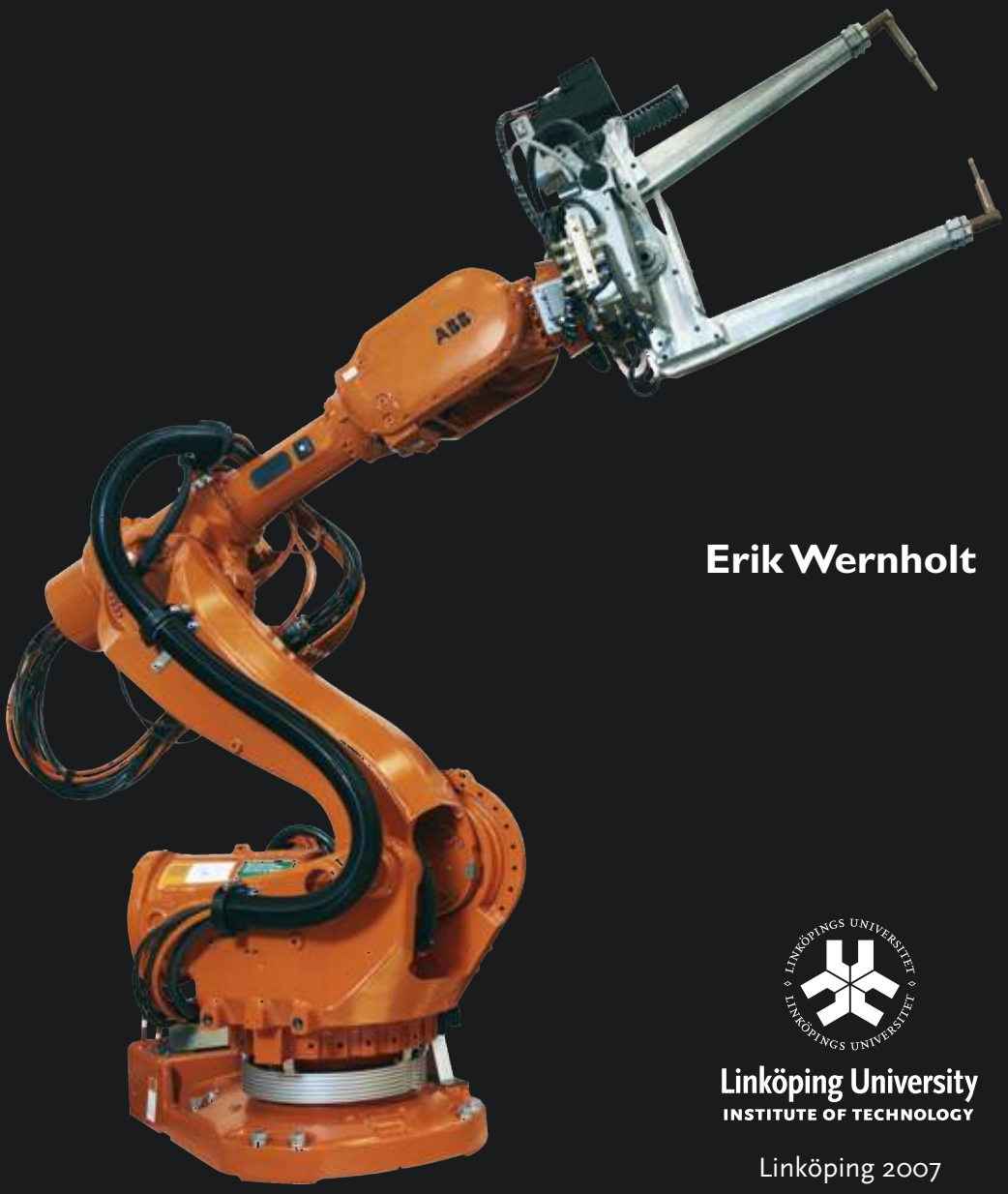
Linköping Studies in Science and Technology. Dissertations. No. 1138

Multivariable Frequency-Domain Identification of Industrial Robots



Erik Wernholt

Linköping University
INSTITUTE OF TECHNOLOGY

Linköping 2007

Linköping Studies in Science and Technology. Dissertations.
No. 1138

Multivariable Frequency-Domain Identification of Industrial Robots

Erik Wernholt

Department of Electrical Engineering
Linköping University, SE–581 83 Linköping, Sweden
Linköping 2007

Linköping Studies in Science and Technology. Dissertations.
No. 1138

Multivariable Frequency-Domain Identification of Industrial Robots

Erik Wernholt

erikw@isy.liu.se
www.control.isy.liu.se
Division of Automatic Control
Department of Electrical Engineering
Linköping University
SE–581 83 Linköping
Sweden

ISBN 978-91-85895-72-4 ISSN 0345-7524

Printed by LiU-Tryck, Linköping, Sweden 2007

To Jessica and Nils

Abstract

Industrial robots are today essential components in the manufacturing industry where they are used to save costs, increase productivity and quality, and eliminate dangerous and laborious work. High demands on accuracy and speed of the robot motion require that the mathematical models, used in the motion control system, are accurate. The models are used to describe the complicated nonlinear relation between the robot motion and the motors that cause the motion. Accurate dynamic robot models are needed in many areas, such as mechanical design, performance simulation, control, diagnosis, and supervision.

