = \frac{3.5 \cdot 0.02 \cdot 2 \cdot 5.2}{2} = 0.23 \quad (51)$$

$$r_{simona} = \frac{3.5 \cdot 0.02}{(1 - 0.23)} = 0.11 \quad (52)$$

Over r_{simona} , [a-simona] and [b-simona] should be guaranteed which is the case in the work space given the operational stroke (not including actuator cushioning part). To guarantee convergence in the whole work space also non-singularity, etc. should be guaranteed further outside the work space which needs lots of calculation (with an extra stroke of 0.15m, singularities can be obtained so $r_{s,min}$ becomes very small).

With a bound on the maximal speed of the actuator it is possible to calculate a minimal update frequency which guarantees $\|d\bar{x}\| < 0.02$.

$$\|\dot{l}\| < 1 \text{ m/s} \quad (53)$$

$$\beta_{simona} \Delta t \dot{l}_{max} < \|d\bar{x}\| \quad (54)$$

$$\Delta t < 1/100 \text{ s} \quad (55)$$

So with an update frequency $f_s > 100$ Hz convergence of the NR-iteration is attained.

Lemma 8.1 *The Stewart platform with the Simona motion system parameters has no singular points in the work space and the NR-iteration with this platform will converge to the right platform pose if the update frequency is larger than 100 Hz.*

9 Application within the control of the motion system

With the dynamic model of the motion system of the Stewart platform (Koekebakker, 1996) a standard feedback linearising control can be constructed.

$$\bar{f}_p = M(\bar{x})(\ddot{\bar{x}} + \epsilon_r) + C(\dot{\bar{x}}, \bar{x}) + G(\bar{x}) \quad (56)$$

An outer loop typically looks as follows:

$$\epsilon_r = K_d(\ddot{\bar{x}}_r - \ddot{\bar{x}}) + K_p(\dot{\bar{x}}_r - \dot{\bar{x}}) \quad (57)$$

The platform force, \bar{f}_p , has to be translated to equivalent actuator forces, \bar{f}_a with the jacobian.

$$\bar{f}_a = J^{-1}(\bar{x})^T \bar{f}_p \quad (58)$$

The platform coordinates are assumed to be known and as only the actuator lengths are measured, a

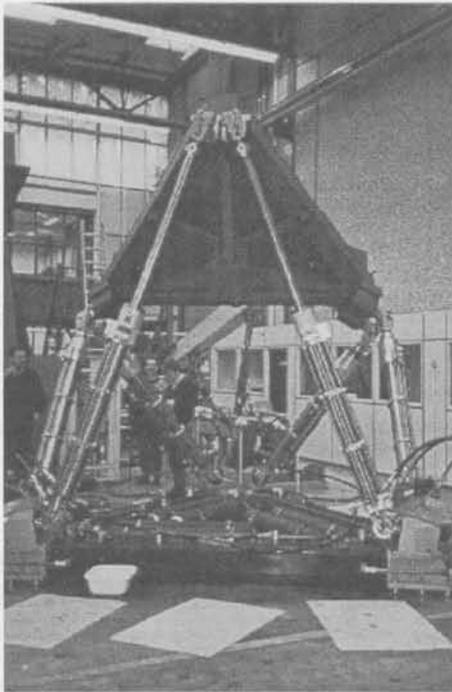


Fig. 8: Simona motion system with dummy platform

scheme like the Newton Raphson iteration scheme ((1)) has to be used. As the iteration has to run as part of the feedback loop, the iteration has to converge to the (local) solution at least to have stability of the loop. In the previous sections it was proved that this is the case with the Simona flight simulator motion system if the iteration is run at at least 100 Hz.

As there is limited time and calculation speed of the digital controller, a limited number of iterations is possible. With a limited number of iterations in one sample of time, the scheme has to show convergence as the system moves with limited speed. Iterations have to be limited as much as possible as inversion of the jacobian takes relatively much computer processor time.

This problem was solved as follows. Two errors forming the total difference between the actual and latest estimate of platform coordinates, now have to be considered. First the error which remains after a limited number of iterations in the previous sample and secondly the error introduced with the motion system which moved during the last sample.

W.r.t. control the error on the estimate preferably should not be larger than the measurement error. This error is ca. .1 mm. With the model based controller running on 1 kHz, the α_{simona} ('first iteration error') becomes quite small as does h. With (C) from the theorem the speed of convergence can be

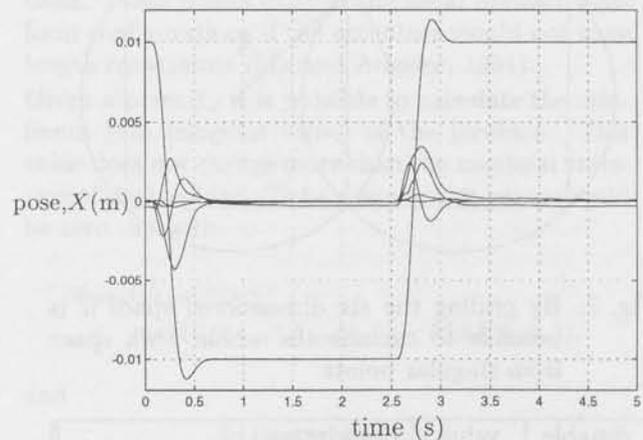


Fig. 9: Step in forward (surge) direction and parasitic pitch motion other states (10 times amplified) calculated on-line with NR-iteration

calculated. Already at two iterations platform coordinates will be calculated up to an error smaller than 0.01 mm.

The scheme was implemented on the Simona motion system currently running with a dummy platform of 2200 kg (see Fig. 8). A step response in surge direction (see Fig. 9) shows that the platform coordinates which are controlled upon, can be monitored on-line.

10 Conclusion

To satisfy a general convergence theorem on Newton-Raphson iteration one of the requirements is the Lipschitz condition on the derivative function which was shown to be satisfied for Stewart platforms. Variations of the jacobian can be bounded by the variations of the platform coordinates.

Next to this requirement the jacobian should not be singular in any point of the work space. With the Lipschitz condition on the jacobian and gridding, volumes can be excluded from singularities.

Although only sufficient (and thereby conservative) conditions could be derived for convergence, it is possible to obtain a result in the practical example of a new built simulator which makes model based feedback possible.

References

- Advani, S.K., E. van Zuylen and S. Bettendorf (1997). Simona - a reconfigurable and versatile research facility. *Proc. AIAA Modeling & Simulation Technologies Conference*.

Dieudonne, J.E., R.V. Parrish and R.E. Bardusch (1972). An actuator extension transformation for a motion simulator and an inverse transformation applying newton-raphson's method. *Technical Note NASA TN D-7067*, NASA Langley Research Center.

Husty, M. (1996). An algorithm for solving the direct kinematics of general stewart-gough platforms. *Mech. Mach. Theory*, 31 365-380.

Koekebakker, S.H., P.C. Teerhuis and A.J.J. van der Weiden (1996). Alternative parametrization in modelling and analysis of a stewart platform. In Bosgra et al. (Eds.), *Sel. Topics in Identif., Modelling and Control*, 9. Delft University Press, pp. 59-68.

Liu, K., M. Fitzgerald, D.M. Dawson and F.L. Lewis (1991). Modeling and control of a stewart platform manipulator. In N. Sadegh and Y.h. Chen (Eds.), *Control of Systems with Inexact Dynamic Models*. American Society of Mechanical Engineers, New York, pp. 83-89. Presented at Winter Annual Meeting ASME, Atlanta, GA, USA.

Luh, C.M., F.A. Adkins, E.J. Haug and C.C. Qiu (1996). Working capability analysis of stewart platforms. *Trans. ASME: Journal of Mechanical Design*, 220-227.

Ma, O. and J. Angeles (1991). Architecture singularities of platform manipulators. *Proc. IEEE Intern. Conf. Robotics and Automation*, pp. 1542-1547.

Merlet, J.P. (1997). Determination of the presence of singularities in a workspace volume of a parallel manipulator. In *Computational methods in mechanisms; Workshop of the NATO Advanced study institute*, pp. 289-295.

Nijmeijer, H. and A.J. van der Schaft (1990). *Non-linear dynamical control systems*. Springer Verlag, Berlin.

Nikravesh, P.E., R.A. Wehage and O.K. Kwon (1985). Euler parameters in computational kinematics and dynamics. *Trans. of the ASME*, 107, 358-369.

Ortega, J.M. and W.C. Rheinboldt (1970). Iterative solution of non-linear equations in several variables. In *Computer Science and Applied Mathematics*. Academic press NY-London.

Stewart, D. (1965). A platform with six degrees of freedom. *Proc. Inst. Mech. Engrs.*, pp. 371-386.

Stoer, J. (1983). *Einführung in die Numerische Mathematik 1*. Springer-Verlag, Berlin.

Suppressing non-periodically repeating disturbances in mechanical servo systems

Rob L. Tousain[†], Jean-Christophe Boissy[§], Meindert L. Norg[‡],
Maarten Steinbuch[‡] and Okko H. Bosgra[†]

[†] *Mechanical Engineering Systems and Control Group
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.
E-mail: r.l.tousain@wbmt.tudelft.nl*

[§] *Philips Electronics N.V., BP 15 94453, Limeil Brevannes, Cedex, France*

[‡] *Philips Electronics N.V., P.O. Box 218, 5600 MD Eindhoven, The Netherlands*

Abstract. A rather specific class of non-stationary disturbances that can occur in mechanical servo systems is that of the non-periodically repeating (*NPR*) disturbances. For a linear control system to deal specifically with this class of disturbances, it must exhibit the same non-periodically repeating behavior. A way to provide the control system with this behavior is by including a *NPR*-disturbance detector and a lookup-table-based feedforward mechanism. Assuming the shape of the *NPR*-disturbance to be unknown a priori, a learning algorithm for the construction of the lookup table must be included as well. Based on the Likelihood Ratio Test, a suitable detector can be derived. Existing theory on Iterative Learning Control (ILC) can be used to design a suitable algorithm for the construction of the lookup table. A realistic simulation example, the shock suppression in a Hard Disk Drive, illustrates the validity of the algorithms.

Keywords. Non-periodically repeating disturbances, feedforward control, detection, iterative learning control, shock suppression, hard disk drive.

1 Introduction

In many control systems, performance is directly related to the amount of disturbance compensation. It is therefore not surprising that many control design techniques take explicitly into account the nature of the disturbances. Examples are LQG control, where some knowledge on the disturbances can be incorporated in the control law, and H_∞ control, where knowledge on the disturbance spectra can be incorporated in the choice of suitable weightings to be included in the "standard plant". Using such techniques, controllers can be designed to deal in a non-conservative way with *stationary* disturbances, which are in their behavior considered as stochastic processes. Non-stationary disturbances can in general not be dealt with. A rather specific class of non-stationary disturbances is that of the *non-periodically repeating (NPR) disturbances*.

With this class, we mean deterministic disturbances with a fixed shape (or a shape that changes slowly in time) that occur randomly in time. Examples of such disturbances are vibrations and shocks in mechanical servo systems, such as Hard Disk Drives and photolithography systems (Yasuda, 1996). Figure 1 shows a representative sequence of 3 realizations of an arbitrary *NPR*-disturbance, d_{NPR} .

Often, standard feedback control will not yield a satisfactory compensation, since the feedback controllers are designed to perform well under nominal conditions¹ in the first place. Adapting the feedback control law in order to improve the suppression of the randomly occurring disturbance will in most cases deteriorate the performance of the control system under nominal conditions, due to well known limitations of linear feedback control such as

¹i.e. absence of the randomly occurring disturbance



Fig. 1: Example of a *NPR*-disturbance

the Bode Sensitivity integral. The need for a different approach to the problem of rejecting *NPR*-disturbances was recognized in e.g. (Yasuda, 1996) and (Simaan, 1990). The *XY*-table control scheme presented in (Yasuda, 1996) inhibits an off-line determined pseudo feedforward disturbance compensation which is activated "at a proper time". The feedforward vector is calculated off-line using measurements of the error that is caused by the *NPR*-disturbance. The drawback of the algorithm is that for every new situation, a new feedforward compensation has to be calculated off-line. Further, it was not clear from the paper how to determine the "proper time" for activating the feedforward. In the process-control-oriented paper (Simaan, 1990), the necessity of detecting *NPR*-disturbances was depicted and a very brief description of a possible detector was included. The question what corrective action should be taken at the moment of detection was not addressed by the authors.

In this paper we present a general control algorithm² for compensation of a particular class of *NPR*-disturbances. The algorithm we propose inhibits a model-based detection and a feedforward compensator that is determined on-line in an adaptive manner. The basic idea is schematically represented in the configuration of Figure 2, where the proposed algorithm is represented as an add-on device. In the block scheme, *P* is the plant, *C* an arbitrary feedback controller, *r* the reference signal, *y* the output, and *e* the error that results from subtracting *y* from *e*. *d* is a stationary disturbance, that perturbs the plant's nominal operation. The main elements in the add-on device are the detector and the mechanism for constructing a suitable feedforward signal, *u_f*, and storing it in a lookup table. The design of both elements is addressed in this paper. The remainder of this paper is organized as follows. First, in Section 2, some restrictions on the

²For the proposed algorithm, a patent request by PHILIPS ELECTRONICS N.V. is running. Official filing number: 97 04939. First filing date: april 22nd.

class of *NPR*-disturbances, derived from the nature of the proposed algorithm, will be pointed out. In Section 3, the solution to the detection problem will be presented. In Section 4, we propose a method for constructing the lookup table on-line. Section 5 illustrates the usefulness of the method with a simulation example considering shock suppression in a Hard Disk Drive. Section 6 ends up with some conclusions.

2 Restrictions on the disturbance class

In this section we will specify more precisely the class of disturbances that is under consideration in this paper. Let us first introduce a *NPR*-disturbance as follows:

$$d_{NPR}(t) = \begin{cases} d_r(t - t_j), & t_j \leq t < (t_j + T) \\ 0, & \text{for all other } t, \end{cases} \quad (1)$$

where *d_r* is a deterministic disturbance of length *T* and *t_j*, *j* = 1, 2, ... are the time instants at which this fixed shape disturbance occurs, randomly distributed according to some probability distribution. For the proposed compensation method to be feasible, *d_r* must be *detectable*. Further, for the feedforward compensation to make sense, the detection time should be fairly smaller than *T*. Besides these demands on the detectability of *d_r*, the rate of change of *d_r* is limited. This can be seen as follows. Due to the stochastic noise disturbance, the detection time will vary along an average number of samples. The effect of a deviation in detection time is a time-shift of the feedforward signal. To evaluate the effect of such a time-shift on the compensation performance, we allow ourselves here to speak of the "frequency content" of the feedforward signal. Suppose that this frequency content is bounded by a frequency *f_{max}* (*Hz*). Then, a deviation in the detection time of one sample will introduce a phase error in the feedforward signal of *T_s* × 2π*f_{max}* for the highest-frequency part of the feedforward, where *T_s* is the sample time. For phase shifts larger than ½π *rad*, the effect of the feedforward compensation may possibly be an error increase instead of decrease. This gives us a *rule of thumb* for the maximum "frequency range" of the feedforward signal,

$$f_{max} \ll \frac{1}{4 \times T_s \times n_{dev}}, \quad (2)$$

where *n_{dev}* is the largest deviation from the mean detection time in samples. Summarizing, we limit the class of disturbances under consideration in this paper to those *NPR*-disturbances that are

- "sufficiently fast" detectable,
- limited in their frequency range by *f_{max}*. (2)

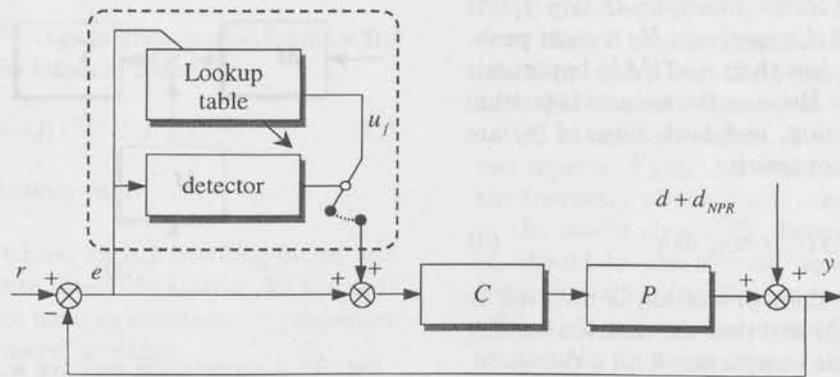


Fig. 2: Schematical representation of the HDD control system with the detector/feedforward algorithm.

3 Detection

The detection problem is to determine as fast as possible the presence of d_r from measurements of the error with a certain desired reliability (quantified by the chances on *false alarms* and *misses*). To deal with this problem in a realistic, though not overly conservative way, we assume here that the shape of the disturbance is known. We will show later that the invalidness of this assumption does not significantly limit the applicability of the detection algorithm. Note that, thanks to linearity, the problem of detecting the presence of d_r is equivalent to detecting the presence of y_r in y , where y_r is the effect of d_r in the output of the system, i.e. $y_r = (1 + PC)^{-1}d_r$. For convenience, we assume here that all disturbances other than d_{NPR} are stochastic and we combine them into one output-equivalent disturbance, d , at the output of the plant. The detection problem thus involves the detection of a known waveform in colored noise, where the known waveform is y_r , and the colored noise is the output of the system due to the stationary stochastic disturbance d . The solution to such problems can be found in literature on communication systems, e.g. (Trees, 1968) and (Poor, 1987), where signal detection and estimation problems are very common. In control literature, the idea of model-based detection of randomly occurring disturbances was posed qualitatively in e.g. (Simaan, 1990). Here, we will derive such a detection algorithm quantitatively.

The Likelihood Ratio Test

For the detection problem, we introduce two hypotheses: H_0 and H_1 , relating to respectively *absence* and *presence* of d_r . Let our detector have an observation window of N samples, then under the two hypotheses the observations of y in a certain

time interval $[t_j, t_j + N]$ are given by:

$$H_0 : y(t_j + i) = y_n(t_j + i); \quad (3)$$

$$H_1 : y(t_j + i) = y_r(i) + y_n(t_j + i), \quad (4)$$

$i = [1, N]$. The Gaussian colored noise stochastic process y_n is defined by

$$y_n = \frac{1}{1 + PC}d. \quad (5)$$

Introducing H_d as a realization filter for d , (5) can be rewritten as

$$y_n = \frac{1}{1 + PC}H_d v, \quad (6)$$

where v is a white noise stochastic process with variance λ . For convenience, we define

$M = (1 + PC)^{-1}H_d$, so that (6) can be written as

$$y_n = Mv. \quad (7)$$

We denote M the *coloring filter* for the colored noise stochastic process y_n .

Attached to the two hypotheses are the two joint probability densities, $py_{(t_j)|H_1}(y(t_j)|H_1)$ and $py_{(t_j)|H_0}(y(t_j)|H_0)$, which define the probability of respectively H_0 and H_1 , given the actual observation vector, $\mathbf{y}(t_j) = [y(t_j+1) \cdots y(t_j+i) \cdots y(t_j+N)]'$.

From e.g. (Trees, 1968) it is known that a suitable decision rule for determining whether H_0 or H_1 is most probable, is the *Likelihood Ratio Test* (LRT). It is defined as:

$$\Lambda(\mathbf{y}(t_j)) = \frac{py_{(t_j)|H_1}(\mathbf{y}(t_j)|H_1)}{py_{(t_j)|H_0}(\mathbf{y}(t_j)|H_0)} \underset{< H_0}{\overset{> H_1}{>}} \eta, \quad (8)$$

where η is a threshold whose value is dependent on the decision criteria that are defined, and $\Lambda(\mathbf{y}(t_j))$ is

called the *Likelihood Ratio*. In words: IF $\Lambda(\mathbf{y}(t_j))$ is greater than η , THAN hypothesis H_1 is most probable, IF $\Lambda(\mathbf{y}(t_j))$ is less than η , THAN hypothesis H_0 is most probable. Because the natural logarithm is a monotonic function, and both sides of (8) are positive, an equivalent test is:

$$\ln \Lambda(\mathbf{y}(t_j)) \underset{H_0}{\overset{H_1}{>}} \ln \eta. \quad (9)$$

We see that all the data processing is involved in computing $\ln \Lambda(\mathbf{y}(t_j))$ and that the decision making mechanism is a simple comparison with a threshold. The derivation of the LRT is given in the Appendix. The result is:

$$\sum_{i=1}^N M^{-1}(q)y(t_j + i)M^{-1}(q)y_r(i) \underset{H_0}{\overset{H_1}{>}} TH, \quad (10)$$

where

$$TH = \frac{1}{2} \sum_{i=1}^N (M^{-1}(q)y_r(i))^2 + \lambda \ln \eta \quad (11)$$

is the threshold.