A trend in industrial robots is toward lightweight robot structures, where the weight is reduced but with a preserved payload capacity. This is motivated by cost reduction as well as safety issues, but results in a weaker (more compliant) mechanical structure with enhanced elastic effects. For high performance, it is therefore necessary to have models describing these elastic effects.

This thesis deals with identification of dynamic robot models, which means that measurements from the robot motion are used to estimate unknown parameters in the models. The measured signals are angular position and torque of the motors. Identifying robot models is a challenging task since an industrial robot is a multivariable, nonlinear, unstable, and resonant system. In this thesis, the unknown parameters (typically spring-damper pairs) in a physically parameterized nonlinear dynamic model are identified, mainly in the frequency domain, using estimates of the nonparametric frequency response function (FRF) in different robot configurations/positions. Each nonparametric FRF then describe the local behavior around an operating point. The nonlinear parametric robot model is linearized in the same operating points and the optimal parameters are obtained by minimizing the discrepancy between the nonparametric FRFs and the parametric FRFs (the FRFs of the linearized parametric robot model).

Methods for estimating the nonparametric FRF from experimental data are analyzed with respect to bias, variance, and nonlinearities. In order to accurately estimate the nonparametric FRF, the experiments must be carefully designed. To minimize the uncertainty in the estimated parameters, the selection of optimal robot configurations/positions for the experiments is also part of the design. Different parameter estimators are compared in the thesis and experimental results show the usefulness of the proposed identification procedure. The identified nonlinear robot model gives a good global description of the dynamics in the frequency range of interest.

The research work is also implemented and made easily available in a software tool for accurate estimation of nonparametric FRFs as well as parametric robot models.

Sammanfattning

Industrirobotar är idag en väsentlig del i tillverkningsindustrin där de bland annat används för att minska kostnader, öka produktivitet och kvalitet och ersätta människor i farliga eller slitsamma uppgifter. Höga krav på noggrannhet och snabbhet hos robotens rörelser innebär också höga krav på de matematiska modeller som ligger till grund för robotens styrsystem. Modellerna används där för att beskriva det komplicerade sambandet mellan robotarmens rörelser och de motorer som orsakar rörelsen. Tillförlitliga modeller är också nödvändiga för exempelvis mekanisk design, simulering av prestanda, diagnos och övervakning.

En trend idag är att bygga lättviktsrobotar, vilket innebär att robotens vikt minskas men att den fortfarande kan hantera en lika tung last. Orsaken till detta är främst att minska kostnaden, men också säkerhetsaspekter spelar in. En lättare robotarm ger dock en vekare struktur där elastiska effekter inte längre kan försummas i modellen om man kräver hög prestanda. De elastiska effekterna beskrivs i den matematiska modellen med hjälp av fjädrar och dämpare.

Denna avhandling handlar om hur dessa matematiska modeller kan tas fram genom systemidentifiering, vilket är ett viktigt verktyg där mätningar från robotens rörelser används för att bestämma okända parametrar i modellen. Det som mäts är position och moment hos robotens alla motorer. Identifiering av industrirobotar är ett utmanande problem bland annat eftersom robotens beteende varierar beroende på armens position. Den metod som föreslås i avhandlingen innebär att man först identifierar lokala modeller i ett antal positioner. Var och en av dessa beskriver robotens beteende kring en viss arbetspunkt. Sedan anpassas parametrarna i en global modell, som är giltig för alla positioner, så att den så väl som möjligt beskriver det lokala beteendet i de olika positionerna.

I avhandlingen analyseras olika metoder för att ta fram lokala modeller. För att få bra resultat krävs att experimenten är omsorgsfullt utformade. För att minska osäkerheten i den globala modellens identifierade parametrar ingår också valet av optimala positioner för experimenten. Olika metoder för att identifiera parametrarna jämförs i avhandlingen och experimentella resultat visar användbarheten av den föreslagna metoden. Den identifierade robotmodellen ger en bra global beskrivning av robotens beteende.

Resultatet av forskningen har även gjorts tillgängligt i ett datorverktyg för att noggrant kunna ta fram lokala modeller och identifiera parametrar i dynamiska robotmodeller.

Acknowledgments

There are several people who helped me during the work of this thesis. First of all I would like to thank my supervisor Professor Svante Gunnarsson for guiding me in my research and always taking time to answer my questions. I am also grateful to Professor Lennart Ljung for giving me the opportunity to join Automatic Control in Linköping, the most inspiring and challenging environment to conduct research in.

The work in this thesis would not have been possible without the support from ABB Robotics. I would especially like to mention my industrial mentor Dr. Torgny Brogårdh and also Stig Moberg for all our profitable discussions on various aspects of robotics and in particular identification of robots. I am also thankful to Sven Hanssen who has provided the simulation models used in the thesis. Moreover, the possibility to use the robot lab at ABB Robotics for various experiments has substantially improved the results of the thesis.