The LRT has the structure of a correlation receiver with additional *whitening filter* (M^{-1}), as is depicted in Figure 3: the observation samples are filtered through M^{-1} , multiplied by the samples of the filtered y_r and the sum of these products added over the observation interval $[1, N]$ is compared with a threshold value to either decide or deny presence of d_r . The correlation receiver can simply be implemented as a discrete time filter.

Robustness against uncertainty in y_r

The output of the *detection filter*, denoted by the left-hand side of (10), is a Gaussian variable. It can easily be shown that in the case the first N samples of y_r are exactly modeled, at time $(t_j + N)$, expectation and covariance are given by respectively $\sum_{i=1}^N (M^{-1}(q)y_r(i))^2$ and $\lambda \sum_{i=1}^N (M^{-1}(q)y_r(i))^2$. Now assume that the actual first N samples of the disturbance are given by $(y_r(i) + \Delta_y(i))$, $i = [1, N]$, where Δ_y is any structured or unstructured uncertainty, and the detector is designed under assumption that y_r is exact. The expectation at time $(t_j + N)$ becomes $\sum_{i=1}^N (M^{-1}(q)y_r(i))^2 + \sum_{i=1}^N (M^{-1}(q)y_r(i))(M^{-1}(q)\Delta_y(i))$, while the covariance remains unaltered. Since the output of the detection filter at time $(t_j + N)$ is a Gaussian variable, the chance on detection at this time instant is solely dependent on the expectation and covariance of the output. The detector is thus robust against varying shape of d_r in the sense that the chance on detection at time $(t_j + N)$ remains unaltered when

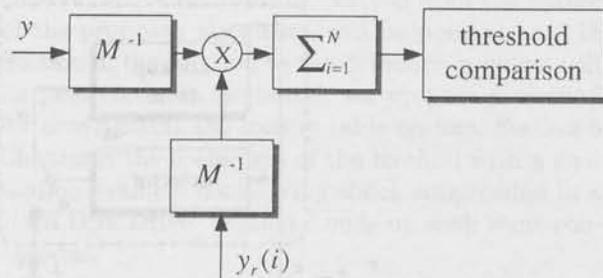


Fig. 3: A correlation receiver with whitening filter

$\sum_{i=1}^N (M^{-1}(q)y_r(i))(M^{-1}(q)\Delta_y(i)) = 0$. When this sum takes values larger than zero, the chance on detection will be larger, for values less than zero it will be smaller.

The real performance of the detector is of course specified by the probability distribution of the detection time, instead of the chance on detection at one time instant. It is to be expected that this distribution shows a larger dependency on the shape of d_r than the performance measure we used, the chance on detection at time $(t_j + N)$. However, for simplicity, we restricted ourselves to the robustness analysis presented above.

4 Construction of the lookup table

For the construction of a suitable feedforward signal we utilized well-known existing techniques of Iterative Learning Control (ILC) and Repetitive Control. ILC has its primary application in the construction of optimal feedforward signals for setpoint trajectory tracking (for a literature survey on the subject, we refer to (Chang, 1997)). Each learning step, the feedforward signal is updated using the most recent error information and stored in a memory. In close resemblance to the principle of learning control, we propose the following on-line procedure for the construction of a suitable lookup table

step 1

each $(l)^{th}$ time the detector outputs H_1 , simultaneously, two operations are initiated:

- a feedforward signal $F^{(l-1)}$ of length V is read from the lookup table and multiplied by a factor k_1 before it is added to the input of controller C :

$$u_f = k_1 F^{(l-1)}; \quad (12)$$

- a vector of the error signal, $E^{(l)}$, of length V is stored in a memory device;

step 2

at the moment $E^{(l)}$ is available, a new feedforward vector, $F^{(l)}$, is calculated as follows:

$$F^{(l)} = Q(F^{(l-1)} + BE^{(l)}), \quad (13)$$

and saved in the lookup table.

In the above procedure, k_1 is a learning factor and Q and B are discrete time filters which for the purpose of convergence have to satisfy some properties which will be discussed hereafter.

Choice of Q , B and k_1

The choice of the so-called *learning filter* B and *robustness filter* Q follows from convergence considerations. For the general update law in learning control (13), literature (e.g. (de Roover, 1996)) provides us the following convergence result (slightly adapted to our case):

Theorem 4.1 Suppose $F, E \in L_2(0, \infty)$, and $F^{(l)}$ is synchronized with d_r for all $l \in \mathbb{N}^+$, then the learning iteration (13) converges to a fixed point $F^* = \lim_{l \rightarrow \infty} F^{(l)}$ if

$$\|Q(1 - BR)\|_{\infty} < 1, \quad (14)$$

where $R = (1 + PC)^{-1}PC$.

For the proof of this theorem, we refer to the relative paper. The following theorem, which is also taken from (de Roover, 1996) teaches us about the quality of the convergence.

Theorem 4.2 Suppose $d_r \neq 0$, and $F^{(l)}$ is synchronized with d_r for all $l \in \mathbb{N}^+$, then the learning iteration (13) converges to a fixed point $E^* = \lim_{l \rightarrow \infty} E^{(l)} = 0$, if and only if (14) is true and $Q = 1$.

Unfortunately, due to the deviation in detection time, $F^{(l)}$ can not be synchronized with d_r for all $l \in \mathbb{N}^+$. Accordingly, Theorems 4.1 and 4.2 do not apply to our learning disturbance compensator. However, we can still use them as rules of thumb for determining suitable filters Q and B . Hence, the following design procedure, taken from the literature on learning control (de Roover, 1996), seems to make perfectly sense:

1. choose $B(e^{i\omega}) \approx R^{-1}(i\omega)$, $\omega \in [0, \omega_c]$, i.e. choose B to be the best possible (approximate) inverse of R , up to some frequency ω_c ;
2. choose $Q(e^{i\omega})$ to be a low-pass filter $\|Q(e^{i\omega})\| = 1$, $\forall \omega \in [0, \omega_c]$, and $\|Q(e^{i\omega})\| = 0$, $\forall \omega > \omega_c$.

Note that the choice of B is intuitively right: the error should be processed through the inverse of the complementary sensitivity function, in order to obtain the reference signal that yields exact compensation of the disturbance. The choice of ω_c contains two aspects. First, ω_c should be lower or equal than the frequency above which considerable uncertainty in the model shows up. Second, by our intuition, ω_c should be considerably lower than $2\pi f_{max}$ (2). Note that since the filter operations are performed off-line, Q may be a non-causal filter which enables us to make its phase zero.

Removing initial error conditions

Probably, when **step 1** of the learning algorithm is initiated, the error will have a nonzero value. As a consequence, the first sample of $E^{(l)}$ will be nonzero for most l . If this is the case, the first sample of $F^{(l)}$ will increase unceasingly with increasing l . Note that this is in fact the result of our unreasonable demand on the error to diminish step-like from a certain non-zero value to zero. To avoid the continuous increase of the first sample of $F^{(l)}$, we propose the following adaptation to the standard procedure described above (13): use as an input to the learning algorithm not the vector $E^{(l)}$, but the vector $E_s^{(l)}$, whose samples are calculated as:

$$E_s^{(l)}(i) = \begin{cases} E^{(l)}(i) - E^{(l)}(1)(W+1-i)/W, & i = 1, 2, \dots, W \\ E^{(l)}(i), & i = W+1, \dots, V. \end{cases} \quad (15)$$

$E_s^{(l)}$ is thus the error signal that we find if we require the output to converge gradually to the setpoint in W samples instead of instantaneously.

5 Shock suppression in a Hard Disk Drive

In this section, we apply the algorithms described in the previous section to the problem of shock suppression in a Hard Disk Drive (HDD). The objective of HDD tracking control is to maintain the read/write head as close as possible to the center of the track despite of disturbances acting on the plant. The most significant disturbances are *output disturbances*, such as eccentricity of the disks, and *shocks*. The HDD control problem is a trade-off between shock suppression performance and performance under nominal conditions. Limitations of linear control, such as the one defined by the well-known Bode Sensitivity Integral Theorem preclude from improving the shock suppression without deteriorating performance under nominal conditions.

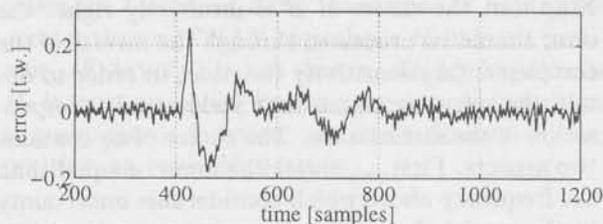


Fig. 4: Measured tracking error during the presence of a shock.

To avoid this trade-off, we propose the following approach to the design of the HDD controller: design a linear controller that yields good performance under nominal conditions, and include a learning feedforward algorithm (as described in the previous sections) to suppress the shocks. To motivate the feasibility of such an approach we must, according to Section 2, prove that (1) a sequence of shocks can be considered as an *NPR*-disturbance, (2) shocks are sufficiently fast detectable and (3) shocks are limited in their frequency range. The first property follows from the fact that in each HDD's application, subsequent shocks will be of approximately the same shape, since this shape will be mainly dependent on the mechanical environment in which the HDD is embedded³. Detectability of the shocks can be denoted from Figure 4, where the measured tracking error (in *track width* [t.w.]) during the presence of a shock is plotted. This measured was obtained from shock table experiments with a real HDD setup. Finally, the Figure 4 shows that the contribution of the shock to the error is mainly located in the low-frequency range (lower than about 200 Hz), so that also the third requirement is satisfied.

The design of the shock suppression algorithm can be split in the design of the detector and the design of the learning feedforward algorithm. Prerequisites are the availability of a model of the plant, a model of the disturbances, and a linear controller C .

Design of the shock detector

Main effort in the design of the shock detector is in the determination of the *coloring* filter M . To solve this problem, we let M have an AR (Auto Regressive) structure. Then, the coefficients of the filter can be obtained by solving the least square problem for linear regression models, taking measurements of the error as identification data set (Ljung, 1987).

³Of course, shocks will have different amplitudes, which requires the addition of a shock size estimator to the algorithm presented in this paper.

Alternatively, the coefficients can be calculated directly from the Toeplitz matrix containing the covariance coefficients of the error stochastic process. The latter can be obtained by inverse Fourier transform of the measured error spectrum, which is available from a experiments with the HDD setup. The choice of the order of the filter is a trade-off between accuracy and compactness. We choose the order of the filter to be 50.

The remaining design freedom is in the choice of y_r , N and TH (or η). Recall that y_r is the contribution of the shock to the error. Since the contribution of the shock to the error is located in the low-frequency range, the easiest way to determine y_r is by low-pass filtering the measurement of the error. N should equal the average detection time (in samples). From simulations with the detector with different values for N and different shock sizes, we found an optimal value of 6. We choose TH equal to 4 times the standard deviation of the output of the correlation receiver (left-hand-side of (10)) under nominal conditions, in order to minimize sufficiently the chance on false alarms.

Design of the learning feedforward algorithm

We determine B as the best approximate of R^{-1} using the MATLAB function *zpetc.m*, which is an algorithm based on (Tomizuka, 1996). The filter we obtain this way is an accurate approximation of R^{-1} up to 800 Hz. To determine the necessary cut-off frequency of the Q filter, we apply the *rule of thumb* (2). Given $T_s = 0.12$ ms and $n_{dev} = 4$ (from simulations with the detector), we find that $f_{max} \ll 500$ Hz, so the cut-off frequency of Q must be fairly smaller than 500 Hz. For both sufficient robustness and satisfactory compensation of the shocks in the most relevant frequency range, we choose the cut-off frequency of Q equal to 200 Hz. For a sufficient amount of averaging, we choose $k_1 = 0.2$.

Simulation results

In order to determine the effectivity of the shock detector, we compare the performance of the detector with that of a simple alternative detection method: a threshold on the absolute error. We choose the value of this threshold in such a way that the reliability of the detector is the same as for our correlation receiver (i.e. 4 times the standard deviation of the error). Figure 5 shows 40 samples of the outputs of both detection algorithms and the thresholds for detection. It can be seen that the correlation receiver (lower part of the figure) yields a faster detection than the simple threshold detector (upper

part of the figure). The figure is representative for the better performance of the correlation receiver: over hundreds of simulations with different shock sizes, the detection speed of the correlation receiver appeared to be almost one time sample faster than that of the simple threshold detector.

To investigate the effectivity of the learning feedfor-

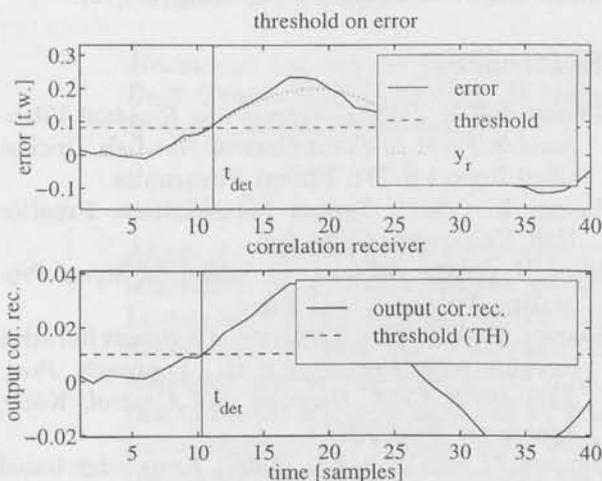


Fig. 5: 40 Samples of the outputs of the simple threshold detector (upper) and the correlation receiver (lower) during the presence of a shock.

ward, we simulated the learning algorithm in combination with the detector for a sequence of 20 shocks, randomly occurring in time. The simulated errors due to respectively shocks 1, 5 and 10 are plotted in Figure 6. As can be seen, the effect of the shocks diminishes thanks to our learning suppression algorithm. The first error peak is decreased with approximately 60%. For the 10th shock, only a small error peak occurs, right after the shock arrival. This can be explained from the fact that the feedforward is activated only at the moment the shock is detected, which is about 6 samples later than the actual shock arrival time. Continuation of the simulation with additional shocks showed that the achieved suppression remained constant for $l > 10$. The feedforward vector that was used for suppression of the 10th shock (and hence constructed in 9 iterations of the update law), is plotted in Figure 7.

6 Conclusions

We introduced the notion of "non-periodically repeating (NPR) disturbances" to describe those non-stationary disturbances that exhibit the same behavior in some randomly distributed time intervals

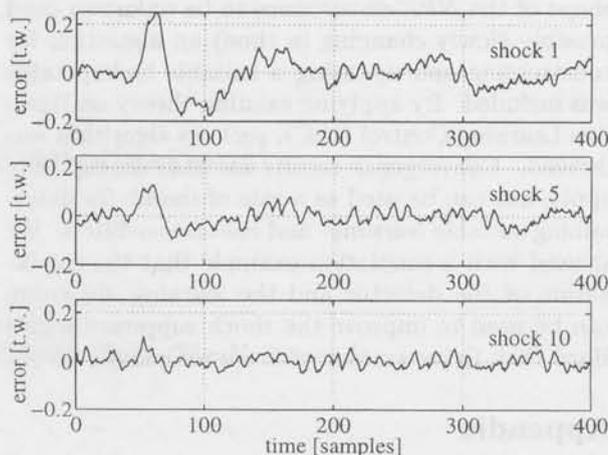


Fig. 6: Simulated errors of the HDD control system for respectively the 1st, 5th and 10nd shock after activating the additional shock suppression algorithm (detector+learning feedforward).

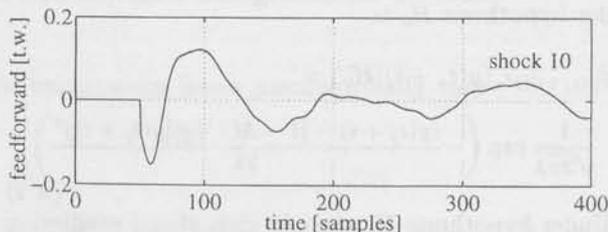


Fig. 7: Simulated feedforward signal after 10 iterations of the HDD learning feedforward shock suppression algorithm.

and are zero elsewhere. Such disturbances may affect the performance of mechanical servo systems unacceptably. Assuming them to be unmeasurable, intuitively, they can be compensated using a lookup-table-based feedforward which is activated when the presence of the disturbance is detected from the output. The restrictions of such an algorithm limit the class of disturbances under consideration in this paper to those NPR-disturbances that are sufficiently fast detectable and limited in their frequency range due to the deviation in detection time. A suitable detector can be derived by introducing hypotheses for respectively presence and absence of the NPR-disturbance and calculating the Likelihood Ratio Test for the detection problem. In implementation, this detector is simply a discrete time filter whose output is compared with a threshold to decide either presence or absence. Since we assumed the

shape of the NPR-disturbance to be unknown (and possibly slowly changing in time) an algorithm for constructing and updating a suitable lookup table was included. By applying existing theory on Iterative Learning Control (ILC), such an algorithm was derived. Convergence results for ILC do not fully apply, but can be used as a rule of thumb for determining suitable *learning*- and *robustness*-filters. We showed with a simulation example that the combination of the detector and the learning algorithm can be used to improve the shock suppression in a Hard Disk Drive system with approximately 60 %.

Appendix

Derivation of the Likelihood Ratio Test

From (Ljung, 1987) we know that the expectation of $y_n(t_j + i)$ is the one step ahead prediction: $\hat{y}_n(t_j + i | t_j + i - 1)$. If $M(q)$ is monic⁴, under hypothesis H_0 , this one step ahead prediction is given by:

$$\hat{y}_n(t_j + i | t_j + i - 1) = [1 - M^{-1}(q)]y(t_j + i). \quad (\text{A.1})$$

Accordingly, the probability density of $y(t_j + i)$ under hypothesis H_0 is

$$p_{y(t_j+i)|H_0}(y(t_j+i)|H_0) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{(y(t_j+i) - [1 - M^{-1}(q)]y(t_j+i))^2}{2\lambda}\right). \quad (\text{A.2})$$

Under hypothesis H_1 the one step ahead prediction of $y_n(t_j + i)$ is given by:

$$\hat{y}_n(t_j + i | t_j + i - 1) = [1 - M^{-1}(q)](y(t_j + i) - y_r(i)), \quad (\text{A.3})$$

and the probability density of $y(t_j + i)$ under hypothesis H_1 is consequently

$$p_{y(t_j+i)|H_1}(y(t_j+i)|H_1) = \frac{1}{\sqrt{2\pi\lambda}} \times \exp\left(-\frac{(y(t_j+i) - [1 - M^{-1}(q)](y(t_j+i) - y_r(i)) - y_r(i))^2}{2\lambda}\right) \quad (\text{A.4})$$

Accordingly, the joint probability densities of $\mathbf{y}(t_j)$ under respectively H_0 and H_1 are given by:

$$p_{\mathbf{y}(t_j)|H_0}(\mathbf{y}(t_j)|H_0) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\lambda}} \times \exp\left(-\frac{(y(t_j+i) - [1 - M^{-1}(q)]y(t_j+i))^2}{2\lambda}\right); \quad (\text{A.5})$$

$$p_{\mathbf{y}(t_j)|H_1}(\mathbf{y}(t_j)|H_1) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\lambda}} \times \exp\left(-\frac{(y(t_j+i) - [1 - M^{-1}(q)](y(t_j+i) - y_r(i)) - y_r(i))^2}{2\lambda}\right). \quad (\text{A.6})$$

⁴i.e. The zeroth coefficient of its impulse response is unity.

Substituting (A.5) and (A.6) in (9), canceling common terms and taking the logarithm we obtain:

$$\ln \Lambda(\mathbf{y}(t_j)) = \frac{1}{\lambda} \sum_{i=1}^N M^{-1}(q)y(t_j+i)M^{-1}(q)y_r(i) - \frac{1}{2\lambda} \sum_{i=1}^N (M^{-1}(q)y_r(i))^2, \quad (\text{A.7})$$

which leads in a straightforward way to (10).

References

- Chang, F.K.K. (1997). *Suppressing Residual Vibrations in Point-to-Point Control*. Nat.Lab. Unclassified Report S-794, Philips Electronics.
- Ljung, L. (1987). *System Identification*. Prentice Hall, Englewood Cliffs, NJ.
- Poor, H. (1987). *Advances in Statistical Signal Processing*, Volume 1. JAI Press.
- Roover, D. de (1996). Synthesis of a robust iterative learning controller using a H_∞ approach, *Proc. 35th IEEE Conf. Decision and Control*, Kobe, Japan, pp. 3044-3049.
- Simaan, M. and Love, P.L. (1990). Knowledge-based detection of out-of-control outputs in process control. *Proc. 29th IEEE Conf. Decision and Control*, Honolulu, Hawaii, pp. 128-129.
- Tomizuka, M. (1996). Zero Phase Error Tracking Algorithm for Digital Control. *J. Dynamic Syst., Measur. Control*, 109, 65-68.
- Trees, L. van (1968). *Detection, Estimation, and Modulation*, part I. John Wiley and Sons.
- Yasuda, Masashi; Osaka, Takahide and Ikeda, Masao (1996). Feedforward control of a vibration isolation system for disturbance suppression. *Proc. 35th IEEE Conf. Decision and Control*, Kobe, Japan, pp. 1229-1233.

A full block S-procedure with applications[‡]

Carsten W. Scherer[§]

*Mechanical Engineering Systems and Control Group
Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands
E-mail: c.scherer@wbmt.tudelft.nl*

Abstract. In this paper we provide a general result that allows to equivalently translate robust performance analysis specifications characterized through a single quadratic Lyapunov function into the corresponding analysis test with multipliers. Just as an illustration we apply the technique to robust quadratic and robust generalized H_2 performance, and we comment on the wide range of its applicability. Finally, we reveal how this technique allows to approach LPV problems in which the control input and measurement output matrix are parameter dependent. The latter is made possible by letting the parameter enter the LPV controller via a kernel representation that generalizes the more conventional LFT structure.

Keywords. S-procedure, structured parametric uncertainty, linear parametrically varying systems.

1 A full block S-procedure

Suppose \mathcal{S} is a subspace of \mathbb{R}^n , $T \in \mathbb{R}^{l \times n}$ is a full row rank matrix, and $\Delta \subset \mathbb{R}^{k \times l}$ is a compact set of matrices of full row rank. Define the family of subspaces

$$\mathcal{S}_\Delta := \mathcal{S} \cap \ker(\Delta T) = \{x \in \mathcal{S} : Tx \in \ker(\Delta)\}$$

indexed by $\Delta \in \Delta$.

In the terminology of the behavioral approach, \mathcal{S} is the system, T picks the interconnection variables that are constrained by the uncertainties, the elements of $\Delta \in \Delta$ define kernel representations of the possible uncertainties, and \mathcal{S}_Δ is the perturbed system.

Suppose $N \in \mathbb{R}^{n \times n}$ is a fixed symmetric matrix. The goal is to render the implicit negativity condition

$$\forall \Delta \in \Delta : N < 0 \text{ on } \mathcal{S}_\Delta$$

[‡]This paper is presented at the 36th IEEE Conference on Decision and Control, 10-12 December 1997, San Diego, CA, USA. Copyright of this paper remains with IEEE.

[§]The author would like to thank Ted Iwasaki (Tokyo Institute of Technology) for helpful discussions.

explicit. We want to relate this property, under certain technical hypotheses, to the existence of a multiplier P that satisfies

$$N + T^T P T < 0 \text{ on } \mathcal{S} \text{ and } P > 0 \text{ on } \ker(\Delta)$$

for all uncertainties $\Delta \in \Delta$.

The required technical condition will be related to a certain well-posedness property; here it amounts to the complementarity of the subspace \mathcal{S}_Δ to a fixed subspace $\mathcal{S}_0 \subset \mathcal{S}$ that is sufficiently large. Moreover, the quadratic form N is supposed to be nonnegative on this subspace; in the applications we have in mind, this is a property on the performance index under consideration that is, interestingly enough, indispensable in reducing the underlying controller design problem to an LMI problem. To be precise, we require

$$\dim(\mathcal{S}_0) \geq k \text{ and } N \geq 0 \text{ on } \mathcal{S}_0.$$

Theorem 1.1 *The condition*

$$\forall \Delta \in \Delta : \mathcal{S}_\Delta \cap \mathcal{S}_0 = \{0\}, N < 0 \text{ on } \mathcal{S}_\Delta$$

holds iff there exists a matrix P that satisfies

$$\forall \Delta \in \Delta : \begin{cases} N + T^T P T < 0 \text{ on } \mathcal{S} \\ P > 0 \text{ on } \ker(\Delta) \end{cases}$$

2 Application to robust performance problems

Consider the first order image representation

$$\begin{pmatrix} \dot{x} \\ z_u \\ z_p \end{pmatrix} = \begin{pmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} x \\ w_1 \\ w_2 \end{pmatrix} \quad (1)$$

of a system (with A Hurwitz) in L_2 . We can assume w.l.o.g. that the third block column of the matrix has full column rank.

Here, w_1 and w_2 are latent variables; z_p are the variables on which we impose the performance specification, and z_u are the interconnection variables to let the parameters enter the system.

We consider the linear parameter-varying (LPV) systems obtained as follows: they are parametrized by all continuous curves

$$\Delta : [0, \infty) \rightarrow \Delta$$

with a given set of values

$$\Delta \subset \mathbb{R}^{k \times l}$$

that captures both the size and the structure of the parameters. We assume that Δ is compact and consists of full row rank matrices only. These parameter curves enter (1) via a kernel representation as

$$\Delta(t)z_u(t) = 0. \quad (2)$$

We will clarify below that this generalizes the more standard LFT structure.

As a first property, we intend to characterize that the representation of LPV systems is *well-posed*:

$$\Delta D_{11} \text{ is nonsingular for every } \Delta \in \Delta. \quad (3)$$

In the case of well-posedness (3), we observe that the LPV system admits the alternative representation

$$\begin{pmatrix} \dot{x} \\ z_p \end{pmatrix} = \begin{pmatrix} A(\Delta(t)) & B(\Delta(t)) \\ C(\Delta(t)) & D(\Delta(t)) \end{pmatrix} \begin{pmatrix} x \\ w_2 \end{pmatrix} \quad (4)$$

where

$$\begin{pmatrix} A(\Delta) & B(\Delta) \\ C(\Delta) & D(\Delta) \end{pmatrix} = \begin{pmatrix} A & B_2 \\ C_2 & D_{22} \end{pmatrix} + \begin{pmatrix} B_1 \\ D_{21} \end{pmatrix} (\Delta D_{11})^{-1} \Delta \begin{pmatrix} C_1 & D_{12} \end{pmatrix}.$$

Given the performance index P_p , the second goal is to guarantee uniform (in the uncertainty) robust exponential stability, and robust quadratic performance:

$$\int_0^\infty z_p(t)^T P_p z_p(t) dt \leq 0 \quad (5)$$

holds for any trajectory of any of the LPV systems with $x(0) = 0$. Let us include the following technical hypotheses:

$$D_{11} \text{ has } k \text{ columns, } D_{21}^T P_p D_{21} \geq 0.$$

The first property is obviously necessary for well-posedness; the second property holds for the standard H_∞ or positive real index and many others.

It is well-known and elementary to show that robust exponential stability and robust quadratic performance is guaranteed by the existence of some $X > 0$ such that

$$\forall \Delta \in \Delta : * \begin{pmatrix} 0 & X & 0 \\ X & 0 & 0 \\ 0 & 0 & P_p \end{pmatrix} \begin{pmatrix} I & 0 \\ A(\Delta) & B(\Delta) \\ C(\Delta) & D(\Delta) \end{pmatrix} < 0. \quad (6)$$

(Note that throughout this paper we will employ the abbreviation $*PM$ for $M^T P M$.)

We will use Theorem 1.1 to equivalently reformulate this characterization as

$$* \begin{pmatrix} 0 & X & 0 & 0 \\ X & 0 & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P_p \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{pmatrix} < 0, \quad (7)$$

where P is a *multiplier* that satisfies

$$\forall \Delta \in \Delta : P > 0 \text{ on } \ker(\Delta). \quad (8)$$

Theorem 2.1 *Well-posedness (3) and (6) hold iff there exists a P with (7) and (8).*

Proof. We just apply Theorem 1.1 to

$$N = \begin{pmatrix} 0 & X & 0 & 0 \\ X & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & P_p \end{pmatrix}, \quad T = (0 \ 0 \ I \ 0),$$

$$S = \text{im} \begin{pmatrix} I & 0 & 0 \\ A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{pmatrix}, \quad S_0 = \text{im} \begin{pmatrix} 0 \\ B_2 \\ D_{12} \\ D_{22} \end{pmatrix}.$$

Before commenting on this result, let us look at the robust generalized H_2 problem that offers an interesting additional insight into the solution of robust mixed problems.

Suppose T_2 and T_∞ are two matrices of full row rank whose number of columns equals the size of z_p . Then we intend to characterize that

$$\|T_\infty z_p\|_\infty \leq \|T_2 z_p\|_2 \quad (9)$$

holds for the whole family of systems (4) with $x(0) = 0$. If we have $T_2 z_p = w_2$, this property amounts to the gain of the mapping $L_2 \ni w_2 \rightarrow T_2 z_p \in L_\infty$ defined by (4) with $x(0) = 0$ being robustly not larger than one. This gain has been called generalized H_2 -norm (Rotea, 1993).

Let us assume that

$$T_\infty D(\Delta) = 0$$

what is indeed required to ensure $\|T_\infty z_p\|_\infty < \infty$ for all $w_2 \in L_2$ in (4). Then the following result is very easy to prove: If there exists an $X > 0$ such that, for all $\Delta \in \Delta$,

$$* \begin{pmatrix} 0 & X & 0 \\ X & 0 & 0 \\ 0 & 0 & -T_2^T T_2 \end{pmatrix} \begin{pmatrix} I & 0 \\ A(\Delta) & B(\Delta) \\ C(\Delta) & D(\Delta) \end{pmatrix} < 0, \quad (10)$$

$$\begin{pmatrix} I \\ C(\Delta) \end{pmatrix}^T \begin{pmatrix} -X & 0 \\ 0 & T_\infty^T T_\infty \end{pmatrix} \begin{pmatrix} I \\ C(\Delta) \end{pmatrix} < 0 \quad (11)$$

then (4) is robustly exponentially stable and (9) holds for any system trajectory.

We end up with two inequalities in the parameter Δ . Therefore, we have to apply Theorem 1.1 to each of these inequalities individually, what leads to two independent multipliers to equivalently reformulate this test.

Theorem 2.2 Suppose $T_2 D_{21} = 0$. Then well-posedness (3) and (10), (11) hold iff there exist multipliers P_1 and P_2 that both satisfy (8) and

$$* \begin{pmatrix} 0 & X & 0 & 0 \\ X & 0 & 0 & 0 \\ 0 & 0 & P_1 & 0 \\ 0 & 0 & 0 & -T_2^T T_2 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{pmatrix} < 0$$

as well as

$$* \begin{pmatrix} -X & 0 & 0 \\ 0 & P_2 & 0 \\ 0 & 0 & T_\infty^T T_\infty \end{pmatrix} \begin{pmatrix} I & 0 \\ C_1 & D_{11} \\ C_2 & D_{21} \end{pmatrix} < 0.$$

Remarks.

- The equivalences of these robust performance characterizations seem not to have appeared in the literature. They extend (Megretski and Rantzer, 1997; Iwasaki et al., 1995) to robust performance problems for general LFT uncertainty descriptions. Comparable robust performance specifications with multipliers that are only indirectly described have been provided in (Tokunaga et al., 1996; Scherer, 1996).
- It is an important structural insight that, in Theorem 2.1, the combined multiplier $\begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix}$

for performance and parameter can be taken block-diagonal.

- If the parameter has a block-diagonal structure, the channel-wise application of the standard S-procedure leads from (6) to (7) with a block-diagonal scaling P . It is known that this step introduces conservatism. Using multipliers which are full and whose structure is not explicitly specified at the outset leads to a reformulation without conservatism. Therefore, we call the technique presented here a *full-block S-procedure*.
- An important aspect is the ease to proceed from (6) to (7) in a formal manner, just by referring to Theorem 1.1. Moreover, the derivation is not only straightforward, but leads to simple formulas that favorably compare with their sometimes pretty intricate counterparts in the literature.
- There are numerous further applications of the full block S-procedure that are currently under investigation. As most prominent ones, we mention that one can straightforwardly extend general robust mixed problems as proposed in (El Ghaoui and Follmer, 1996; Masubuchi et al., 1996) to full block scalings what reduces conservatism; see also (Tokunaga et al., 1996). Moreover, the techniques apply to analysis problems with parameter dependent Lyapunov functions along the lines of (Feron et al., 1995).

3 Application to LPV control

For the discussion of LPV control we concentrate on the quadratic performance specification with index P_p that is, in addition, non-singular. In contrast to robust control, in LPV control it is assumed that the parameter curve is on-line measurable. These design problems can be approached either by directly using the analysis test (6) (Apkarian et al., 1995; Becker et al., 1993; Köse et al., 1996) or by proceeding with the multiplier version (7) (Packard, 1994; Apkarian and Gahinet, 1995; Becker and Packard, 1994; Helmersson, 1995; Scorletti and El Ghaoui, 1995; Scherer, 1996).

The former suffers from the disadvantage that the matrices defining the control input and the measured output are not allowed to depend on the parameter. In the latter, usually a restricted class of structured scalings is employed. One of the main motivations for the full block S-procedure is to overcome these restrictions in LPV control.

Since we need dualization, the parameter dependent system is assumed to admit a slightly more special

description as

$$\begin{pmatrix} \dot{x} \\ w_u \\ z_u \\ w_p \\ z_p \\ y \end{pmatrix} = \begin{pmatrix} A & B_1 & B_2 & B \\ 0 & I & 0 & 0 \\ C_1 & D_{11} & D_{12} & E_1 \\ 0 & 0 & I & 0 \\ C_2 & D_{21} & D_{22} & E_2 \\ C & F_1 & F_2 & 0 \end{pmatrix} \begin{pmatrix} x \\ w_u \\ w_p \\ u \end{pmatrix} \quad (12)$$

with parameters entering as

$$\Delta(t) \begin{pmatrix} w_u(t) \\ z_u(t) \end{pmatrix} = 0. \quad (13)$$

As usual, u denotes the control input variable and y the measured output variable.

We assume that the controller is described as

$$\begin{aligned} \dot{x}_c &= A_c x_c + B_c \begin{pmatrix} y \\ w_c \end{pmatrix} \\ \begin{pmatrix} u \\ z_c \end{pmatrix} &= C_c x_c + D_c \begin{pmatrix} y \\ w_c \end{pmatrix} \end{aligned} \quad (14)$$

and a specific parameter curve $\Delta(\cdot)$ enters as

$$\Delta_c(\Delta(t)) \begin{pmatrix} w_c(t) \\ z_c(t) \end{pmatrix} = 0 \quad (15)$$

for a to be constructed scheduling function $\Delta_c : \Delta \rightarrow \mathbb{R}^{k_c \times l_c}$.

The description of the controlled system is obtained by interconnecting the LTI systems (12), (14) to get

$$\begin{pmatrix} \dot{x} \\ w_u \\ z_u \\ w_c \\ z_c \\ w_p \\ z_p \end{pmatrix} = \begin{pmatrix} A & B_1 & B_2 & B_3 \\ 0 & I & 0 & 0 \\ C_1 & D_{11} & D_{12} & D_{13} \\ 0 & 0 & I & 0 \\ C_2 & D_{21} & D_{22} & D_{23} \\ 0 & 0 & 0 & I \\ C_3 & D_{31} & D_{32} & D_{33} \end{pmatrix} \begin{pmatrix} x \\ w_u \\ w_c \\ w_p \end{pmatrix} \quad (16)$$

and letting the parameters enter via (13), (15).

The LPV problem now reads as follows: Find an LTI controller (14) and a scheduling function Δ_c such that the controlled system (16), (13), (15) is robustly exponentially stable and robustly satisfies the performance specification (5).

Robust stability and robust performance is characterized through Theorem 2.1 by employing multipliers P that satisfy

$$\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} > 0 \text{ on } \ker \begin{pmatrix} \Delta & 0 \\ 0 & \Delta_c(\Delta) \end{pmatrix} \quad (17)$$

for all $\Delta \in \Delta$. Note that one can dualize this test (Iwasaki et al., 1995; Scherer, 1996) to arrive at a

formulation with the dual performance index \tilde{P}_p and the dual multipliers \tilde{P} that fulfill

$$\begin{pmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{21} & \tilde{P}_{22} \end{pmatrix} < 0 \text{ on } \text{im} \begin{pmatrix} \Delta^T & 0 \\ 0 & \Delta_c(\Delta)^T \end{pmatrix}.$$

The duality coupling for the performance index and the multipliers is $\tilde{P}_p = P_p^{-1}$ and $\tilde{P} = P^{-1}$ respectively.

The synthesis inequalities for the LPV problem at hand are obtained along standard lines (Packard, 1994; Apkarian and Gahinet, 1995; Becker and Packard, 1994; Helmersson, 1995; El Ghaoui and Følcher, 1996; Scherer, 1996): Start with the primal and dual analysis inequalities for the controller system which involve the Lyapunov matrices \mathcal{X} and \mathcal{X}^{-1} . Then eliminate the controller parameters for which one requires to compute basis matrices Φ and Ψ of

$$\ker(B^T E_1^T E_2^T), \ker(C F_1 F_2).$$

Due to the particular structure, the resulting two inequalities simplify considerably; one ends up with the LMIs

$$\begin{aligned} & \begin{pmatrix} 0 & X & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^T & R & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ A & B_1 & B_2 \\ 0 & I & 0 \\ C_1 & D_{11} & D_{12} \\ 0 & 0 & I \\ C_2 & D_{21} & D_{22} \end{pmatrix} \Psi < 0 \quad (18) \end{aligned}$$

$$\begin{aligned} & \begin{pmatrix} 0 & Y & 0 & 0 & 0 & 0 \\ Y & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tilde{Q} & \tilde{S} & 0 & 0 \\ 0 & 0 & \tilde{S}^T & \tilde{R} & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{Q}_p & \tilde{S}_p \\ 0 & 0 & 0 & 0 & \tilde{S}_p^T & \tilde{R}_p \end{pmatrix} \begin{pmatrix} A^T & C_1^T & C_2^T \\ -I & 0 & 0 \\ B_1^T & D_{11}^T & D_{21}^T \\ 0 & -I & 0 \\ B_2^T & D_{12}^T & D_{22}^T \\ 0 & 0 & -I \end{pmatrix} \Phi > 0 \quad (19) \end{aligned}$$

in the symmetric matrices X, Y coupled as

$$\begin{pmatrix} X & I \\ I & Y \end{pmatrix} > 0 \quad (20)$$

and in

$$P_{11} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}, \tilde{P}_{11} = \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix}$$

which satisfy

$$\forall \Delta \in \Delta : \begin{cases} P_{11} > 0 \text{ on } \ker(\Delta) \\ \tilde{P}_{11} < 0 \text{ on } \text{im}(\Delta^T) \end{cases} \quad (21)$$

We observe that other parts of the multipliers simply drop out and do not occur in this result. In this way one proves the necessity part of the following theorem.

Theorem 3.1 *There exists an LPV controller (14), (15) for (12), (13) such that the controlled system satisfies the condition for robust performance in Theorem 2.1 if and only if there exist $X, Y, P_{11}, \tilde{P}_{11}$ that fulfill (18)-(20) and (21).*

As an important novel aspect, we make no assumption on the multipliers. Indeed, this causes the main difficulties in proving the reverse direction by constructing a suitable LPV controller.

Let us briefly describe how to construct such a controller. For that purpose suppose that $X, Y, P_{11}, \tilde{P}_{11}$ satisfy the synthesis conditions.

The most difficult step in the construction is covered by the following theorem.

Theorem 3.2 *Suppose the matrices P_{11}, \tilde{P}_{11} fulfill (21). Then there exists a continuous function $\Delta_c(\Delta)$ and an extension P_{12}, P_{21}, P_{22} such that*

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \text{ satisfies } P^{-1} = \begin{pmatrix} \tilde{P}_{11} & * \\ * & * \end{pmatrix}$$

and such that (17) holds for all $\Delta \in \Delta$.

This results allows to find extended scalings and a suitable scheduling function. If we observe that (18)-(20) are nothing but the synthesis inequalities for the quadratic performance problem with index $\begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix}$, the construction of the LTI part of the controller can be obtained by a standard nominal design procedure.

Remark. Note that the synthesis inequalities for designing a robust controller (w_c, z_c are absent) are given by (18)-(20), (21) including the duality coupling

$$\tilde{P}_{11} = P_{11}^{-1}.$$

This relation renders these conditions, as well-known, non-convex in the variables P_{11} and \tilde{P}_{11} . Under specific structural hypotheses on the multipliers (as discussed below), this procedure has been followed in (Scherer, 1996) and is worked out in full detail in (Scherer, 1997). A full discussion of the novel procedure including all the proofs will be available in a forthcoming paper.

Note that the multipliers in (21) are described by infinitely many linear matrix inequalities. They can be reduced to finitely many inequalities by gridding the parameter space Δ . Instead, however, we propose to constrain the scalings, possibly involving conservatism, such that one can exploit convexity in order to reduce the test to finitely many LMIs that are amenable to standard software.