I would also like to thank Stig Moberg and Dr. Johan Löfberg for our fruitful cooperation in some of the included papers.

Parts of the thesis have been proofread by Professor Svante Gunnarsson, Dr. Torgny Brogårdh, Stig Moberg, Dr. Martin Enqvist, Dr. Jacob Roll, and Dr. Thomas Schön. Your comments have improved the quality of this thesis substantially and I am very grateful for that. A special thanks goes to Dr. Martin Enqvist for always having time to discuss various issues with me and for being such a good friend. Dr. Mikael Norrlöf has been helping me with various problems and giving me valuable comments on my work. Thanks also to Dr. Peter Lindskog for providing the code for the time-domain nonlinear gray-box identification.

I would also like to thank everyone else in the Automatic Control group for a productive and enjoyable climate and for the interesting lunch discussions on various topics. Ulla Salaneck deserves extra gratitude for helping me with various practical issues and Gustaf Hendeby has helped me with many LaTeX-related questions.

In addition, I would like to thank Professor Johan Schoukens for giving me valuable comments on my work and for letting me use some of his time for discussions.

The financial support for my research has been provided by the VINNOVA Center of Excellence ISIS (Information Systems for Industrial Control and Supervision) at Linköping University, which is gratefully acknowledged.

My family and friends deserve my deepest gratitude for always being there for me. Finally, I would like to thank Jessica and our son Nils for all their love, support, and patience with me during intensive times of writing and thinking. You bring joy into my life!

Erik Wernholt

Linköping, October 2007

Contents

1 Introduction **1**
1.1 Background . . . 1
1.2 Industrial robots . . . 2
1.3 Problem Statement . . . 3
1.4 Outline . . . 5
1.4.1 Outline of Part I . . . 5
1.4.2 Outline of Part II . . . 5
1.5 Contributions . . . 9

I Overview **11**

2 Robotics **13**
2.1 Introduction . . . 13
2.1.1 Manipulator Structures . . . 13
2.1.2 Facts and Figures . . . 16
2.1.3 Historical Background . . . 18
2.2 Modeling . . . 20
2.2.1 Actuators and Sensors . . . 20
2.2.2 Kinematics . . . 23
2.2.3 Rigid Body Dynamics . . . 24
2.2.4 Flexible Body Dynamics . . . 25
2.3 Control . . . 28
2.3.1 Motion Planning . . . 28
2.3.2 Trajectory Generation . . . 28
2.3.3 Trajectory Tracking . . . 30

3 System Identification **33**
3.1 Preliminaries . . . 33
3.2 The System Identification Procedure . . . 36
3.3 Model Structures . . . 37
3.4 Calculating the Estimate . . . 39
3.4.1 Parametric Time-Domain Methods . . . 39
3.4.2 Nonparametric Frequency-Domain Methods . . . 41
3.4.3 Parametric Frequency-Domain Methods . . . 43
3.4.4 Time-Domain versus Frequency-Domain Methods . . . 44
3.5 Closed-Loop Identification . . . 45
3.6 Bias and Variance . . . 46
3.6.1 Parametric Methods . . . 47
3.6.2 Nonparametric Methods . . . 48
3.7 Experiment Design . . . 49
3.7.1 Selection of Power Spectrum . . . 49
3.7.2 Selection of Excitation Signal . . . 51
3.7.3 Nonlinear Gray-Box Models . . . 55
3.7.4 Dealing with Transients . . . 55
3.8 Nonparametric FRF Estimation of Nonlinear Systems . . . 55
3.8.1 Analysis Using Volterra Series . . . 56
3.8.2 Excitation Signals . . . 56
3.8.3 Example: Nonlinear Two-Mass Model . . . 57

4 System Identification in Robotics **61**
4.1 Introduction . . . 61
4.2 Kinematics . . . 62
4.3 Rigid Body Dynamics . . . 64
4.3.1 Measurements . . . 64
4.3.2 Base Parameters . . . 65
4.3.3 Estimators . . . 65
4.3.4 Experiment Design . . . 65
4.4 Flexibilities and Nonlinearities . . . 66
4.4.1 Identification of Friction and Backlash . . . 67
4.4.2 Identification of Flexible Joint Models . . . 68
4.4.3 Identification Using Additional Sensors . . . 69
4.4.4 Identification of Extended Flexible Joint Models . . . 70
4.5 EXPDES—An Identification Toolbox for Industrial Robots . . . 72

5 Concluding Remarks **75**
5.1 Conclusion . . . 75
5.2 Future Research . . . 76

Bibliography **79**

II Publications **89**

A Frequency-Domain Gray-Box Identification of Industrial Robots **91**
1 Introduction . . . 93
2 Problem Description . . . 94
3 Robot Model . . . 96
4 FRF Estimation . . . 97
5 Parameter Estimation . . . 98
5.1 Estimators . . . 99
5.2 Optimal Positions . . . 101
5.3 Solving the Optimization Problem . . . 101
6 Experimental Results . . . 102
7 Concluding Discussion . . . 108
References . . . 109

B Analysis of Methods for Multivariable Frequency Response Function Estimation in Closed Loop **113**
1 Introduction . . . 115
2 FRF Estimation Methods . . . 116
3 Open-Loop Error Analysis . . . 119
3.1 Noise Assumptions . . . 119
3.2 Bias and Covariance . . . 119
4 Closed-Loop Error Analysis . . . 121
4.1 Bias . . . 121
4.2 Covariance . . . 123
5 Numerical Illustration . . . 125
6 Conclusion . . . 127
References . . . 128

C Experimental Comparison of Methods for Multivariable Frequency Response Function Estimation **133**
1 Introduction . . . 135
2 Measurement Setup . . . 136
3 FRF Estimation Methods . . . 137
3.1 Basic Idea . . . 137
3.2 Some Classical Estimators . . . 138
3.3 Estimators Based on Nonlinear Averaging Techniques . . . 139
4 Excitation Signals . . . 140
5 Experimental Results . . . 141
5.1 Measurement Data . . . 141
5.2 Evaluation of FRF Estimators . . . 142
5.3 Discussion . . . 145
6 Conclusion . . . 148
References . . . 148

D Estimation of Nonlinear Effects in Frequency-Domain Identification of Industrial Robots **151**
1 Introduction . . . 153
2 Measurement Setup . . . 154
3 Estimation Method . . . 155
3.1 Introduction . . . 156
3.2 SISO Measurements . . . 157
3.3 MIMO Measurements . . . 158
4 Experimental Results . . . 159
4.1 SISO Measurements . . . 160
4.2 MIMO Measurements . . . 161
5 Conclusion . . . 164
References . . . 165

E Experiment Design for Identification of Nonlinear Gray-Box Models with Application to Industrial Robots **167**
1 Introduction . . . 169
2 Problem Description . . . 170
3 The Information Matrix . . . 171
4 The Experiment Design Problem . . . 173
5 Solving the Experiment Design Problem . . . 175
6 Numerical Illustration . . . 176
6.1 Robot Model . . . 176
6.2 Measurements . . . 177
6.3 Solution using YALMIP and SDPT3 . . . 179
6.4 Results . . . 179
7 Conclusion . . . 182
References . . . 183

F Nonlinear Gray-Box Identification of a Flexible Manipulator **185**
1 Introduction . . . 187
2 Problem Description . . . 188
3 Nonlinear Gray-Box Identification . . . 190
4 Three-Step Identification Procedure . . . 191
4.1 Step 1: Rigid Body Dynamics and Friction . . . 191
4.2 Step 2: Initial Values for Flexibilities . . . 192
4.3 Step 3: Nonlinear Gray-Box Identification . . . 193
5 Data Collection . . . 193
6 Results . . . 194
6.1 Step 1 . . . 194
6.2 Step 2 . . . 195
6.3 Step 3 . . . 196
6.4 Discussion . . . 197
7 Conclusion . . . 198
References . . . 199

Notation **201**

1

Introduction

This thesis deals with identification of unknown parameters in dynamic models of industrial robots. This involves experiment design, nonparametric frequency-domain identification, and identification of unknown parameters in physically parameterized models (often called gray-box models).

This first chapter gives a brief background to the project in Section 1.1, followed by a short introduction to industrial robots in Section 1.2. The problem is stated in Section 1.3 and the outline of the thesis is provided in Section 1.4. Finally, a summary of the contributions is given in Section 1.5.

1.1 Background

This project has been carried out within the VINNOVA Center of Excellence ISIS (Information Systems for Industrial Control and Supervision) at Linköping University. Within the center that started 1996, research is performed in different areas, focusing on problems that are important for the industry. One of the companies that joined as a partner is ABB Robotics. Having them as a partner in research is both challenging and rewarding. The real world is often too complicated to be properly described by mathematical models. Moreover, obtaining general results for such complicated systems is a real struggle. On the other hand, the result that *do* come out of the project is relevant for an important application and it is possible to see its immediate use. That is great fun!

Another good thing with this setup is that all experiments have been carried out in the research lab of ABB Robotics, Västerås, Sweden, using their hardware as well as expertise. Spending time there has proved to be a good way of transfer knowledge from the research community to the company as well as vice versa.

The project has also resulted in a software tool, written by the author of this thesis, to enable experiment design, preprocessing of experimental data, and accurate estimation of both nonparametric and parametric models. With the software tool, most of the research

work in this thesis is implemented and made easily available. The code has been developed over a period of more than three years, with an effective time of production of about 3–5 months.

1.2 Industrial robots

A standard industrial robot (see Figure 1.1) consists of a mechanical arm with a number of links connected by joints (also called axes), where each joint is actuated by an (electric) motor via a transmission. The movements are controlled by a computer system. Usually the robot has six joints, giving six degrees-of-freedom and the ability to control both the Cartesian position of the tool and its orientation in the workspace. Each combination of the six joints is called a *robot configuration*, or simply *robot position*.