Let us illustrate this technique and the benefit of the presented approach over existing ones by briefly

resorting to the standard LFT description of uncertain systems; in that case the parameters enter as

$$w_u(t) = \delta(t)z_u(t) \quad (22)$$

where

$$\delta = \begin{pmatrix} \delta_1 I_1 & & 0 \\ & \ddots & \\ 0 & & \delta_m I_m \end{pmatrix}$$

with the $\dim(I_j)$ times repeated scalar parameters δ_j varying in $[-1, 1]$. Note that (22) can be written as

$$(I - \delta(t)) \begin{pmatrix} w_u(t) \\ z_u(t) \end{pmatrix} = 0$$

such that it nicely fits in our more general scenario. If we recall that

$$\ker(I - \delta) = \text{im} \begin{pmatrix} \delta \\ I \end{pmatrix},$$

the constraints (21) on the scalings hence read as

$$\begin{pmatrix} \delta \\ I \end{pmatrix}^T \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \begin{pmatrix} \delta \\ I \end{pmatrix} > 0 \quad (23)$$

and

$$\begin{pmatrix} I \\ -\delta^T \end{pmatrix}^T \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \begin{pmatrix} I \\ -\delta^T \end{pmatrix} < 0 \quad (24)$$

for all $\delta_j \in [-1, 1]$. (Note that δ happens to be symmetric; since the generalization to a possibly non-symmetric structure δ is straightforward, we neglect this extra property.)

Let us now denote the extreme points of the set of all δ by δ^j . If we impose the strong constraint $Q < 0$ and $\tilde{R} > 0$ on parts of the multipliers, we infer that (23), (24) hold for all δ iff they hold for all extreme points δ^j . This is the situation considered in our previous work (Scherer, 1996). However, it is simple to relax the strong negativity/positivity condition by referring to a partial convexity argument. Indeed, if partitioning

$$Q = \begin{pmatrix} Q_1 & * \\ & \ddots \\ * & Q_m \end{pmatrix}, \quad \tilde{R} = \begin{pmatrix} \tilde{R}_1 & * \\ & \ddots \\ * & \tilde{R}_m \end{pmatrix}$$

according to δ, δ^T , we can work with the relaxed negativity/positivity constraint

$$Q_j < 0, \quad \tilde{R}_j > 0 \quad (25)$$

imposed on the diagonal blocks only. This implies that the left-hand sides of (23), (24) are partially concave, convex functions of δ respectively. Hence

one can again conclude that (23), (24) are satisfied for all δ iff they are fulfilled for the extreme points δ^j .

The set of synthesis inequalities then consists of (18)-(20), (23)-(24) for the extreme points, and (25), such that the feasibility test amounts to a standard LMI problem.

In the talk we will provide an example which reveals that this relaxation can considerably reduce the conservatism.

4 Conclusion

We have given a general full block S-procedure in order to rewrite robust performance tests formulated in terms of a constant Lyapunov matrix into the corresponding multiplier test without conservatism. As an application, we have given a full solution to the corresponding LPV synthesis problem where the multipliers are in no way restricted. This has been made possible by proposing a novel scheme to schedule the LPV controller; the parameters are viewed to define a kernel representation of a static system that is interconnected with the LTI part of the controller.

References

- Apkarian, P. and P. Gahinet (1995). A convex characterization of gain-scheduled \mathcal{H}_∞ controllers, *IEEE Trans. Automat. Control*, AC-40, 853-864.
- Apkarian, P., P. Gahinet, and G. Becker (1994). Self-scheduled \mathcal{H}_∞ control of linear parameter-varying systems, *Proc. Amer. Control Conf.*, pp. 856-860.
- Feron, E., P. Apkarian, and P. Gahinet (1996). Analysis and synthesis of robust control systems via parameter-dependent Lyapunov functions, *IEEE Trans. Automat. Control*, AC-41, 1041-1046.
- Becker, G., A. Packard, D. Philbrick, and G. Balas (1993). Control of parametrically-dependent linear systems: A single quadratic Lyapunov approach, *Proc. Amer. Control Conf.*, San Francisco, CA, pp. 2795-2799.
- Becker, G. and A. Packard (1994). Robust performance of linear parametrically varying systems using parametrically dependent linear feedback, *Systems Control Lett.*, 23, 205-215.
- Boyd, S., L. El Ghaoui, E. Feron, and V. Balakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory*. SIAM Studies in Applied Mathematics 15, SIAM, Philadelphia.
- Gahinet, P., A. Nemirovski, A. Laub, and M. Chilali (1994). *LMI Control Toolbox*. The Mathworks Inc.
- El Ghaoui, L. and J.P. Folcher (1996). Multiobjective robust control of LTI systems subject to unstructured perturbations, *Systems Control Lett.*, 28, 23-30.
- Helmersson, A. (1995). *Methods for Robust Gain-Scheduling*. PhD thesis, Linköping University, Sweden.
- Iwasaki, T., S. Hara, and T. Asai (1995). Well-posedness theorem: a classification of LMI/BMI-reducible robust control problems, *preprint*.
- Köse, I.E., F. Jabbari, and W.E. Schmittendorf (1996). A direct characterization of L_2 -gain controllers for LPV systems, *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, pp. 3990-3995.
- Masubuchi, I., N. Suda, and A. Ohara (1996). Robust multi-objective controller design via convex optimization, *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, pp. 263-264.
- Packard, A. (1994). Gain scheduling via linear fractional transformations, *System and Control Letters*, 22, 79-92.
- Megretski, A. and A. Rantzer (1997). System analysis via integral quadratic constraints, *IEEE Trans. Automat. Control*, AC-42, 819-830.
- Rotea, M.A. (1993). The generalized \mathcal{H}_2 control, *Automatica*, 29, 373-385.
- Scherer, C.W. (1996). Robust generalized H_2 control for uncertain and LPV systems with general scalings, *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, pp. 3970-3975.
- Scherer, C.W. (1997). Robust mixed control and LPV control with full block scalings, *preprint*.
- Scorletti, G. and L. El Ghaoui (1995). Improved linear matrix inequality conditions for gain-scheduling, *Proc. 34th IEEE Conf. Decision Contr.*, New Orleans, MA, pp. 3626-3631.
- Tokunaga, H., T. Iwasaki, and S. Hara (1996). Multi-objective robust control with transient specifications, *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, pp. 3482-3483.

An internal-model-based framework for the analysis and design of repetitive and learning controllers[‡]

D. de Roover[§] and O.H. Bosgra

*Mechanical Engineering Systems and Control Group
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.
E-mail: d.derover@wbmt.tudelft.nl, o.h.bosgra@wbmt.tudelft.nl*

Abstract. Repetitive and iterative learning control are two modern control strategies, used in tracking systems in which the signals are periodic in nature. Both schemes are in some sense based on the internal model principle applied to periodic signals. Because of the great number of successful applications, both schemes have been investigated in their own right, largely independent of the results available for the internal model principle. This has clouded some basic properties of these schemes, with consequences for their analysis and design. This paper returns to the origin of repetitive and learning control by analyzing and designing both controllers within a general internal model based framework. A link is made with several existing approaches, for which the design is shown to depend on modifications of the internal model.

Keywords. Iterative learning control, repetitive control, internal model principle, duality, synthesis.

1 Introduction

In practice, many tracking systems have to deal with periodic reference and/or disturbance signals, for example computer disk drives, rotating machine tools, or robots that have to perform their tasks repeatedly. It is well known that any periodic signal can be generated by an autonomous system consisting of a time-delay element inside a positive feedback loop. Therefore, in view of the internal model principle, (Francis and Wonham, 1975), it might be expected that accommodation of these periodic signals can be achieved by duplicating this model inside a feedback loop. In the literature, two types of compensators can be found which accomplish this: the *repetitive controller*, see for example Inoue *et al.* (1981), Hara *et al.* (1988), Tomizuka *et al.* (1989) and Sadegh (1991), and the *iterative learning con-*

troller, see for example Arimoto *et al.* (1984), Moore *et al.* (1992) and Horowitz (1993).

Although it has been recognized that both schemes differ in the way periodic compensation is performed, (Hara *et al.*, 1988; Horowitz, 1993), still the impression exists that both schemes are equivalent. However, in a recent paper, it was shown that the schemes are not equivalent, but are related by duality, which is a consequence of the difference in location of the internal model inside the compensator, (de Roover and Bosgra, 1997). It was shown that a repetitive controller has the structure of a servo compensator-with the internal model located at the system output-while a learning controller has the structure of a disturbance observer, with the internal model located at the system input. In this paper we use the general framework given in de Roover and Bosgra (1997) to set up a general framework for the synthesis of (MIMO) repetitive and learning controllers. It is shown that a number of existing repetitive and learning control schemes can be put into this framework, according to specific modifica-

[‡]This paper was presented at the 36th IEEE Conference on Decision and Control, 10-12 December 1997, San Diego, CA, USA. Copyright of this paper remains with IEEE.

[§]The work of Dick de Roover is financially supported by Philips Research Laboratories, Eindhoven, The Netherlands.

tions in the internal model. Throughout this paper, \mathbb{R} denotes the field of real numbers. Let n_u denote the dimension of the vector u , then \mathbb{R}^u denotes the set of all n_u -vectors with elements in \mathbb{R} . Likewise, $\mathbb{R}^{u \times y}$ denotes the set of all $n_u \times n_y$ matrices with elements in \mathbb{R} , and I_u denotes the $n_u \times n_u$ identity matrix. Furthermore, z denotes the discrete-time delay operator, and $\mathbb{R}(z)$ denotes the set of all rational functions with real coefficients in z . Let $M \in \mathbb{R}^{n \times m}$ then $\rho(M)$ denotes the rank of M , and $\rho(M) \leq \min\{n, m\}$.

2 The robust periodic control problem

Any periodic signal can be generated by an autonomous system consisting of a time-delay element inside a positive feedback loop, corresponding to the periodicity and with appropriate initial conditions, see Figure 1. For example, a discrete time periodic

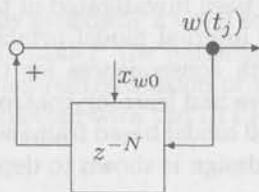


Fig. 1: Periodic signal w generated by an autonomous system with appropriate initial conditions x_{w0} and N denoting the number of delays.

signal of length N can be generated by:

$$\begin{aligned} x_w(t_{j+1}) &= A_w x_w(t_j), & x_w(t_0) &= x_{w0} \\ w(t_j) &= C_w x_w(t_j), \end{aligned} \quad (1)$$

with

$$A_w = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{N \times N}$$

$$C_w = [1 \ 0 \ 0 \ \cdots \ 0].$$

Note that the spectrum of A_w consists of N roots equally spaced on the unit disk, *i.e.* an internal model of this periodic signal is simply given by:

$$\phi_w(z) = 1 - z^{-N}. \quad (2)$$

Next, consider a discrete time linear time-invariant (LTI) plant $P(z) \in \mathbb{R}(z)^{y \times u}$ with input signal $u = u_p + w_u \in \mathbb{R}^u$, and output signal $y = Pu \in \mathbb{R}^y$, according to Figure 2. Given a desired output sig-

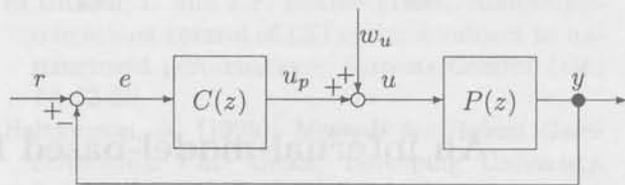


Fig. 2: General periodic control problem; u_p and w_u denote a periodic control signal and periodic input disturbance, respectively.

nal which is periodic with a period of N samples: $r(t_j) = r(t_{j+N\Delta T})$, $j = 0, \Delta T, 2\Delta T, \dots$, with ΔT denoting the sampling time. Let $e = r - y$ be the tracking error. Then we define the *robust periodic control problem* as:

Definition 2.1 *The robust periodic control problem is to find a feedback compensator $C(z)$ for the system $P(z)$ such that:*

1. *The resulting compensated system is exponentially stable.*
2. *The tracking error e tends to zero asymptotically, for all periodic references r and periodic disturbances w_u satisfying (1)¹.*
3. *Properties 1. and 2. are robust, i.e. they also hold in case the dynamics of P are perturbed.*

Repetitive and learning control are two strategies which attempt to solve this problem.

3 Synthesis in an internal-model-based framework

In this section, we give necessary and sufficient conditions, for both a repetitive and learning type of controller to solve the problem, which are a direct consequence of the general framework given in de Roover and Bosgra (1997).

3.1 Repetitive control

A prototype repetitive controller which might solve the periodic control problem is given by:

$$C(z) = R(z) (\phi_w^{-1}(z) I_y), \quad (3)$$

with $R(z) \in \mathbb{R}(z)^{u \times y}$ denoting a stabilizing repetitive control gain, and $\phi_w(z) = 1 - z^{-N}$ is the minimal polynomial of the autonomous system (1). Note that due to n_y times duplication of $\phi_w(z)$, the compensator (3) has a state space realization which has

¹Note that there is no fundamental difference between an error resulting from a disturbance at the output or from a reference input r ; we will only consider r .

the structure of a servo compensator; this duplication is necessary to obtain robustness property 3. of Definition 2.1. The following theorem gives a necessary and sufficient condition for the repetitive controller (3) to solve the robust repetitive control problem:

Theorem 3.1 Let $\{A, B, C, D\}$ be a minimal realization of $P(z) \in \mathbb{R}(z)^{y \times u}$, let $\{A_r, B_r\}$ be a controllable realization of the n_y -fold duplication of $\phi_w^{-1}(z) = 1/(1-z^{-N})$. Then there exists a repetitive control system (3) that solves the robust periodic control problem, defined in Definition 2.1, if and only if

$$\rho \left(\begin{bmatrix} \lambda I - A & B \\ -C & D \end{bmatrix} \right) = n_x + n_y, \quad \forall \lambda \in \sigma(A_w). \quad (4)$$

This result can be directly derived from the results available for a general servocompensator, by noting that a repetitive controller is a servocompensator for periodic signals (see de Roover and Bosgra, 1997). In fact, condition (4) guarantees the controllability of the series connection of $P(z)$ followed by $\phi_w^{-1}(z)I_y = I_y(\phi_w(z)I_y)^{-1}$. It implies that the system P does not have transmission zeros located at the spectrum of A_w , and that P has at least as many inputs as outputs. If the conditions of Theorem 3.1 hold, the control input u_p can be designed according to:

$$u_p(t_j) = K_r x_r(t_j) + K x_c(t_j),$$

with $x_r \in \mathbb{R}^r$ the state-vector of the n_y -fold duplication of the internal model, and $x_c \in \mathbb{R}^c$ the state-vector of any plant stabilizing compensator, and $[K \ K_r]$ a stabilizing state feedback for the series connection of plant and internal model. For example, an observer-based compensator can be designed, which stabilizes the system $\{A, B, C, D\}$. In this case, a state space realization for the resulting compensator $C(z)$ is given by:

$$\begin{aligned} A_c &= \begin{bmatrix} A_r & 0 \\ (B+LD)K_r & A+BK+L(C-DK) \end{bmatrix} \\ B_c &= \begin{bmatrix} B_r \\ K \end{bmatrix}, \quad C_c = [K_r \ -L], \quad D_c = 0, \end{aligned} \quad (5)$$

with L denoting an observer gain, designed such that $(A+LC)$ is stable. According to the separation principle, L and $[K \ K_r]$ can be designed independently.

3.2 Learning control

There might be situations where condition (4) does not hold, for example if the system P has more outputs than inputs. Consequently there does not

exist a repetitive controller that can solve the robust periodic control problem defined in Definition 2.1. Notwithstanding, a *dual* repetitive controller might exist that can solve the robust periodic control problem for periodic disturbances w_u at the system input, and for periodic reference signals r under certain controllability restrictions (see de Roover and Bosgra, 1997). A candidate dual repetitive controller is given by:

$$C(z) = (\phi_w^{-1}(z)I_u) L(z), \quad (6)$$

with $L(z) \in \mathbb{R}(z)^{u \times y}$ called the dual repetitive control gain. We call this dual repetitive controller a *continuously learning controller*—or simply *learning controller*—as opposite to iterative learning controllers, because it continuously updates (learns) the input signal u_p . Dual to Theorem 3.1 the following result analyses under which conditions a learning controller (6) exists that solves the periodic control problem:

Theorem 3.2 Let $\{A, B, C, D\}$ be a minimal realization of $P(z) \in \mathbb{R}(z)^{y \times u}$, let $\{A_l, C_l\}$ be an observable realization of the n_u -fold duplication of $\phi_w^{-1}(z) = 1/(1-z^{-N})$, and let r be available for compensation. Then there exists a learning controller (6) that solves the robust periodic control problem, defined in Definition 2.1, if and only if

$$\rho \left(\begin{bmatrix} \lambda I - A & B \\ -C & D \end{bmatrix} \right) = n_x + n_u, \quad \forall \lambda \in \sigma(A_w). \quad (7)$$

and

$$\rho(B) + n_y = \rho \left(\begin{bmatrix} B \\ D \end{bmatrix} \right). \quad (8)$$

In fact condition (7) guarantees the observability of the series connection of $\phi_w^{-1}(z)I_u = (\phi_w(z)I_u)^{-1}I_u$ followed by $P(z)$.

Remark 3.3 The first condition of Theorem 3.2 states that P may not have transmission zeros located at the spectrum of A_w , and that P must have at least as many outputs as inputs. However, the second condition—which is necessary for asymptotic tracking of r , but not for rejection of w_u —is in general only true if P has at least as many inputs as outputs. Therefore, asymptotic tracking of r using a dual repetitive controller is in general only possible if P is square. Note also that the condition that r should be available for feedback, is a necessary restriction, which is not required for the repetitive controller (3).

In de Roover and Bosgra (1997), it was shown that a learning controller (6) has the structure of a disturbance observer, which is a consequence of the internal model being located at the system input. Therefore, if the conditions of Theorem 3.2 are satisfied,

the control input u_p can be constructed according to:

$$u_p(t_j) = K\hat{x}(t_j) + C_l x_l(t_j),$$

with K denoting a state feedback gain, designed such that $(A+BK)$ is stable, \hat{x} denoting the reconstructed system state, and x_l denoting the state-vector of the n_u -fold duplication of the internal model. In this case, the resulting compensator $C(z)$ has the following state space realization:

$$A_c = \begin{bmatrix} A_l & L_l(C+DK) \\ 0 & A+BK+L(C-DK) \end{bmatrix}, \quad (9)$$

$$B_c = \begin{bmatrix} L_l \\ -K \end{bmatrix}, \quad C_c = [C_l \quad L], \quad D_c = 0,$$

with $[L \quad L_l]^T$ a stabilizing observer gain for the series connection of plant and internal model. According to the separation principle, K and $[L \quad L_l]^T$ can be designed independently. By investigating the solutions (5) and (9), it is directly verified that the repetitive controller and the learning controller are related by duality.

3.3 Design philosophy

Application of the internal model framework to the concepts of repetitive and learning control provides insight into two fundamental aspects of these concepts, which have not been discussed before. First, it gives necessary and sufficient conditions for the existence of a solution to the periodic control problem, which guarantee the controllability and observability of the series connection of plant and internal model. For a MIMO plant, these conditions provide insight into the choice between a repetitive or a learning controller, *i.e.* if the system has more inputs than outputs, a repetitive controller should be used for both tracking and disturbance rejection, whereas for a system with more outputs than inputs a learning controller should be used for rejection of disturbances at the system input; in this case, tracking of r is in general not possible, unless condition (8) applies. For square systems, either a repetitive or a learning controller can be used for tracking and disturbance rejection, however, for learning control the reference r should be available for compensation, which is not necessary for repetitive control. Second, once the existence of a solution has been verified, the *design* of a repetitive or learning controller is nothing more than the design of a stabilizing compensator for the series connection of plant and internal model. In general, any technique can be used for this design, *e.g.* observer-based state feedback (LQG, H_2), or robust control oriented techniques like QFT, H_∞ , or μ -synthesis. The final compensator then consists of the series connection of in-

ternal model and stabilizing compensator, compare (3), (6). In the next section we show that several existing repetitive and learning control approaches can be explained from the internal model framework by a particular choice of the internal model. Consequently, the same design philosophy applies to these approaches as well, *i.e.* synthesis of a repetitive or learning controller is nothing more than the choice of a stabilizing compensator for the series connection of the plant and a particular internal model.

4 Connection with existing repetitive and learning control schemes

For simplicity, in this section only tracking of a periodic reference signal r is considered. Without loss of generality we assume that the system under consideration is stable, which gives a bit more insight into the discussed mechanisms, since equations become more simple.

4.1 Current-error feedback versus past-error feedforward

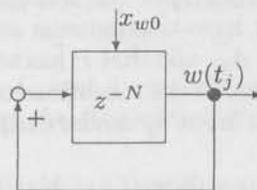


Fig. 3: Discrete time internal model with delay element in forward path.