The number of sensors used for control is usually kept at the minimum level of only measuring the motor angular positions, even though the control objective is that the tool follows the programmed path. The relationship between the motors and the tool is described using kinematic and dynamic models. For the design of an advanced robot controller, accurate robot models are therefore crucial to obtain high performance, accuracy, and disturbance rejection.

The development rate of new industrial robots is also high, with several kinds of robots to tune each year. For top performance, there could also be a need to tune each individual robot, as well as re-tune robots at the customer site due to wear or other changing conditions. This means that there is an increasing need for good procedures to estimate accurate robot models. Good models are also needed for supervision, diagnosis, mechanical design, performance simulation, and so on. Model based diagnosis of robots is, for example, important in order to increase reliability and reduce maintenance time.

Historically, the dynamic models used for control are either entirely rigid (An et al., 1988), or only flexible joint models are considered, i.e., elastic gear transmission and rigid links (Albu-Schäffer and Hirzinger, 2000; Spong, 1987). The trend in industrial robots is toward lightweight robot structures with a reduced mass but with preserved payload capabilities. This is motivated by cost reduction as well as safety issues, but results in lower mechanical resonance frequencies inside the controller bandwidth. The sources of elasticity in such a manipulator are, e.g., gearboxes, bearings, elastic foundations, elastic payloads, as well as bending and torsion of the links. In Öhr et al. (2006) it is shown that there are cases when these other sources of flexibilities can be of the same order as the gearbox flexibilities for a modern industrial robot. Accurate dynamic models that also describe these elastic effects are therefore needed in order to obtain high performance.

The industrial robot poses a challenging modeling problem due to both the system complexity and the required model accuracy. The dynamics of the six joints is strongly coupled (moving one motor will cause motion also in the other joints), giving a truly multivariable system. The dynamics is nonlinear, both with respect to the rigid body dynamics and other things such as non-ideal motors and sensors, and a transmission with friction, backlash, hysteresis, and nonlinear stiffness. The system is resonant due to elastic effects and, in addition, experimental data must usually be collected while the robot controller is operating in closed loop since the open-loop system is unstable.

Two main routes to obtain models are physical modeling and system identification.

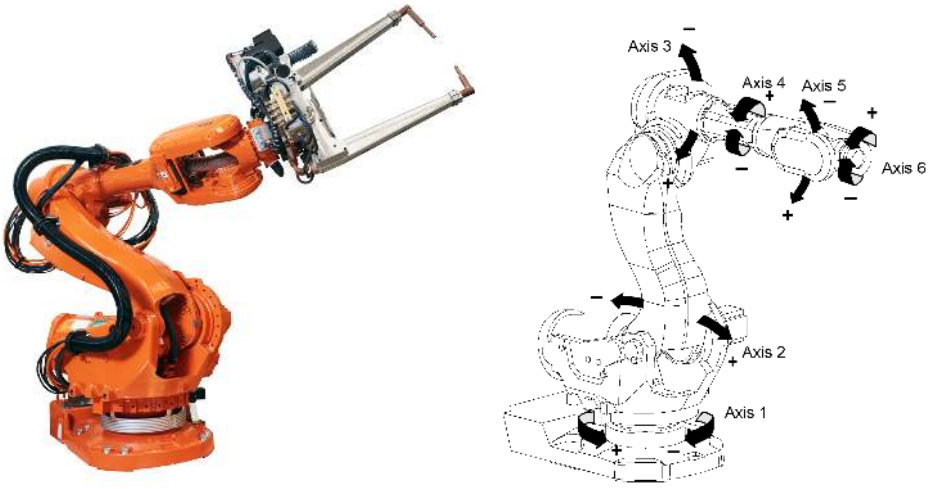


Figure 1.1: *The ABB IRB6600 robot. Left: equipped with a spot-welding gun. Right: illustration with the six axes.*

For the modeling route, basic physical laws and other well-established relationships are used to come up with a model. System identification, on the other hand, uses experimental data to adjust parameters in a selected model structure. In this thesis, a combination of these two routes is used where unknown parameters in a physically parameterized model are estimated from experimental data, often called gray-box identification.

1.3 Problem Statement

The **main problem considered in this thesis** is about identification of unknown parameters in a nonlinear dynamic model of an industrial robot, here represented by a nonlinear state-space model

$$\dot{x}(t) = f(x(t), u(t), \theta), \tag{1.1a}$$
$$y(t) = h(x(t), u(t), \theta), \tag{1.1b}$$

with state vector $x(t)$, input vector $u(t)$, output vector $y(t)$, parameter vector θ, and nonlinear functions $f(\cdot)$ and $h(\cdot)$ that describe the dynamics. The model must be *global*, i.e., valid throughout the whole workspace (all robot configurations/positions), as well as *elastic*, which here means that resonances due to elastic effects are captured by the model. The elastic effects are modeled through a lumped parameter approach (Khalil and Gautier, 2000) where each rigid body is connected by spring-damper pairs. The model is of gray-box type and the rigid body parameters of the model are usually assumed to be known from a CAD model or prior rigid body identification. The main objective is identification of elasticity parameters (spring-damper pairs) but other parameters can be added, such as the location in the robot structure of the spring-damper pairs and a few

unknown rigid body parameters. It is also possible to include nonlinear descriptions of selected quantities (e.g., the gearbox stiffness) and identify those by a linearization for each operating point. An additional constraint in this thesis is that no extra sensors are used for the identification. This means that only measurements of motor angular positions and motor torques (actually the reference signals to the torque controllers) will be used as output y and input u, respectively. This constraint is used both to limit the complexity and to see how much that can be achieved using only the standard sensors.

The real challenge for system identification methods is that the industrial robot is multivariable, nonlinear, unstable, and resonant at the same time. Usually, in the literature, at least one of the first three topics is left out. Identification of such a complex system is therefore a huge task, both in finding suitable model structures and efficient identification methods.

One solution could be to apply a nonlinear prediction error method (Ljung, 1999, pp. 146–147), where measured input-output data are fed to the model and the predicted output from the model is compared with the measured output. This is treated in Paper F for axis one of the industrial robot, which means a stable scalar system (axis one is not affected by gravity). Extending these results to a multivariable and unstable system would involve, for example: finding a stable predictor, numerical problems, and handling large data sets. The last two problems stem from the fact that the system is resonant and numerically stiff, as well as large in dimension both with respect to the number of states and parameters. Moreover, the choice of model structure (parameters) and handling local minima in the optimization are also parts of the problem. Apart from all these problems, such a solution would really tackle our main problem.

Due to the complexity of the industrial robot, it is common practice to estimate approximate models for various purposes. By, for example, using a low-frequency excitation, elastic effects have a minor influence and a nonlinear model of the rigid body dynamics can be estimated using least squares techniques. This is a much studied problem in the literature, see, e.g., Kozlowski (1998) for an overview. Taking elastic effects into account makes the identification problem much harder. The main reason is that only a subset of the state variables now are measured such that linear regression cannot be used. One option could be to add sensors during the data collection to measure all states, even though accurate measurements of all states are not at all easy to obtain (if even possible) and such sensors are probably very expensive (for example laser trackers). The data collection will also be more involved with additional signals to measure, mounting of sensors, and so on, which will make such a solution harder to use as a simple way of tuning numerous robots, as well as re-tuning robots at the customer site.

It is common to study the local dynamic behavior around certain operating points (x_0, u_0) and there estimate parametric or nonparametric linear models (see, e.g., Albu-Schäffer and Hirzinger, 2001; Behi and Tesar, 1991; Johansson et al., 2000; Öhr et al., 2006). One application area for these linear models is control design, where a global (feedback) controller is achieved through gain scheduling. The linear models can also be used for the tuning of elastic parameters in a global nonlinear robot model, which is **the adopted solution in this thesis:**

- The local behavior is considered by estimating the nonparametric frequency response function (FRF), $\widehat{G}^{(i)}(\omega_k)$, with frequencies $\omega_k,\ k = 1, \ldots, N_f$, of the

system in a number of operating points, $(x_0^{(i)}, u_0^{(i)}),\ i = 1, \ldots, Q$.

- Next, the nonlinear parametric robot model (1.1) is linearized in each of these operating points, resulting in Q parametric FRFs, $G^{(i)}(\omega_k, \theta),\ i = 1, \ldots, Q$.

- Finally, the optimal parameter vector $\hat{\theta}$ is obtained by minimizing the discrepancy between the parametric FRF, $G^{(i)}(\omega_k, \theta)$, and the estimated nonparametric FRF, $\widehat{G}^{(i)}(\omega_k)$, for all the N_f frequencies and Q operating points.

Using an FRF-based procedure allows for data compression, unstable systems are handled without problems, it is easy to validate the model such that all important resonances are captured, and model requirements in the frequency domain are easily handled. The proposed procedure also has some possible problems. The choice of model structure (parameters) and handling local minima in the optimization are problems here as well. In addition come some difficulties with biased nonparametric FRF estimates due to closed-loop data and nonlinearities. There are also cases when even a small perturbation around an operating point can give large variations due to the nonlinearities, which makes a linear approximation inaccurate. Various aspects of this procedure are treated in Papers A–E. For details on the parametric FRFs, see also Section 4.4.4.

1.4 Outline

Part I contains an overview of robotics and system identification theory related to the results presented in the thesis. Parts of the material has previously been published in Wernholt (2004). Part II consists of a collection of papers.

1.4.1 Outline of Part I

The application studied in this thesis is the industrial robot. Chapter 2 gives an introduction to the robotics area, including modeling and control. Chapter 3 presents some system identification methods and model structures that are relevant for this thesis. A survey on system identification in robotics can be found in Chapter 4. Finally, Chapter 5 provides a conclusion and some ideas for future research.

1.4.2 Outline of Part II

This part consists of a collection of edited papers, introduced below. A summary of each paper is given, together with a short paragraph describing the background to the paper and the contribution of the present author.