In practical applications, the number of delays N in the internal model can be rather large, and consequently the dimension of the repetitive control gain in (3) can be rather high, although the internal model itself has a rather simple structure. In the literature, often an alternative stability analysis is used for the repetitive control system (3), see for example Tomizuka *et al.* (1989) and Hara *et al.* (1988), where the delay chain is put into the forward path of the controller, see Figure 3. Hence, the internal model changes to:

$$\phi_w^{-1} = \frac{z^{-N}}{1 - z^{-N}}. \quad (10)$$

In fact, the internal model (10) delays the direct transmission from the current error to the input u_p . In the literature on iterative learning control, schemes based on the model (2) are called *current-error feedback*, see for example Owens (1993) and Goh and Yan (1996), while schemes based on the

model (10) are called *past-error feedforward*, see for example Amann *et al.* (1996), Padieu and Su (1990) and Moore *et al.* (1992). The latter name indicates the open loop nature of the control input u_p during a trial. In the theory on repetitive control, also both schemes can be found, see for example Hillerström (1994) for current-error feedback, and Tomizuka *et al.* 1989 and Hara *et al.* (1988) for past-error feedforward.

The stability analysis suggested in the literature, now proceeds with isolating the delay chain of the internal model (10) in an *equivalent* system representation, (Hara *et al.*, 1988). For the repetitive controller (3), this is shown in Figure 4. Since the

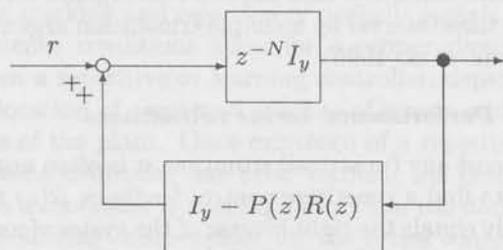


Fig. 4: Equivalent system representation of repetitive control system.

delay chain has magnitude equal to one, the small gain theorem can be used, which states that the following condition is *sufficient* for the equivalent system to be stable:

$$\|I_y - P(z)R(z)\|_i < 1, \quad (11)$$

for some induced i -norm. Equation (11) motivates to choose the repetitive control gain as $R(z) = P^{-1}(z)$, *i.e.* equal to the (right) inverse of the system $P(z)$; consequently, the dimension of R is now determined by the system P , and not by the number N of the internal model anymore. Note the difference with current-error feedback schemes, which require

$$\|(I_y - P(z)R(z))^{-1}\|_i < 1, \quad (12)$$

for closed loop stability, *i.e.* $R(z)$ should be high-gain. This shows the advantage of past-error feedforward schemes over current-error feedback schemes: the frequency up to which (11) holds in practical situations, is in general 2-3 times larger than the frequency up to which (12) is valid, see for example de Roover *et al.* (1996).

Although the use of the small gain theorem may yield low-dimensional repetitive gains, the resulting controller might be overly conservative in a sense that there might exist no $R(z)$ for which condition (11) or (12) hold, although condition (4) might still

be true, *i.e.* there still might exist a solution to the robust periodic control problem. This result is formalized in the following:

Corollary 4.1 Let $\{A, B, C, D\}$ be a minimal realization of $P(z) \in \mathbb{R}(z)^{y \times u}$, and let (2), ((10), respectively) be an internal model of (1), then condition (4) is true if condition (11) (condition (12), respectively) is true.

This result shows the power of the design philosophy based on the internal model framework: one can choose either (2) or (10) as internal model, and design a stabilizing compensator for the series connection of this specific model and the plant. In de Roover (1996) an alternative synthesis technique is proposed—based on the small gain theorem—which directly minimizes (11) using an H_∞ approach. With this approach, also model uncertainty can be easily incorporated into the design, and a stabilizing compensator can be computed with a μ -synthesis.

4.2 Asymptotic versus finite time tracking

The internal model framework of Section 3 is asymptotic in nature, *i.e.* theoretically it takes infinite time before the tracking and disturbance rejection objective are achieved. In practice, however, tracking and disturbance rejection are achieved whenever the servo error e has settled within some user-defined bounds, which is definitely not at infinity. This renders the infinite time framework to be valid, even for finite time objectives. For repetitive control, this is not an issue, since the output of the repetitive controller acts continuously as a control input. However, iterative learning control deals with batch-wise updating of the control input u_p after each successive trial of the reference r , where the duration of one batch—the length of one trial—is of finite time. Consequently, the demand on asymptotic stability of the controlled system is replaced by the demand on asymptotic convergence of the control signal to some fixed signal. This updating process can be derived from the internal model based scheme. If we progress on the dual repetitive control system and substitute (10) into (6), we obtain for the plant input signal:

$$\begin{aligned} u_p(t_j) &= (\phi_w(z)^{-1} I_u) L(z) e(t_j) \\ &= \left(\frac{z^{-N}}{1 - z^{-N}} I_u \right) L(z) e(t_j) \Leftrightarrow \\ u_p(t_j) &= (z^{-N} I_u) (u_p(t_j) + L(z) e(t_j)). \end{aligned} \quad (13)$$

Now suppose the reference signal of length N is repeated an indefinite number of times, and at each new trial of the reference the initial output of the

system is reset to the initial value of the reference, i.e. $y(kN) = r(0)$ where $k = 0, 1, 2, \dots$ denotes the number of trials. Then the control input (13) can be written as:

$$u_p^{k+1}(t_j) = u_p^k(t_j) + L(z)e^k(t_j), \quad (14)$$

which shows that the input signal is updated after each trial of the reference signal on the basis of the tracking error. As a matter of fact, Equation (14) is a general update law used in many past-error feedforward learning control schemes, see for example Horowitz (1993), Kavli (1992) and Moore *et al.* (1992). Due to this discretization of the repetitive learning process, no necessary condition for stability of the learning system can be derived. However, in Figure 5 the equivalent system representation is

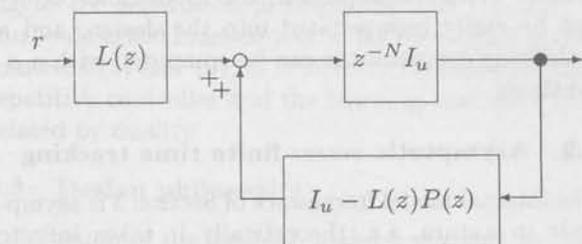


Fig. 5: Equivalent system representation of dual repetitive control system.

given. Using the small gain theorem for this system, a sufficient condition for convergence is given by:

$$\|I_u - L(z)P(z)\|_i < 1, \quad (15)$$

for some induced i -norm, which is dual to (11). This convergence analysis applies equally well to the update law (14) as to the stability analysis of the input signal (13), see for example Moore *et al.* (1992) and de Roover (1996). Likewise, by substituting (2) into (6), the following update law is derived for current-error feedback schemes, see for example Owens (1993), Goh and Yan (1996) and Amann *et al.* (1996):

$$u_p^{k+1}(t_j) = u_p^k(t_j) + L(z)e^{k+1}(t_j).$$

Again using the small gain theorem, a sufficient condition for convergence is given by:

$$\|(I_u - L(z)P(z))^{-1}\|_i < 1. \quad (16)$$

Dual to Corollary 4.1, we claim that the analysis and synthesis procedure for the gain L , based on the small gain theorem might be conservative with respect to the general results presented in Section 3:

Corollary 4.2 Let $\{A, B, C, D\}$ be a minimal realization of $P(z) \in \mathbb{R}(z)^{y \times u}$, and let (2), ((10), respectively) be an internal model of (1), then condition (7) is true if condition (15) (condition (16), respectively) is true.

It should be noted that the use of frequency domain expressions, like (15) and (16), intrinsically presume an infinite trial length in the learning process, which seems only valid for theoretical and not for practical purpose. However, if the trial length is long with respect to the dynamics of the system and the transient behaviour of the reference signal, this assumption seems justified. Moreover, if the system is causal on each trial, an infinite time convergence condition also implies convergence on any finite time interval by a simple truncation argument (Amann *et al.* 1996).

4.3 Performance versus robustness

In almost any (practical) situation, it is often impossible to find a repetitive control feedback $R(z)$ that exactly equals the right inverse of the system (condition (15)), either because this inverse does not exist, for example if P is strictly proper or behaves non-minimum phase, or because only an approximate description of P is available. Therefore, in the literature, a modified repetitive control system is suggested, which uses an internal model of the form:

$$\phi_m^{-1}(z) = \frac{q(z)z^{-N}}{1 - q(z)z^{-N}}, \quad (17)$$

with $q(z)$ being a (low order) low pass filter with magnitude equal to one at low frequencies, see also Tomizuka *et al.* (1989) and Hara *et al.* (1988). The idea is to choose the bandwidth of $q(z)$ up to a value of z where $R(z)$ is still a good approximation of $P^{-1}(z)$. The filter $q(z)$ provides an easy way to tune robustness of the closed loop to high frequent model errors, at the cost of a nonzero tracking error at high frequencies. Likewise, the model (2) is modified to $\phi_m(z) = 1 - q(z)z^{-N}$, because in general it is impossible to fulfil condition (12) for all values of z . In the literature on learning control systems, the same (dual) synthesis procedure can be found for the gain $L(z)$, see for example Kavli (1992) and de Roover (1996). In this situation, $L(z)$ should be a close approximation of the left inverse of the system $P(z)$. Due to the same practical limitations of this choice, a low-pass filter $Q(z)$ is used, which changes the update law (14) to

$$u_p^{k+1}(t_j) = Q(z)(u_p^k(t_j) + L(z)e^k(t_j)),$$

and consequently changes the sufficient condition for stability to

$$\|Q(z)(I_u - L(z)P(z))\|_i < 1,$$

which increases the robustness of the learning system to high frequent model errors, at the cost of a nonzero tracking error at high frequencies. Naturally, the same argument applies for current-error feedback learning control schemes, (see Goh and Yan, 1996; Amann *et al.*, 1996).

5 Conclusions

This paper gives a general framework for the analysis and design of repetitive and learning controllers, explicitly derived from results available for the internal model principle. The internal model framework gives necessary and sufficient conditions for existence of a solution to the problem of robust asymptotic tracking and rejection of periodic signals. The existence conditions allow for a proper choice between a repetitive or learning controller, dependent on location of zeros and number of inputs and outputs of the plant. Once existence of a repetitive or learning controller has been verified, the design of such a controller is nothing more than the design of a stabilizing compensator for the series connection of the plant and an internal model of the periodic signal, using any design technique. It is shown that a number of existing repetitive and learning control approaches can be put into this framework by making a specific choice of the internal model. Consequently, the analysis and design of these approaches can be generalized to the powerful analysis and design procedure of the internal model framework.

References

- Amann, N., D.H. Owens, E. Rogers and A. Wahl (1996). An H_∞ approach to linear iterative learning control design. *Int. J. of Adaptive Control and Signal Processing*, 10, 767-781.
- Arimoto, S., S. Kawamura and F. Miyazaki (1984). Bettering operation of robots by learning. *J. Robotic Systems*, 1, 123-140.
- Francis, B.A. and W.M. Wonham (1975). The internal model principle for linear multivariable regulators. *J. Applied Mathematics and Optimization*, 2, 175-194.
- Goh, C.J. and W.Y. Yan (1996). An H_∞ synthesis of robust current error feedback learning control. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 118, 341-346.
- Hara, S., Y. Yamamoto, T. Omata and M. Nakano (1988). Repetitive control system: a new type servo system for periodic exogenous signals. *IEEE Trans. Automat. Contr.*, AC-33, 659-668.
- Hillerström, G. (1994). *On Repetitive Control*. Dr. Dissertation, Control Engineering Group, Dept. of Computer Science and Electrical Engineering, Lulea Univ. of Techn., Lulea, Sweden.
- Horowitz, R. (1993). Learning control of robot manipulators. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 115, 402-411.
- Inoue, T., M. Nakano, T. Kubo, S. Matsumoto and H. Baba (1981). High accuracy control of a proton synchrotron magnet power supply. *Proc. 8th IFAC World Congress*, vol. 6, pp. 3137-3142.
- Kavli, T. (1992). Frequency domain synthesis of trajectory learning controllers for robot manipulators. *J. Robotic Systems*, 9, 663-680.
- Moore, K.L., M. Dahleh and S.P. Bhattacharyya (1992). Iterative learning control: a survey and new results. *J. Robotic Systems*, 9, 563-594.
- Owens, D.H. (1993). Iterative learning control using adaptive high gain feedback. *Proc. 1993 European Control Conf.*, pp. 195-198.
- Padieu, F. and R. Su (1990). An H_∞ approach to learning control systems. *Int. J. Adapt. Contr. and Signal Proc.*, 4, 465-474.
- Roover, D. de (1996). Synthesis of a robust iterative learning controller using an H_∞ approach. *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, pp. 3044-3049.
- Roover, D. de, O.H. Bosgra, F.B. Sperlina and M. Steinbuch (1996). High-performance motion control of a flexible mechanical servomechanism. In: O.H. Bosgra *et al.* (Eds.), *Selected Topics in Identif., Modelling and Control*. Delft University Press, Vol. 9, pp. 69-78.
- Roover, D. de, and O.H. Bosgra (1997). Dualization of the internal model principle in compensator and observer theory with application to repetitive and learning control. *Proc. American Control Conf.*, Albuquerque, New Mexico, USA, pp. 3902-3906.
- Sadegh, N. (1991). Synthesis and stability analysis of repetitive controllers. *Proc. American Control Conf.*, Boston, Massachusetts, USA, pp. 2634-2639.
- Tomizuka, M., T.-C. Tsao and K.-K. Chew (1989). Analysis and synthesis of discrete time repetitive controllers. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 111, 353-358.

Point-to-point control of a high accuracy positioning mechanism[†]

D. de Roover^{§§} and F.B. Sperling[#]

[§] *Mechanical Engineering Systems and Control Group
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.*

E-mail: d.derover@wbmt.tudelft.nl

[#] *Philips Research, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands*

E-mail: sperling@natlab.research.philips.com

Abstract. This paper compares three open-loop command generating methods for the point-to-point control of a wafer stage. Whereas the theoretical time-optimal bang-bang force command results in excessive residual vibration due to excitation of resonant modes, shaped force profiles have the property not to excite these modes while still allowing for fast movements. The most successful shaping method turned out to be the one which was best able to handle uncertainty in our knowledge of the resonant mode dynamics, and the way this knowledge is used to shape the spectral content of the force command. Besides comparing three shaping methods, it is also investigated how the feedback and feedforward compensator affect the point-to-point motion.

Keywords. point-to-point control, input shaping, mechanical servomechanism, residual vibration.

1 Introduction

In order to be competitive, modern chip manufacturing machines are required to perform both fast and accurately. In particular the servomechanism performing the positioning of chips, is required to move the chips as fast and as accurately as possible. Due to the inherent flexibility of the mechanical construction of most positioning devices, these performance requirements are conflicting, *i.e.* the faster the system moves, the less accurate it will be, due to large vibrations induced by fast movements and large acceleration forces.

To overcome this problem, a high-performance motion control system can be designed, which provides steering and regulation of the positioning mechanism. In general, a motion control system of a ser-

vomechanism has three degrees of freedom, see Figure 1. One way to cope with the flexibilities, is to design a feedback controller, C , which attenuates the induced vibrations in closed-loop. Because of stability requirements, a major drawback of this approach is the limited use of high gain inside the loop, in particular in the high frequency range where the flexible dynamics are more likely to occur. Therefore, in this paper, in addition to the closed-loop, an open-loop approach is followed at set-point level: suitable motion forces, f , and reference trajectories, r , are designed, which minimize vibration of flexible dynamics at the end of a movement. In the literature, a large number of techniques is presented which solve this so-called *point-to-point* control problem with minimal residual vibration, see for example Bhat and Miu (1990), Meckl and Seering (1985), Singer and Seering (1990), and the references therein.

To speed up tracking of r , the motion force f and the reference r are often related by a *feedforward com-*

[†]This paper was presented at the 1997 American Control Conference, June 4-6, 1997, Albuquerque, New Mexico, USA. Copyright of this paper remains with IEEE.

[§]The work of Dick de Roover is financially supported by Philips Research Laboratories, Eindhoven, The Netherlands.

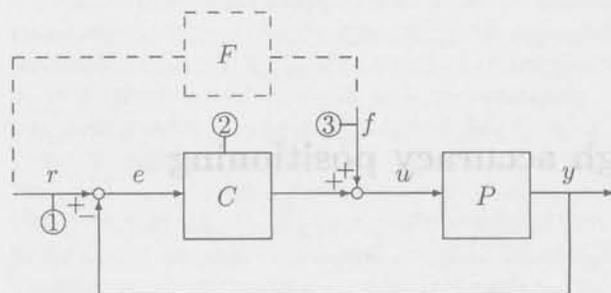


Fig. 1: General 3 degree-of-freedom (3-DOF) motion control system; P , C , f , r , u , y , e , denote the plant, feedback compensator, force input, output reference, system input, system output, and tracking error, respectively.

compensator, F (see dashbox in Figure 1). Because of the closed-loop implementation of f and r , synthesis of these signals is not trivial: one can either start with designing r and derive f according to $f = F_i r$, with F_i some *inverse* compensator which may range from a series of differentiators – simply relating position to force – to complex models describing the inverse dynamics of the system P , see for example Kwon and Book (1990), Moulin and Bayo (1991) and Devasia (1996), or design f and obtain r via $r = F_f f$, with F_f some *forward* compensator, e.g. a series of integrators or a model of the dynamics of P – see for example Meckl and Kinceler (1994), Singer and Seering (1990) and Bhat and Miu (1991). Clearly, the choice of F_i or F_f influences the tracking of r . Which approach is favourable, depends on the given application. In this paper we follow the latter one, because the inverse dynamics of P are unstable.

The goal of this paper is twofold. First, three different open-loop force command generating methods are compared: a standard approach, which limits the jerk of the acceleration profiles to minimize residual vibration, is compared to two modern approaches, which use knowledge of the system's resonant frequencies to minimize residual vibration. Second, it is experimentally analyzed how the feedback compensator C and feedforward compensator F contribute to reduction of the tracking error e . To this end, the next section describes the modelling and feedback control of the experimental set-up, as a starting point for subsequent sections. In Section 3, the point-to-point control problem of this particular system is mathematically formulated, and a three-step solution is proposed. Next, Section 4 shows the experimental results of three different command generating methods, and Section 5 experimentally

analyses the influence of C and F on the error e .

2 The wafer stage experimental set-up

Figure 2 shows a schematic view of a prototype wafer stage experimental set-up, used for the experiments shown in this paper. A wafer is placed

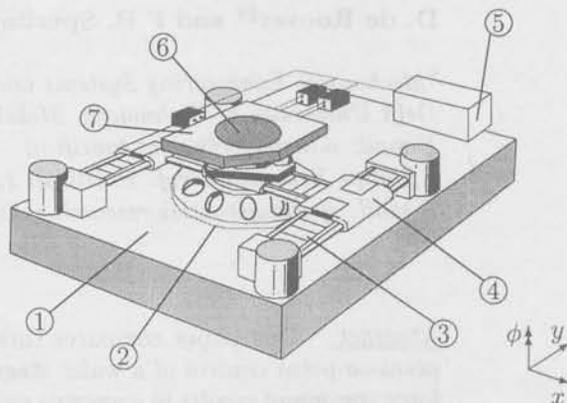


Fig. 2: Schematic view of a wafer stage; 1: granite block, 2: airfoot, 3: stator, 4: translator, 5: laser interferometer, 6: wafer chuck, 7: mirror block.

on a wafer chuck in the middle of the stage. The stage, consisting of airfoot, translator, wafer chuck and mirror block, is driven by the linear motor in x -direction. The stator of that motor is fixed to the translators of two other linear motors, hence driving the stage in y -direction; by driving these two motors independently, also a slight rotation ϕ of the stage is possible, which freedom is used to correct for misalignment of the wafer. The position of the stage in the horizontal plane is measured with three laser interferometers, one in x -direction and two in y -direction; the measurement resolution is approximately 13 *nanometre* [nm]. Thus, the positioning system is multivariable, having three actuators and three sensors, henceforth denoted as *inputs* and *outputs*, respectively. For sake of clarity, in this paper only results from one input to one output are shown. After alignment, the wafer is alternately stepped and exposed according to a certain prescribed pattern. In this paper we only consider stepping of the wafer from one position to another.

In Callafon *et al.* (1996) an experimental model was derived for the stage, using frequency-domain identification techniques; a linearly parametrized time-invariant model, denoted \hat{P} , was fitted to a frequency response of the system, estimated with a Hewlett-Packard signal analyzer. Figure 3(a) shows the computed frequency response, together with the resulting 16th order model fitted to this response. In

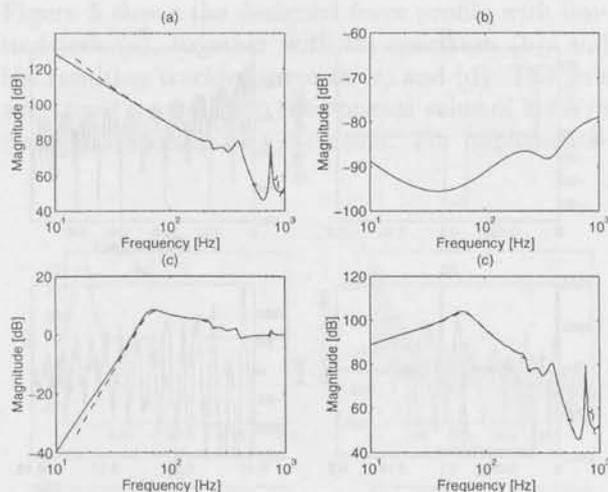


Fig. 3: Bode magnitude of (a) 16th order model \hat{P} (solid) and estimated frequency response (dashed), (b) feedback controller C , (c) resulting output sensitivity S , and (d) plant multiplied by sensitivity $S\hat{P}$.

the fit procedure, extra weights were applied which emphasized the mid-frequency range, important for control design. This figure shows a typical response of a general mechanical servomechanism: at low frequencies, the response has the shape of a double integrator according to Newton's 2nd law, and at middle and high frequencies some resonances show up due to the flexible components.

Because the system is marginally stable (double integrator), first a feedback controller is designed to stabilize the system; besides, feedback control enables the suppression of external disturbances, like electric actuator noise, and enhances robustness against modelling errors. In de Groen (1996) a MIMO feedback design has been performed for the wafer stage using QFT. Figure 3(b) shows a magnitude Bode plot of the resulting controller, and Figures 3(c) and 3(d) show a magnitude Bode plot of the resulting output sensitivity transfer function S – defined as $S = (I + \hat{P}C)^{-1}$ – and S multiplied by \hat{P} , respectively, which is the transfer function from f to e . Note that to some extent damping of the open loop resonant frequencies has been increased in this closed loop.

3 Point-to-point control problem

3.1 Problem formulation

The task of a wafer stage is to move a wafer from one chip position to another, typically a distance of 10 mm. Hence, the point-to-point control problem

for a wafer stage can be formulated in terms of actuator inputs, u , and sensor outputs, y , which we assume to be the real system inputs and outputs respectively:

Design a control $u(t), t \in [t_0, t_1]$, which moves the stage from $y(t_0)$ to $y(t_1)$, with $y(t_1) - y(t_0) = 1e-2$ m, such that $t_1 - t_0$ is minimal, and the following constraints are satisfied:

$$\begin{aligned} \dot{y}(t) &\leq 0.12 \text{ m/s} \quad \forall t \in [t_0, t_1], \\ u(t) &\leq 62 \text{ N} \quad \forall t \in [t_0, t_1], \\ \dot{u}(t) &\leq 5.0e4 \text{ N/s} \quad \forall t \in [t_0, t_1], \end{aligned} \quad (1)$$

$$y(t) = y(t_1) \pm 52 \text{ nm} \quad \forall t > t_1, \quad (2)$$

Note that constraints (1) reflect the limitations of sensors and actuators, while constraint (2) is a performance demand in the end-point; in this light, constraint (2) serves as a criterion for a successful move.

3.2 Three-step solution

To solve this problem, we suggest the following three-step approach:

Step 1: Obtain a linear model, denoted \hat{P} , which describes the relevant system dynamics, such that $y \approx \hat{P}u$.

Step 2: Design a pair $\{f, r\}$ such that $f(t)$ fulfils the constraints (1) on $u(t)$, and $r(t)$ fulfils the constraints (1),(2) on $y(t)$.

Step 3: Implement the pair $\{f, r\}$ in the 3-DOF nominal tracking scheme of Figure 1.

Step 1 has been carried out in Section 2, resulting in the 16th order model shown in Figure 3(a). Step 2 concerns the choice of an input design method – which will be the subject of the next subsection – and the choice of a feedforward compensator F . The actual implementation of Step 3 will be the topic of Section 4.

3.3 Different input design methods

To accomplish Step 2, three different open-loop command generating methods are used. The first, somewhat heuristic method, considers a so-called 'bang-bang' input signal with limited slope, see for example Lewin (1994) and Miu and Bhat (1991). The idea is to minimize the high frequency content of the input signal by decreasing the slope of this signal. As a consequence, the transition times between the zero and peak input levels are increased, resulting in a smoother command signal of longer duration. The method is easy to apply since no knowledge of the flexible dynamics is required at

all. The optimal jerk is obtained by manual tuning with cycle time as criterion.

Two other methods are investigated which do require some knowledge of the dynamics of the flexibilities. One method concerns the design of a finite impulse response (FIR) filter, to preshape an existing command signal (see Singer and Seering, 1990; Singhose *et al.*, 1995). The idea is to synthesize an FIR filter which removes the energy contribution of the command signal at the system resonant frequencies. The FIR filter has to be convolved with an existing command signal, for example a time optimal bang-bang input, and preserves its vibration reducing properties after convolution. The knowledge required to use this method, is the location of the natural frequency of the flexibility together with its damping ratio, *i.e.* the location of the complex poles of a 2nd order system describing the flexibility. In Bhat and Miu (1990) it was shown that the FIR filter has the Laplace domain interpretation of placing zeros at the locations of the resonant poles. The other method concerns the synthesis of a series of 'versines', which approximate a bang-bang command signal with limited velocity (see Meckl and Seering, 1988). These versine functions allow energy removal from the input signal in a narrow band surrounding the system natural frequencies. This technique neglects the damping of the resonant frequencies, hence only minimizing the energy contribution of the input signal at the natural frequencies in the Fourier domain; consequently, the only knowledge of the system required, are the locations of the natural frequencies of the resonant modes.

4 Experimental results

To illustrate the need for carefully designed input signals, Figure 4 shows a bang-bang command (a) together with its spectrum (b), and the resulting tracking error e after implementation of this signal in the set-up of Figure 1, (c), (d). The residual vibration in the endpoint – a consequence of the excitation of all flexible modes – is clearly visible. These experimental results were obtained with the forward compensator F_f being a simple double integrator with time-delay, *i.e.* :

$$F = F_f = \frac{e^{-sT_d}}{Ms^2}, \quad (3)$$

with $T_d = 1.15\Delta T$ denoting the delay-time, and $\Delta T = 3e-4s$ being the sampling time of the controller; $M = 10.4kg$ denotes the mass of the stage. Besides, to compensate for viscous friction during stage movement at maximum velocity, a velocity dependent term has been added to the actual control signal. The actual control signal, which has to be

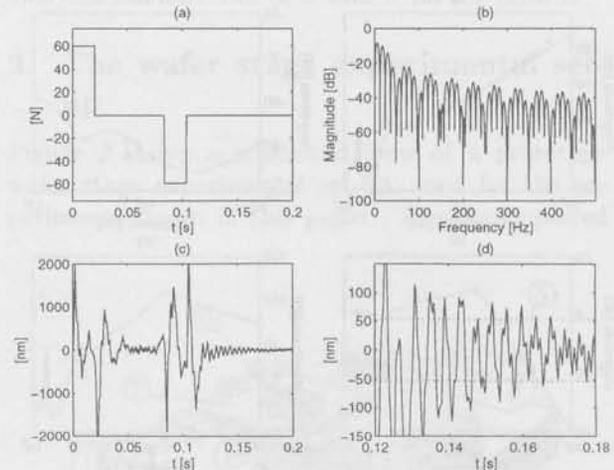


Fig. 4: Command response bang-bang force profile for 10 mm step in x -direction; (a) force profile, (b) normalized spectrum of force profile, (c) tracking error, (d) tracking error with performance bounds during settling.

scaled within a ± 1 range, is therefore given by:

$$u(t) = 0.1 \left(\frac{1}{M} + \frac{1}{Vs} \right) f(t), \quad \forall t \in [t_0, t_1], \quad (4)$$

with $V = 3.83kgs$ denoting a constant velocity gain. To systematically design proper signals f and r , we first analyzed the resonant mode properties of the 16th order model, see Table 1. It is rather surprising

ω_n [Hz]	ζ	A_y [%]
1.42e2	2.49e-2	23.3
1.55e2	2.46e-2	73.6
2.28e2	4.50e-2	2.14
3.03e2	5.97e-2	0.66
3.98e2	6.12e-2	0.26
7.57e2	3.19e-3	0.03
1.10e3	2.33e-2	0.01

Table 1: Resonant Mode properties of 16th order model; ω_n denotes natural frequency, ζ denotes relative damping, and A_y denotes relative static error in the output after a step input.

that two hardly visible resonance modes near 150Hz are responsible for about 97% of the static error of residual vibration at the system output after a step input. Therefore, we decided to design command signals which minimize residual vibration stemming from these low frequent modes.

Figure 5 shows the designed force profile with limited jerk (a), together with its spectrum (b), and the resulting tracking error in (c) and (d). The jerk was tuned manually to the optimal value of $5e3N/s$ for a 10 mm step in x -direction. For implementa-

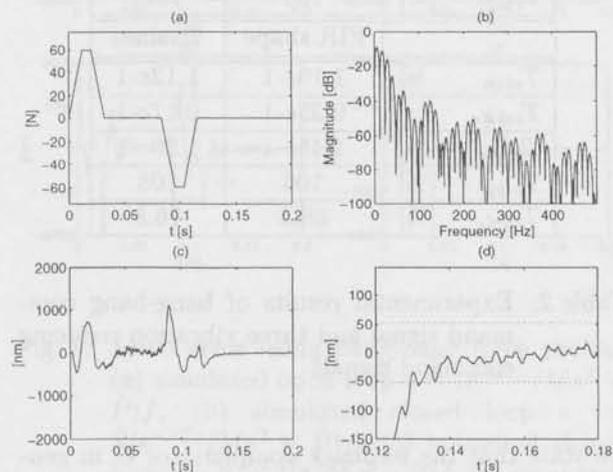


Fig. 5: Command response limited jerk force profile for 10 mm step in x -direction; (a) force profile, (b) normalized spectrum of force profile, (c) tracking error, (d) tracking error with performance bounds during settling.

tion we used the same forward compensator (3) and friction compensation (4). Clearly, the residual vibration has been removed from the error, compare Figure 4(c),(d) with Figure 5(c),(d). This can be explained from the spectrum of the limited jerk signal, which has small values at the system resonant modes. Although the method is easy to apply, it might be overly conservative in a sense that energy has been removed from the input spectrum, not only at frequency locations of resonant modes, but also at other frequencies. Therefore, also two model-based input design techniques are considered.

Figure 6 shows the results of the FIR preshape method. We designed one series of three pulses (robust ZVD shaper (see Singhose *et al.*, 1995) for a fictitious resonance at 150Hz with relative damping $\zeta = 2.47e-2$, and convolved this pulse series with the bang-bang command of Figure 4(a). Concluding from the input spectrum, the 150Hz component has been completely removed, while at other frequencies the spectrum is equal to the original spectrum of the bang-bang input, preserving a relatively fast force command. However, still residual vibration is present, stemming from the resonant mode at 228 Hz. Apparently, the force input has too much energy at this frequency.

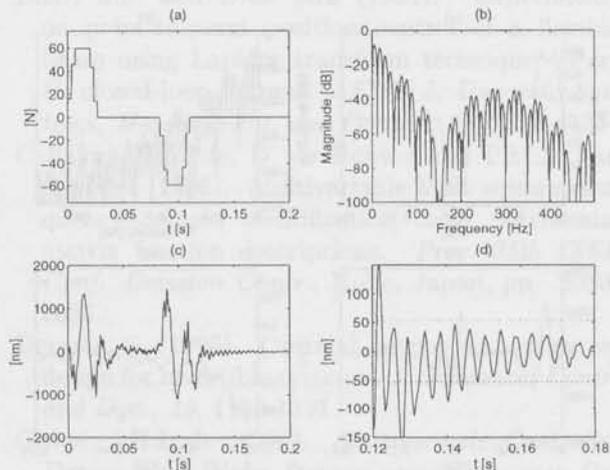


Fig. 6: Command response FIR preshaped force profile for 10 mm step in x -direction; (a) force profile, (b) normalized spectrum of force profile, (c) tracking error, (d) tracking error with performance bounds during settling.

Figure 7 shows the synthesized versine series. Using the software described in Meckl and Seering (1988), we designed a series of 5 versine basis functions with a notch surrounding 5% of a fictitious resonance mode at dimensionless frequency $150 \cdot 0.02 = 3$. Whereas the FIR preshape method *removes* energy from a given spectrum, the versine series *adds* energy to its spectrum by expanding more basis functions into its series; in our case, a total of 5 basis functions adds energy to the signal up to approximately 200 Hz. As a result, a relatively fast force command is obtained, with the property of suppressing almost all of the original residual vibration, compare Figures 4 and 7.

Table 2 quantitatively summarizes the obtained results. In this table, T_{step} denotes the duration of the command signal, T_{settle} denotes the time between T_{step} and the point at which the tracking error has settled within the performance band around the end-point, and T_{cycle} is the sum of T_{step} and T_{settle} . Good results are obtained with the heuristic limited jerk approach, but things can be improved using more sophisticated model-based techniques. In this case, the series of versines performed best, because it was best able to shape the spectrum of the force input. Although the FIR preshape method preserved the fastest force command, extra shaping near 228Hz is necessary to suppress remaining vibration, and consequently the force command will slow down.

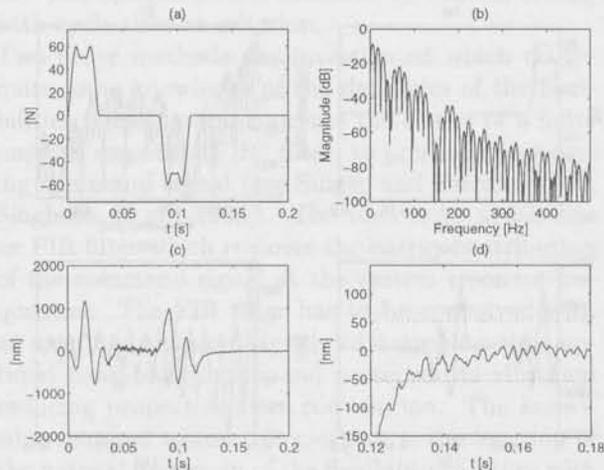


Fig. 7: Command response versus force profile for 10 mm step in x -direction; (a) force profile, (b) normalized spectrum of force profile, (c) tracking error, (d) tracking error with performance bounds during settling.

5 Role of feedback and feedforward compensator

In the previous section we have seen the success of smart input design. An important question which remains is: *What is the role of the feedback and feedforward compensators C and F , respectively?* This will be addressed in this section.

To investigate the role of feedback in the suppression of residual vibration, we compared open-loop and closed-loop simulations. Consider the general configuration of Figure 1. Clearly, if the forward compensator F is equal to the plant P , the resulting tracking error e is equal to zero for all t . Hence, if F is equal to the simple model (3) – which we used in our experiments – a non-zero tracking error is to be expected resulting from the resonant modes, which will be regulated by the feedback compensator C . Figure 8(a) shows the simulated open-loop difference between the simple model (3) and the complex 16th order model shown in Figure 3(a), convolved with the bang-bang force shown in Figure 4(a). Clearly, the effect of the flexible modes is visible in the error. Moreover, the feedback compensator will respond to this error according to the shape of the sensitivity function S , shown in Figure 3(c), *i.e.* for frequencies where $|S(i\omega)| < 0$ dB, the open-loop error will be suppressed, but for frequencies where $|S(i\omega)| > 0$ dB, the open-loop tracking error will be amplified. Since the maximum obtainable bandwidth¹ is limited to approximately 100Hz,

¹We denote the bandwidth as the first value of ω where

		bang-bang	lim. jerk
T_{step}	[s]	1.04e-1	1.16e-1
T_{settle}	[s]	0.64e-1	0.16e-1
T_{cycle}	[s]	1.68e-1	1.32e-1
T_{step}	[%]	100	112
T_{cycle}	[%]	100	78.5
		FIR shape	versines
T_{step}	[s]	1.10e-1	1.12e-1
T_{settle}	[s]	0.35e-1	0.17e-1
T_{cycle}	[s]	1.45e-1	1.29e-1
T_{step}	[%]	106	108
T_{cycle}	[%]	86.3	76.8

Table 2: Experimental results of bang-bang command signal and three vibration reducing command signals.

we state that the feedback compensator C in general *works against* the suppression of residual vibration. This is confirmed, both by simulations and measurements, see Figure 8(b) and (c), respectively. Although there is still some discrepancy between the modelled and measured settling behaviour, there is a remarkable resemblance between the closed-loop simulation and measurement, which amplifies our line of reasoning. The adverse effect of feedback in suppressing residual vibration, in addition, stresses the need for careful design of force commands.

This observation strongly motivates to replace the simple forward compensator (3) by the complex 16th order model, to avoid reaction of C to the open-loop tracking error. Experiments have been performed for all previous mentioned force commands, with F equal to the 16th order model \hat{P} . For the bang-bang force command, this experiment is shown in Figure 8(d). Although the error has been reduced to 50% of the original error during step time, the error during settling time has remained unchanged. This is explained from the fact that the filtered reference signal, $r = \hat{P}f$, was set to its final value after reaching the desired end-point, *i.e.* the system output y vibrates in the end-point, while the reference was fixed to the desired end-point value. This tendency also showed up in the experiments with shaped force commands. Note that the error during step time is not completely zero, which must be explained from the fact that the model \hat{P} is not a perfect match of the real plant P . Hence, the choice of a complex forward compensator F only affects the error during step time, but, sad to say, hardly

$|S(i\omega)| = 0$ dB

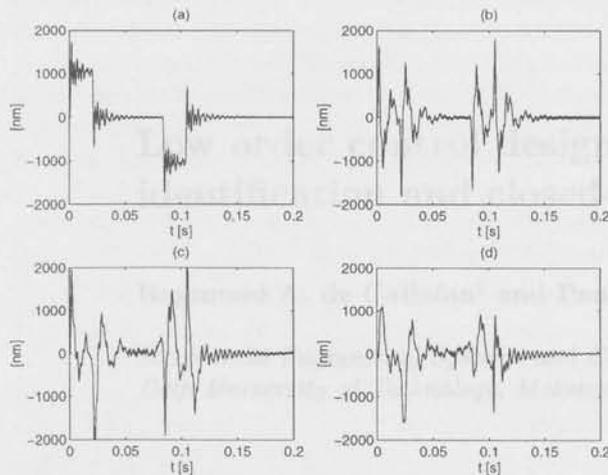


Fig. 8: Servo error using bang-bang force profile; (a) simulated open-loop $e = (e^{-sT_d}/Ms^2 - \hat{P})f$, (b) simulated closed loop $e = \hat{S}(e^{-sT_d}/Ms^2 - \hat{P})f$, (c) measured closed loop $e = S(e^{-sT_d}/Ms^2 - P)f$, (d) measured closed loop $e = S(\hat{P} - P)f$.

influences the error during settling.

6 Conclusions

In this paper we investigated the point-to-point control of a wafer stage. The key issue in obtaining fast force commands with vibration free movements, is the shaping of the spectrum of the force command only at those frequencies where the resonant modes are located. A comparison of three open-loop force command generating methods showed a preference for the model-based approximation of a bang-bang command with a series of versines, because it was best able to shape the force spectrum.

In addition, the role of the feedback and feedforward compensator in the suppression of residual vibration has been analyzed. Simulations and experiments showed the adverse effect of feedback in case the magnitude of the sensitivity is larger than one at flexible mode locations. The use of a high-order forward compensator instead of a simple approximation, improves the error suppression during motion, but hardly affects the error during settling.

References

Bhat, S.P. and D.K. Miu (1990). Precise point-to-point positioning control of flexible structures. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 112, 667-674.

Bhat, S.P. and D.K. Miu (1991). Experiments on point-to-point position control of a flexible beam using Laplace transform technique - Part II: closed-loop. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 113, 438-443.

Callafon, R.A. de, D. de Roover and P.M.J. Van den Hof (1996). Multivariable least squares frequency domain identification using polynomial matrix fraction descriptions. *Proc. 35th IEEE Conf. Decision Contr.*, Kobe, Japan, pp. 2030-2035.

Devasia, S. (1996). Optimal output trajectory redesign for invertible systems. *J. Guidance, Contr. and Dyn.*, 19, 1189-1191.

Groen, J.H.I. de (1996). *Multivariable Controller Design for a Wafer Stepper using Frequency Domain Techniques*. MSc. Thesis A-756, Mech. Eng. Systems and Control Group, Delft University of Technology, The Netherlands.

Kwon, D.-S. and W.J. Book (1994). An inverse dynamic method yielding flexible manipulator state trajectories. *Proc. American Control Conf.*, San Diego, CA, USA, pp. 186-193.

Lewin, C. (1994). Motion control gets gradually better. *Machine Design*, 90-94.