Paper A: Frequency-Domain Gray-Box Identification of Industrial Robots

> Wernholt, E. and Moberg, S. (2007b). Frequency-domain gray-box identification of industrial robots. Technical Report LiTH-ISY-R-2826, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Summary: This paper describes the proposed identification procedure, where unknown parameters (mainly spring-damper pairs) in a physically parameterized nonlinear dynamic model are identified in the frequency domain, using estimates of the nonparametric FRF in different robot configurations. Different parameter estimators are compared and experimental results show the usefulness of the proposed identification procedure. The weighted logarithmic least squares estimator achieves the best result and the identified model gives a good global description of the dynamics in the frequency range of interest.

Background and contribution: The basic idea of using the nonparametric FRF for the estimation of the parametric robot model is described in Öhr et al. (2006), where the author of this thesis mainly contributed with experiment design and the nonparametric FRF estimation. The author of this thesis has continued to analyze various aspects of the identification procedure such as the nonparametric FRF estimation, the selection of optimal experiment positions, and the choice of parameter estimator. Some of these results are presented in this paper. Stig Moberg has done much of the initial work on the identification procedure (see Öhr et al., 2006). In this paper, he has served as a discussion partner, performed most of the experiments, and helped out with the experimental evaluation.

Paper B: Analysis of Methods for Multivariable Frequency Response Function Estimation in Closed Loop

> Wernholt, E. and Gunnarsson, S. (2007a). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Summary: Different methods for estimating the nonparametric multivariable FRF are proposed in the literature. In this paper, some of these methods are analyzed, both in open loop and closed loop. Expressions for the bias and covariance are derived and the usefulness of these expressions is illustrated in simulations of an industrial robot where the different estimators are compared.

Background and contribution: Studying the properties of nonparametric FRF estimators, applied to closed-loop data from industrial robots, has played an important role throughout this thesis project. An early version of the paper is also published in Wernholt and Gunnarsson (2004). Professor Svante Gunnarsson mainly assisted in discussions as thesis advisor.

Paper C: Experimental Comparison of Methods for Multivariable Frequency Response Function Estimation

> Wernholt, E. and Moberg, S. (2007a). Experimental comparison of methods for multivariable frequency response function estimation. Technical Report LiTH-ISY-R-2827, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Summary: Five different nonparametric estimation methods for the multivariable FRF are experimentally evaluated using closed-loop data from an industrial robot. Three classical estimators (H_1, joint input-output, arithmetic mean) and two estimators based on nonlinear averaging techniques (harmonic mean, geometric/logarithmic mean) are considered. The estimators based on nonlinear averaging give the best results, followed by the arithmetic mean estimator, which gives a slightly larger bias. The joint input-output estimator, which is asymptotically unbiased in theory, turns out to give large bias errors for low frequencies. Finally, the H_1 estimator gives the largest bias for all frequencies.

Background and contribution: After writing Paper B, the immediate question was how the different estimators would turn out in practice. Stig Moberg mainly assisted by collecting all experimental data.

Paper D: Estimation of Nonlinear Effects in Frequency-Domain Identification of Industrial Robots

> Wernholt, E. and Gunnarsson, S. (2007b). Estimation of nonlinear effects in frequency-domain identification of industrial robots. *Accepted for publication in IEEE Transactions on Instrumentation and Measurement.*

Summary: This paper deals with how the nonparametric FRF estimate is affected by nonlinearities in the system. A method for detection and estimation of nonlinear distortions in nonparametric FRF estimates is applied to experimental data from the industrial robot. The results show that nonlinear distortions are indeed present and cause larger variability in the nonparametric FRF than the measurement noise contributions.

Background and contribution: After a seminar by Professor Johan Schoukens in March 2005, we got interested in studying how nonlinearities affect the estimated nonparametric FRF of industrial robots. This resulted in the conference paper Wernholt and Gunnarsson (2006a), which is extended in Paper D. Professor Svante Gunnarsson mainly assisted in discussions as thesis advisor.

Paper E: Experiment Design for Identification of Nonlinear Gray-Box Models with Application to Industrial Robots

> Wernholt, E. and Löfberg, J. (2007). Experiment design for identification of nonlinear gray-box models with application to industrial robots. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Summary: The adopted solution in the thesis is based on nonparametric FRF estimates in a number of robot configurations/positions. The selection of these positions in an optimal way is treated in this paper. From the derived Fisher information matrix, a convex optimization problem is posed. By considering the dual problem, the experiment design is efficiently solved with linear complexity in the number of candidate positions, compared to cubic complexity for the primal problem. In the numerical illustration, using an industrial robot, the parameter covariance is reduced by a factor of six by using the 15 optimal positions compared to using the optimal single position in all experiments.

Background and contribution: The optimal selection of robot configurations/positions for the experiments has been an open question for a long time during the project. During the second half of 2006, the question was finally formulated as a convex optimization problem by introducing the candidate positions and studying the Fisher information matrix. Dr. Johan Löfberg assisted in how to solve the convex optimization problem using YALMIP (Löfberg, 2004). He also came up with the idea of considering the dual problem.

Paper F: Nonlinear Gray-Box Identification of a Flexible Manipulator

> Wernholt, E. and Gunnarsson, S. (2007c). Nonlinear gray-box identification of a flexible manipulator. *Submitted to Control Engineering Practice*.