Meckl, P.H. and R. Kinceler (1994). Robust motion control of flexible systems using feedforward forcing functions. *IEEE Trans. Control Systems Technology*, CST-2, 245-254.

Meckl, P.H. and W.P. Seering (1988). Controlling velocity-limited systems to reduce residual vibration. *Proc. IEEE Int. Conf. on Robotics and Automation*, Philadelphia, PA, USA, pp. 1428-1433.

Meckl, P.H. and W.P. Seering (1985). Minimizing residual vibration for point-to-point motion. *Trans. ASME J. Vibration, Acoustics, Stress, and Reliability in Design*, 107, 378-382.

Miu, D.K. and S.P. Bhat (1991). Minimum power and minimum jerk position control and its applications in computer disk drives. *IEEE Trans. on Magnetics*, MAG-27, 4471-4475.

Moulin, H. and E. Bayo (1991). On the accuracy of end-point trajectory tracking for flexible arms by noncausal inverse dynamic solutions. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 113, 320-324.

Singer, N.C. and W.P. Seering (1990). Preshaping command inputs to reduce system vibration. *Trans. ASME J. Dynamic Systems, Measurement, and Control*, 112, 76-82.

Singhose, W.E., L.J. Porter and N.C. Singer (1995). Vibration reduction using multiple-hump extra-insensitive input shapers. *Proc. American Control Conf.*, Seattle, Washington, USA, pp. 3830-3834.

Low order control design by feedback relevant identification and closed-loop controller reduction[‡]

Raymond A. de Callafon[§] and Paul M.J. Van den Hof

*Mechanical Engineering Systems and Control Group,
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands*

Abstract. An approach is presented that can be used to obtain low complexity controllers for an unknown system. In this approach, the identification of a set of models is used to represent the incomplete knowledge of the system. Subsequently, the set is used for the synthesis of a robust controller. In order to design low complexity controllers, the aim is to find a low complexity representation of the set. Additionally, a closed loop reduction tool can be used to decrease the controller complexity further. This approach will be illustrated by an application to a multivariable positioning mechanism present in a wafer stepper.

Keywords. System identification; robust control; servomechanisms; multivariable systems.

1 Introduction

Industrial systems need feedback control to meet enhanced accuracy or performance requirements. In many applications the plant to be controlled is partly known, whereas limited complexity controllers are required due to hardware limitations. Both the inadequate knowledge of a plant to be controlled and the restriction on the complexity of the controller to be used makes the design of such a feedback controller a challenging task. In this paper, an approach is presented that can be used to obtain such low complexity (low order) linear feedback controllers for an unknown system.

To deal with the lack of information on the plant, the approach in this paper starts with the estimation of a set of models by means of system identification techniques, such that the unknown plant is an element of the set. Such a set of models is unavoidable

as the data used for identification purposes only represents a finite time, possibly disturbed, observation of the plant causing the knowledge of the plant to remain incomplete. As a consequence, a set of models consists of all models that are either validated (Ljung, 1987) or cannot be invalidated (Smith *et al.*, 1997) by the observations obtained from the plant.

Subsequently, a robust controller can be designed on the basis of this set of models. For that purpose, the set should be built up from a nominal model along with an allowable model perturbation (Boyd and Barrat, 1991). A general representation of a set of models can be written in terms of linear fractional transformation (LFT) based model perturbation (Boyd and Barrat, 1991). Such an LFT, based on a (dual) Youla-Kucera parametrization, is being estimated in this paper and shown to be particularly useful for both identification and control design purposes (de Callafon and Van den Hof, 1997).

To restrict the complexity (McMillan degree) of the controller, the aim is to estimate a low complexity representation of the LFT via an approximate identification. This is due to the fact that this LFT directly influences the order of a robust controller being computed (Boyd and Barrat, 1991; Zhou *et*

[‡]This paper was presented at the 2nd IFAC Symposium on Robust Control Design (ROCOND'97), 25-27 June 1997, Budapest, Hungary. Copyright of this paper remains with IFAC.

[§]The research of Raymond de Callafon was supported by the Dutch Institute of Systems and Control (DISC); current address: Department AMES, University of California at San Diego, La Jolla, CA 92093-0411, USA.

al., 1996). For further reduction of the controller complexity, an additional controller reduction can be employed. In this paper, a closed loop reduction that is based on the work by Ceton *et al.* (1993) is shown to be useful for reducing the complexity of the controller.

The subsequent steps of approximate identification of an LFT and the design of a robust controller followed by a closed loop controller reduction will be illuminated in this paper. To illustrate the approach, the application to a multivariable positioning mechanism present in a wafer stepper has been included.

2 Preliminaries

2.1 Norm-based feedback design

Let the notation P and C be used to denote finite dimensional, linear time invariant (FDLTI) (possibly unstable) systems, where C is used to indicate a controller. For notational convenience a control objective function is denoted by $J(P, C) \in \mathbb{RH}_\infty$ and the notion of performance will be characterized by the value of the norm $\|J(P, C)\|_\infty$: a smaller value of $\|J(P, C)\|_\infty$ indicates better performance (Van den Hof and Schrama, 1995).

A feedback connection of a system P and a controller C is denoted by $\mathcal{T}(P, C)$ and defined as the connection structure depicted in Figure 1. It is as-

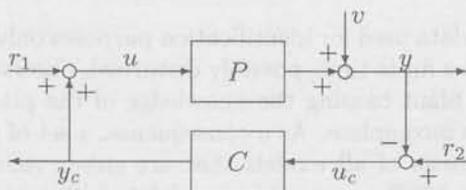


Fig. 1: Feedback connection structure $\mathcal{T}(P, C)$.

sumed that a connection $\mathcal{T}(P, C)$ is well posed, that is $\det(I + CP) \neq 0$ (Boyd and Barrat, 1991). The mapping from the signals $\text{col}(r_2, r_1)$ onto $\text{col}(y, u)$ is given by the transfer function matrix $T(P, C)$ with

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

Note that $\mathcal{T}(P, C)$ is internally stable if and only if $T(P, C) \in \mathbb{RH}_\infty$ (Schrama and Bosgra, 1993). In order to maintain generality, $J(P, C)$ is taken to be a weighted form of $T(P, C)$:

$$\|J(P, C)\|_\infty := \|U_2 T(P, C) U_1\|_\infty \quad (2)$$

where U_2 and U_1 are (square) weighting functions. The performance characterization (2) is fairly general and will be used for analysis purposes in this

paper. In this perspective, the performance objective function $J(P, C)$ as given in (2) will be used to evaluate both the identification of a set of models \mathcal{P} and the additional reduction of a robust controller designed based on the set \mathcal{P} . For that purpose, the set of models \mathcal{P} as used in this paper is discussed below.

2.2 Model uncertainty set

As indicated in Section 1, the incomplete knowledge of a plant P_o is represented by means of a set of models \mathcal{P} . An (upper) LFT

$$\mathcal{F}_u(Q, \Delta) := Q_{22} + Q_{21} \Delta (I - Q_{11} \Delta)^{-1} Q_{12} \quad (3)$$

provides a general notation to represent all models $P \in \mathcal{P}$ as follows

$$\mathcal{P}_i = \{P \mid P = \mathcal{F}_u(Q, \Delta) \\ \text{with } \Delta \in \mathbb{RH}_\infty \text{ and } \|\Delta\|_\infty < \gamma^{-1}\}$$

where Δ indicates an unknown (but bounded by γ^{-1}) uncertainty that reflects the incomplete knowledge of the plant P_o . The entries of the coefficient matrix Q in (3) indicate how the set of models \mathcal{P} has been structured, where $\hat{P} := \mathcal{F}(Q, 0) = Q_{22}$ denotes the nominal model of the set \mathcal{P} .

In this paper, the coefficient matrix Q is formed by employing the knowledge of any (possibly unstable) controller denoted by \hat{C} , that is used to form a stabilizing feedback connection $\mathcal{T}(P_o, \hat{C})$. In many practical situations, the presence of such a stabilizing controller \hat{C} is unavoidable due to instability of the plant P_o or additional safety requirements during operation.

Employing the knowledge of such a stabilizing feedback controller \hat{C} and using the algebraic theory of fractional model representations (Vidyasagar, 1985), the coefficient matrix Q in (3) is formed by considering a model perturbation that is structured similar to a (dual) Youla-Kucera parametrization:

$$\mathcal{P} = \{P \mid P = (\hat{N} + D_c \hat{\Delta})(\hat{D} - N_c \hat{\Delta})^{-1} \\ \text{with } \hat{\Delta} \in \mathbb{RH}_\infty \text{ and } \|\hat{V} \hat{\Delta} \hat{W}\|_\infty < \gamma^{-1}\} \quad (4)$$

where (N_c, D_c) and (\hat{N}, \hat{D}) respectively denote a right coprime factorization (rcf) of the controller \hat{C} and a nominal model \hat{P} , that satisfies $T(\hat{P}, \hat{C}) \in \mathbb{RH}_\infty$. \hat{V}, \hat{W} denote stable and stably invertible weighting functions used to normalize the upper bound on $\hat{V} \hat{\Delta} \hat{W}$ to γ^{-1} . It can be verified that the coefficient matrix Q in the LFT of (3) reads as follows.

$$Q = \left[\begin{array}{c|c} \hat{W}^{-1} \hat{D}^{-1} N_c \hat{V}^{-1} & \hat{W}^{-1} \hat{D}^{-1} \\ \hline (D_c + \hat{P} N_c) \hat{V}^{-1} & \hat{P} \end{array} \right] \quad (5)$$

It should be noted that in order to guarantee that $P_o \in \mathcal{P}$, additional prior information on the plant P_o must be introduced. This is due to the fact that $P_o \in \mathcal{P}$ cannot be validated solely on the basis of finite time, possibly disturbed, observations coming from the plant P_o (Mäkilä *et al.*, 1995; Ninness and Goodwin, 1995). Such information is in accordance with the uncertainty modelling procedure of Hakvoort (1994), that is used in this paper to bound the uncertainty $\bar{\Delta}$ in (4).

2.3 Evaluation of performance

The theory of fractional model representations provides a unified approach to handle both stable and unstable models and controllers within the set \mathcal{P} of (4). Additionally, the set \mathcal{P} has some favourable properties that can be illuminated by evaluating the performance objective function $J(P, C)$ for all $P \in \mathcal{P}$.

Lemma 2.1 Consider the set \mathcal{P} defined in (4) and a controller C such that the map $J(P, C) = U_2 T(P, C) U_1$ is well-posed for all $P \in \mathcal{P}$. Then

$$J(P, C) = \mathcal{F}_u(M, \Delta) \quad \forall P \in \mathcal{P}$$

where the entries of M are given by

$$\begin{aligned} M_{11} &= -\hat{W}^{-1}(\hat{D} + C\hat{N})^{-1}(C - \bar{C})D_c\hat{V}^{-1} \quad (6) \\ M_{12} &= \hat{W}^{-1}(\hat{D} + C\hat{N})^{-1} [C \ I] U_1 \\ M_{21} &= -U_2 \begin{bmatrix} -I \\ C \end{bmatrix} (I + \hat{P}C)^{-1}(I + \hat{P}\bar{C})D_c\hat{V}^{-1} \\ M_{22} &= U_2 \begin{bmatrix} \hat{N} \\ \hat{D} \end{bmatrix} (\hat{D} + C\hat{N})^{-1} [C \ I] U_1 \end{aligned}$$

Proof: By algebraic manipulation, see de Callafon and Van den Hof (1997). \square

It can be observed from (6) that substitution of $C = \bar{C}$ yields $M_{11} = 0$. This implies that when a controller C (equal to the controller \bar{C} used in the construction of the set \mathcal{P} in (4)) is applied to the set \mathcal{P} , stability robustness is satisfied, regardless of the value of γ in (4). This advantage, observed also by Sefton *et al.* (1990), is not shared by alternative uncertainty characterizations such as an open loop additive uncertainty description. Moreover, for $C = \bar{C}$ the upper LFT $\mathcal{F}_u(M, \Delta)$ modifies into

$$M_{22} + M_{21}\Delta M_{12} \quad (7)$$

which is an affine expression in Δ . As a result, when the controller \bar{C} is applied to the plant P_o , finding the smallest possible allowable model perturbation

Δ such that $P_o \in \mathcal{P}$ (via system identification techniques) will effectively minimize the worst case performance (de Callafon and Van den Hof, 1997). This property can be exploited to formulate a (control relevant) identification problem to estimate a set of models by employing the knowledge of a stabilizing controller \bar{C} that is currently being implemented on the (unknown) plant P_o .

3 Estimation of a set of models

3.1 Control relevant identification

In order to design an enhanced performing robust controller, it is preferable to use a set of models \mathcal{P} for which

$$\sup_{P \in \mathcal{P}} \|J(P, C)\|_{\infty}$$

is minimized. Clearly, this makes the modelling of a set of models \mathcal{P} and the design of a robust controller interrelated (Skelton, 1989). To deal with the interrelation between modelling and control design, knowledge of a controller \bar{C} that is implemented on the unknown plant P_o , similar as in (4), can be exploited to estimate a set of models \mathcal{P} . In that case, a set of models \mathcal{P} subjected to the condition $P_o \in \mathcal{P}$ should be estimated such that

$$\sup_{P \in \mathcal{P}} \|J(P, \bar{C})\|_{\infty} \quad (8)$$

is minimized. In this way, a set of models is found for which the worst case performance for the controller \bar{C} is minimized.

As the controller \bar{C} is assumed to be known, the unknown variables in the coefficient matrix Q of (5) are the factorization (\hat{N}, \hat{D}) of a nominal model and the weighting functions (\hat{V}, \hat{W}) . Minimizing (8) using these variables simultaneously is (as yet) unfeasible. Therefore, minimization of (8) is tackled by estimating the *ref* (\hat{N}, \hat{D}) and the pair (\hat{V}, \hat{W}) separately. In this way, (standard) tools for the identification of a nominal factorization and an uncertainty bound can be employed.

3.2 Estimation of a nominal model

Estimation of a nominal model involves the estimation of $\hat{P} = \hat{N}\hat{D}^{-1}$, subjected to internal stability of the feedback connection $\mathcal{T}(\hat{P}, \bar{C})$, such that (8) is being minimized. At this stage, the variables \hat{V} and \hat{W} are unknown and assumed to vary freely in order to satisfy $P_o \in \mathcal{P}$. Consequently, the set \mathcal{P} is still unknown and (8) cannot be computed. However, for any $P \in \mathcal{P}$ the following upper bound for $\|J(P, \bar{C})\|_{\infty}$ can be given.

$$\|J(P_o, \bar{C})\|_{\infty} + \|J(P, \bar{C}) - J(P_o, \bar{C})\|_{\infty}$$

As $\|J(P_o, \bar{C})\|_\infty$ in (3.2) does not depend on the nominal model \hat{P} , a *rcf* (\hat{N}, \hat{D}) of a nominal model can be found by minimizing

$$\|J(P, \bar{C}) - J(P_o, \bar{C})\|_\infty \quad (9)$$

Estimation of a *rcf* of a nominal model of limited complexity by minimizing (9) on the basis of closed loop experiments obtained from the connection $\mathcal{T}(P_o, \bar{C})$, has been studied in Van den Hof *et al.* (1995). An approach to minimize (9) on the basis of frequency domain data can be found in de Callafon and Van den Hof (1995).

3.3 Estimation of uncertainty bounds

Estimation of an allowable model perturbation involves the characterization of an upper bound on $\bar{\Delta}$ in (4) via (\hat{V}, \hat{W}) such that (8) is being minimized and $P_o \in \mathcal{P}$. For that purpose, first a frequency dependent upper bound on the allowable model perturbation $\bar{\Delta}$ in (5) is determined such that $P_o \in \mathcal{P}$. For that purpose, any uncertainty estimation procedure can be used, as the input and output data of the allowable model perturbation $\bar{\Delta}$ can be accessed simply by a filtering of the input u and output y signals present in the feedback connection $\mathcal{T}(P_o, \bar{C})$ (de Callafon and Van den Hof, 1997).

Similar to the approach presented in (Lee *et al.*, 1993), the availability of the input and output signals of $\bar{\Delta}$ gives rise to an open loop identification problem of the stable dual Youla-Kucera parameter. However, the estimation is being used here to find an upper bound on $\bar{\Delta}$. An uncertainty estimation routine such as the procedure described by Hakvoort (1994) can be used to obtain a frequency dependent upper bound for Δ

$$\|\Delta_i(\omega)\| \leq \delta(\omega) \text{ with probability } \geq \alpha \quad (10)$$

where α is a prechosen probability. In the multivariable case, the upper bound (10) can be obtained for each transfer function. Subsequently, stable and stably invertible weighting filters $\hat{V}(\omega)$ and/or $\hat{W}(\omega)$ of limited complexity can be constructed to over bound $\delta(\omega)$ (Hakvoort, 1994).

4 Controller design

The set of models \mathcal{P} represents the incomplete knowledge on the plant P_o and can be used for subsequent control design. Again taking into account the performance specification (2), a controller C can be designed by minimizing

$$\sup_{P \in \mathcal{P}} \|J(P, C)\|_\infty \quad (11)$$

where \mathcal{P} denotes the set of models being estimated. For $J(P, C) = U_2 T(P, C) U_1$, (11) constitutes a

(standard) \mathcal{H}_∞ -norm based control design, wherein the worst case performance is being optimized. For that purpose, a μ -synthesis via a so-called D-K iteration (Zhou *et al.*, 1996) can be used. In order to use the available techniques on μ -synthesis, the transfer function M in Lemma 2.1 should be represented as a lower fractional transformation $\mathcal{F}_l(G, C)$, where the controller C to be designed has been extracted. An expression for G can be found by standard algebraic manipulations.

5 Closed loop reduction

The design of a controller as mentioned in Section 4 generally leads to full order controllers, although limited complexity of the coefficient matrix Q in (5) can be enforced by the approximate identification of a *rcf* (\hat{N}, \hat{D}) and the weighting filters (\hat{V}, \hat{W}) .

In light of the performance objective function $J(P, C)$ given in (2), a reduction of the controller may be required, that takes account of this performance function. For that purpose, a closed loop balanced reduction, as proposed by Ceton *et al.* (1993), is well suited. In Ceton *et al.* (1993), a similarity transformation that balances the states of a stable feedback connection is used for partial balancing of the (unstable) controller states (Wortelboer, 1993). As a result, an (unstable) controller can be reduced in closed loop, taking into account the closed loop operation of the controller.

The closed loop configuration in Ceton *et al.* (1993) is slightly different from the one used in this paper. However, the results of Ceton *et al.* (1993) can be readily carried over to perform closed loop reduction of the controller C in the feedback connection $\mathcal{T}(P, C)$, incorporating the performance weightings U_2 and U_1 .

6 Application to waferstepper

6.1 Description of the positioning mechanism

The approach outlined in this paper has been applied to a multivariable positioning mechanism, denoted by the wafer stage, present in a wafer stepper. A wafer stepper is a fast and high accuracy positioning machine, used in chip manufacturing processes; a schematic view is depicted in Figure 2. The position of the wafer chuck on the horizontal surface of a granite block is measured by means of three laser interferometry measurements, whereas three linear motors are used to position the wafer chuck. The three currents to the linear motors denote the input u , whereas the three position measurements denote the output y of the system.

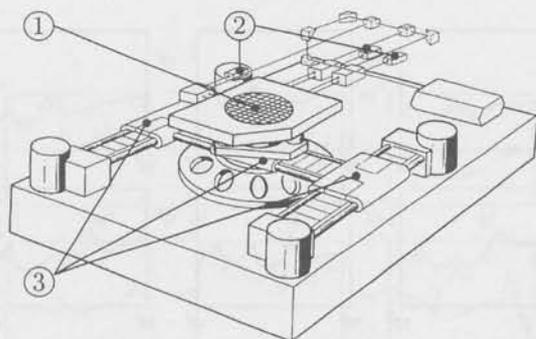


Fig. 2: Schematic view of a wafer stage; 1:wafer chuck, 2:laser interferometers, 3:linear motors.

A diagonal PID controller is used as an initial controller \bar{C} to stabilize and position the wafer chuck for experimental purposes. External references signals r_1 and r_2 are used to excite the closed loop similar to Figure 1. Time and frequency domain data were gathered for identification purposes. The aim is to design a low complexity controller that is able to attain a high bandwidth, tracking and suppression of residual vibrations. For that purpose, only relatively simple (diagonal) weighting functions U_2 and U_1 are used to enforce a controller with high gain at low frequencies.

6.2 Estimation of a nominal factorization

First a MIMO nominal $ref\ col(\hat{N}, \hat{D})$ having 6 outputs and 3 inputs is estimated. For that purposes, frequency measurements are used to curve fit a factorization (\hat{N}, \hat{D}) of 30th order using the procedure described in de Callafon and Van den Hof (1995). This procedure requires an initial estimate for the non-linear optimization which is found by a MIMO least squares curve fitting (de Callafon *et al.*, 1996). An amplitude Bode plot of the result is presented in Figures 3 and 4. It can be observed from these figures that the frequency domain data has only been approximated by the factorization (\hat{N}, \hat{D}) , as more accurate modelling would require a much higher order model.

6.3 Estimation of model uncertainty

Given the nominal factorization (\hat{N}, \hat{D}) and a normalized $ref(N_c, D_c)$ of the controller \bar{C} , an estimation of the allowable model perturbation $\bar{\Delta}$ in (4) is performed. For that purpose, the uncertainty estimation as presented in (Hakvoort, 1994) has been applied to estimate a frequency dependent upper bound on $\bar{\Delta}$. Due to space limitations only the result is presented in Figure 5.

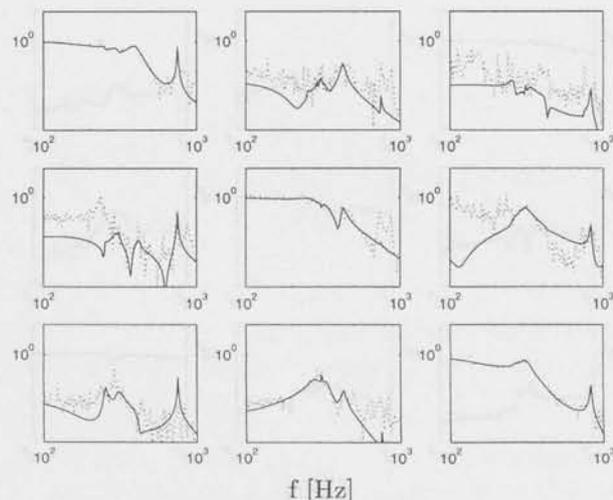


Fig. 3: Amplitude Bode plot of nominal numerator factor \hat{N} (-), and the corresponding frequency domain data (\dots)

It can be observed from Figure 5 that the upper bound of the frequency domain estimation of $\bar{\Delta}$ is crossing the upper bound $\delta(\omega)$. Partly, this is due to the fact the upper bound only holds within a prespecified probability of 95%.

6.4 Full order controller and reduction

On the basis of the nominal factorization (\hat{N}, \hat{D}) and (only) a single stable and stable invertible weighting filter \hat{V} that over-bounds the upper bounds $\delta(\omega)$ depicted in Figure 5, a robust controller has been designed by means of a μ -synthesis. An amplitude Bode plot of the controller has been depicted in Figure 6.

Despite of the low complexity modelling, the full order controller being designed still has a McMillan degree of 74. Additional reduction of the controller as described in section 5 enables the controller to be reduced to 32nd order. The additional closed loop reduction deteriorates the performance robustness only by 2.12%. The 32nd order controller has been applied to the wafer stepper mechanism successfully.

7 Conclusions

In this paper a systematic approach to find a low complexity controller for a unknown system has been presented. The approach consists of a system identification technique to estimate a model uncertainty set, followed by a robust controller design and an additional controller reduction. In all these steps, the performance and the closed loop operation of

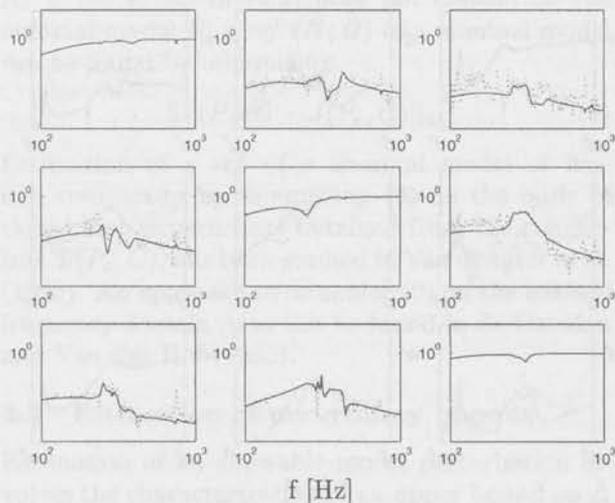


Fig. 4: Amplitude Bode plot of nominal denominator factor \hat{D} (-) and the corresponding frequency domain data ($\cdot\cdot\cdot$)

both the uncertainty set and the low complexity controller being constructed is taken into account. The approach has been illustrated on a highly complex multivariable mechanical servo system present in a wafer stepper. This has resulted in a relatively low order controller that successfully has been applied.

Acknowledgments

The authors would like to acknowledge the support of Philips Research Laboratories, Eindhoven, the Netherlands for providing the wafer stepper experimental setup, and Dick de Roover for fruitful discussions and his contributions to this paper.

References

- Boyd, S.P. and C.H. Barrat (1991). *Linear Controller design - Limits of Performance*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.
- de Callafon, R.A., D. Roover and P.M.J. Van den Hof (1996). Multivariable least squares frequency domain identification using polynomial matrix fraction descriptions. *Proc. 35th IEEE Conf. Decis. Control*, Kobe, Japan, pp. 2030-2035.
- de Callafon, R.A. and P.M.J. Van den Hof (1995). Control relevant identification for H_∞ -norm based performance specifications. *Proc. 34th IEEE Conf. Decis. Control*, New Orleans, LA, pp. 3498-3503.
- de Callafon, R.A. and P.M.J. Van den Hof (1997). Suboptimal feedback control by a scheme of it-

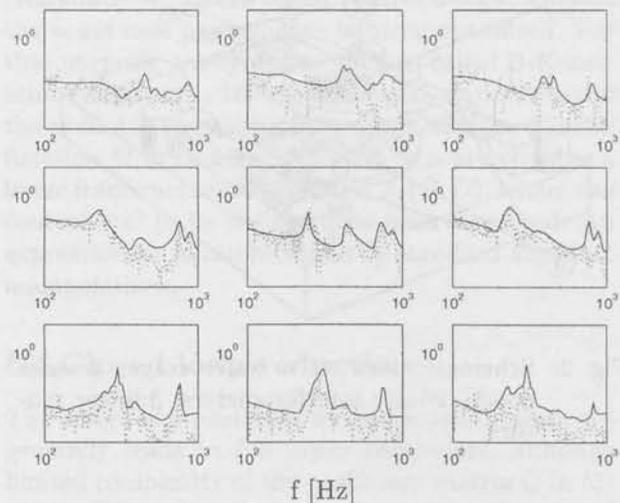


Fig. 5: Amplitude Bode plot of uncertainty bound $\delta(\omega)$ (-) and frequency domain estimate of $\hat{\Delta}$ ($\cdot\cdot\cdot$)

erative identification and control design. *Mathematical Modelling of Systems*, 3, 77-101.

- Ceton, C., P.M.R. Wortelboer and O.H. Bosgra (1993). Frequency weighted closed-loop balanced reduction. *Proc. 2nd European Control Conference*, Groningen, the Netherlands, pp. 697-701.
- Hakvoort, R.G. (1994). *System Identification for Robust Process Control: Nominal Models and Error Bounds*. PhD thesis, Delft Univ. Techn., Delft.
- Lee, W.S., B.D.O. Anderson, R.L. Kosut and I.M.Y. Mareels (1993). A new approach to adaptive robust control. *Int. J. Adapt. Contr. Sign. Proc.*, 7, 183-211.
- Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.
- Mäkilä, P.M., J.R. Partington and T.K. Gustafson (1995). Worst-case control-relevant identification. *Automatica*, 31, 1799-1819.
- Ninness, B. and G.C. Goodwin (1995). Estimation of model quality. *Automatica*, 31, 1771-1797.
- Schrama, R.J.P. and O.H. Bosgra (1993). Adaptive performance enhancement by iterative identification and control design. *Int. J. Adapt. Cont. Sign. Proc.*, 7, 475-487.
- Sefton, J.A., R.J. Ober and K. Glover (1990). Robust stabilization in the presence of coprime factor perturbations. *Proc. 29th IEEE Conf. Decis. Control*, Honolulu, HI, pp. 1197-1198.
- Skelton, R.E. (1989). Model error concepts in control design. *Int. J. of Control*, 49, 1725-1753.
- Smith, R., G. Dullerud, S. Rangan and K. Poola (1997). Model validation for dynamically uncer-

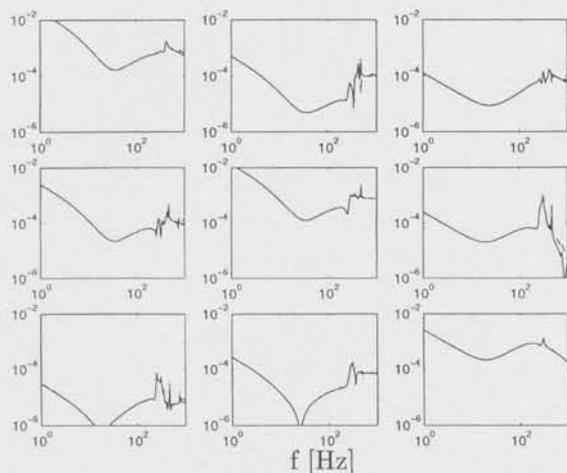


Fig. 6: Amplitude Bode plot of full order controller (- -) and closed loop reduced controller (-)

- tain systems. *Mathematical Modelling of Systems*, 3, 43-58.
- 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.J.P. Schrama, R.A. de Callafon and O.H. Bosgra (1995). Identification of normalised coprime plant factors from closed-loop experimental data. *Europ. J. of Control*, 1, 62-74.
- Vidyasagar, M. (1985). *Control System Synthesis: A Factorization Approach*. MIT Press, Cambridge, MA.
- 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.
- Zhou, K., J.C. Doyle and K. Glover (1996). *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ.

Doctoral dissertations published in 1997

Gert-Wim van der Linden

High Performance Robust Manipulator Motion Control - Theory, Tools and Experimental Implementation

6 January 1997

Advisors: Okko Bosgra and Ton van der Weiden

Current address:

SCSolutions, systems and control division, 3211
Scott Boulevard, Santa Clara, CA 95054, USA.

E-mail: linden@scsolutions.com

Hans Heintze

Design and Control of a Hydraulically Actuated Industrial Brick Laying Robot

25 March 1997

Advisors: Okko Bosgra and Ton van der Weiden

Gert van Schothorst

Modelling of Long-Stroke Hydraulic Servo-Systems for Flight Simulator Motion Control and System Design

9 September 1997

Advisors: Okko Bosgra, Bob Mulder and Ton van der Weiden

Current address:

Philips Research Laboratories, Prof. Holstlaan 4,
5656 AA Eindhoven, The Netherlands.

E-mail: schothor@natlab.research.philips.com

Dick de Roover

Motion Control of a Wafer Stage - A Design Approach for Speeding Up IC Production

2 December 1997

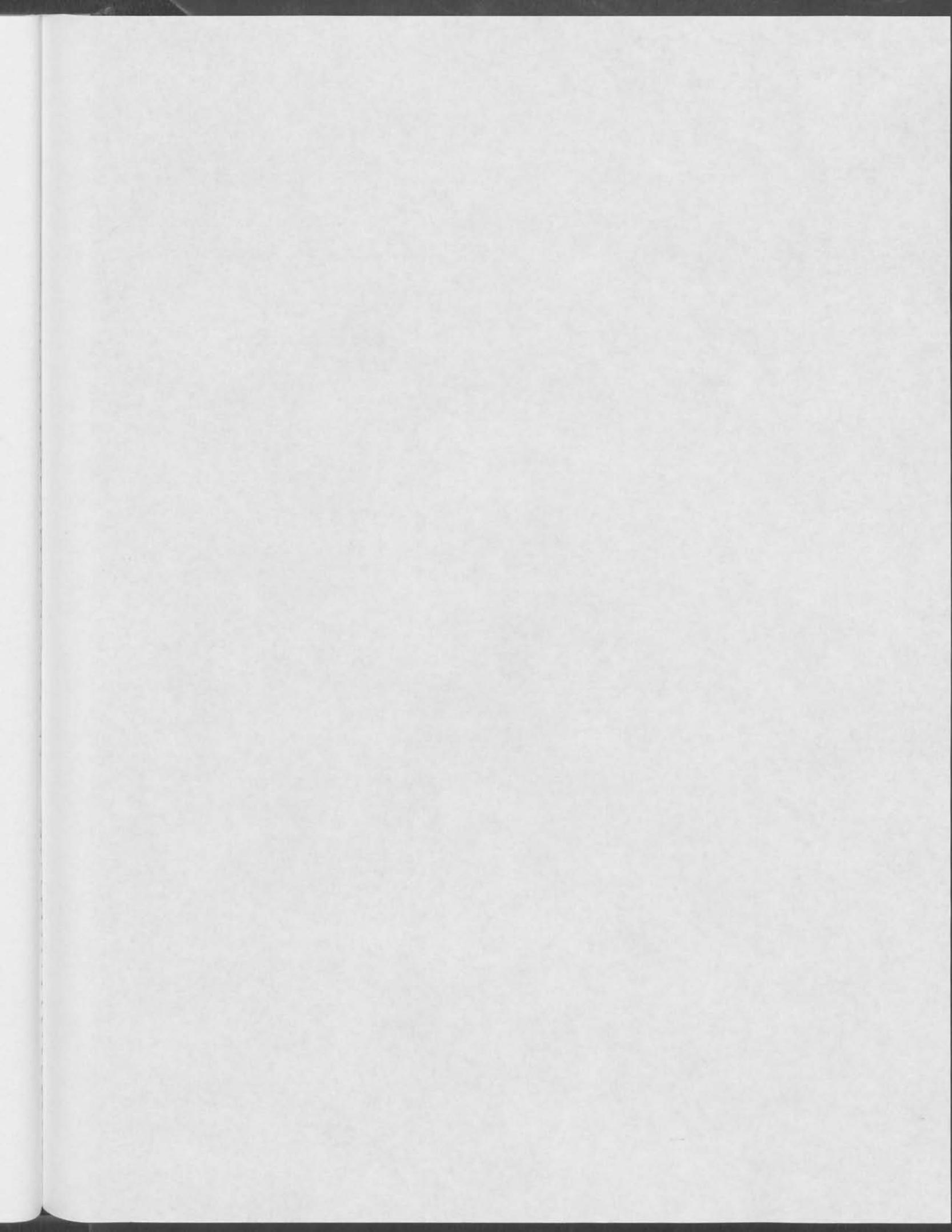
Advisor: Okko Bosgra

Current address:

SCSolutions, systems and control division, 3211
Scott Boulevard, Santa Clara, CA 95054, USA.

More information on these dissertations can be
found on WWW-page:

www-sr.wbmt.tudelft.nl/sr/thesis



Doctoral dissertations published in 1997

Gert Wijn van der Linden
Real Performance Robot Manipulator Motion Control - Theory, Test and Experimental Implementation
1997
6 January 1997
Advisors: Olof Bengt and Tim van der Weiden

Current address:
SCSolutions, systems and control division, 1211
Scott Boulevard, Santa Clara, CA 95054, USA
E-mail: industrial@scsolutions.com

Rene Holman
*Control and Control of a Hydrokinetic Actuated In-
cubator Drive Loading Robot*
1997
21 March 1997
Advisors: Olof Bengt and Tim van der Weiden

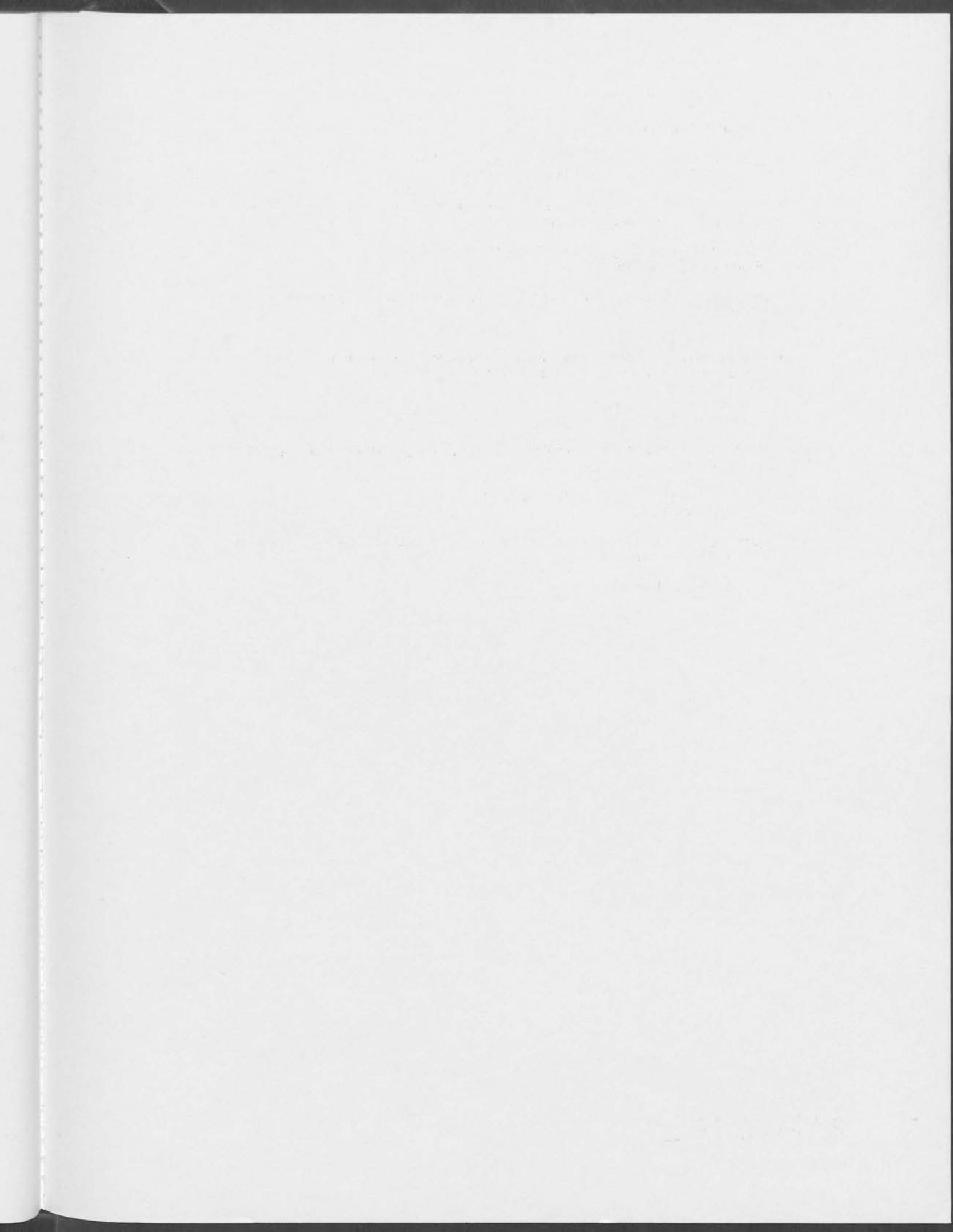
Gert van Schoonhoven
*Modeling of Long-Dwell Hydraulic Servo Systems
for Flight Simulator Motion Control and System De-
sign*
1 September 1997
Advisors: Olof Bengt, Rob Mulder and Tim van
der Weiden

Current address:
Phillips Research Laboratories, Fred. Idenburg 4,
9601 AA Breda, The Netherlands
E-mail: wholman@philips.com

Dick de Boer
*Motion Control of a Robot Drive - A Design Ap-
proach for Servoing DC-DC Converters*
1 December 1997
Advisor: Olof Bengt

Current address:
SCSolutions, systems and control division, 1211
Scott Boulevard, Santa Clara, CA 95054, USA

More information on these dissertations can be
found at www.scsolutions.com
www.scsolutions.com/theses



CONTENTS Volume 10, December 1997

Identification of a fluidized catalytic cracking unit: an orthonormal basis function approach <i>E.T. van Donkelaar, P.S.C. Heuberger and P.M.J. Van den Hof</i>	1
Closed-loop identification of uncertainty models for robust control design: a set membership approach <i>M. Milanese, M. Taragna and P.M.J. Van den Hof</i>	9
Mixed objectives MIMO control design for a Compact Disc player <i>M. Dettori, V. Prodanovic and C.W. Scherer</i>	19
Indirect position measurement and singularities in a Stewart platform with an application to model-based control <i>S.H. Koekebakker</i>	27
Suppressing non-periodically repeating disturbances in mechanical servo systems <i>R.L. Tousain, J.C. Boissy, M.L. Norg, M. Steinbuch and O.H. Bosgra</i>	39
A full block S-procedure with applications <i>C.W. Scherer</i>	47
An internal-model-based framework for the analysis and design of repetitive and learning controllers <i>D. de Roover and O.H. Bosgra</i>	53
Point-to-point control of a high accuracy positioning mechanism <i>D. de Roover and F.B. Sperling</i>	61
Low order control design by feedback relevant identification and closed-loop controller reduction <i>R.A. de Callafon and P.M.J. Van den Hof</i>	69
Doctoral dissertations published in 1997	76