Summary: A three-step procedure for time-domain nonlinear gray-box identification of an industrial manipulator containing flexibilities is studied. The aim of the first two steps is to obtain good initial values for the third prediction error minimization step. In the first step, rigid body dynamics and friction are identified using a separable least-squares method. In the second step, initial values for flexibilities are obtained using an inverse eigenvalue method. Finally, in the last step, the remaining parameters of a nonlinear gray-box model are identified directly in the time domain using prediction error minimization. The identification procedure is exemplified using experimental data from axis one of an industrial robot. The estimated physical parameters have realistic numerical values and give a model with good correspondence to FRF measurements.

Background and contribution: This work started as a small project in 2004 together with David Törnqvist in a graduate course on system identification held by Professor Lennart Ljung. The author of this thesis continued the work by collecting experimental data, extending the model structure, and introducing the first two steps in the procedure. Early versions of the paper are also presented in Wernholt and Gunnarsson (2005, 2006b). Professor Svante Gunnarsson has mainly assisted in discussions as thesis advisor.

Related Publications

Publications of related interest, but not included in the thesis:

> Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.
>
> Gunnar, J., Wernholt, E., Hovland, G., and Brogårdh, T. (2006). Nonlinear grey-box identification of linear actuators containing hysteresis. In *Proc. 2006 IEEE International Conference on Robotics and Automation*, pages 1818–1823, Orlando, Florida.
>
> Wernholt, E. (2004). *On Multivariable and Nonlinear Identification of Industrial Robots*. Licentiate thesis no. 1131, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden.

1.5 Contributions

The main contributions of the thesis are:

- The refinement, analysis, and experimental comparison of the identification procedure in Paper A, first introduced in Öhr et al. (2006), for the identification of unknown parameters in nonlinear gray-box models using frequency-domain data.
- The analysis of asymptotic properties for the weighted nonlinear least squares estimator and the weighted logarithmic least squares estimator in Paper A.
- The analysis of multivariable nonparametric FRF estimators in Paper B. Bias and variance expressions are derived, both for open-loop and closed-loop data.
- The experimental comparison of different multivariable nonparametric FRF estimators in Paper C.
- Insight into the choices of excitation signals and averaging techniques when using nonparametric FRF estimators for closed-loop identification of an industrial robot.
- The experimental results in Paper D where it is shown that errors in the nonparametric FRF estimate due to nonlinearities are larger than the measurement noise contributions. Therefore the orthogonal random phase multisine signal is recommended together with averaging over experiments with different realizations of the random phases.
- The results in Paper E on experiment design for nonlinear gray-box models, where the selection of optimal robot configurations/positions for experiments is solved as a convex optimization problem.
- The proposed three-step identification procedure in Paper F for combined identification of rigid body dynamics, friction, and flexibilities using continuous-time nonlinear gray-box identification.

Part I

Overview

2

Robotics

In this chapter, some important properties of industrial robots will be examined. First an introduction to robotics is given in Section 2.1. Modeling of industrial robots is described in Section 2.2 and finally some aspects of the robot control problem are presented in Section 2.3.

2.1 Introduction

The word *robot* was first introduced by the Czech playwright Karel Capek in his 1920 play *Rossum's Universal Robots*, the word *robota* being the Czech word for work. The term has been applied to a great variety of devices, such as humanoids (trying to mimic humans), domestic robots like robot vacuum cleaners and robot lawn movers, underwater vessels, military missiles, autonomous land rowers, etc. Almost anything that operates with some degree of autonomy, usually under computer control, has at some point been called a robot. This thesis deals with industrial robots (see Figure 1.1 for an example), which consist of a mechanical arm with a number of joints, where each joint is actuated by an (electric) motor via a transmission. The movements are controlled by a computer system. This type of robot is often called *robot manipulator* or just *manipulator*. A widely accepted definition of a robot according to the Robot Institute of America (Spong and Vidyasagar, 1989) is:

> A robot is a re-programmable multi-functional manipulator designed to move materials, parts, tools or specialized devices through variable programmed motions for the performance of a variety of tasks.

2.1.1 Manipulator Structures

The mechanical structure of a robot manipulator consists of a sequence of rigid bodies (*links*) connected by revolute or prismatic *joints*, also called *axes*, where each joint is

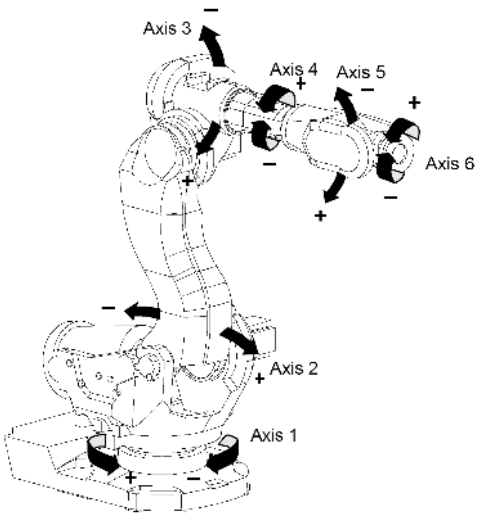


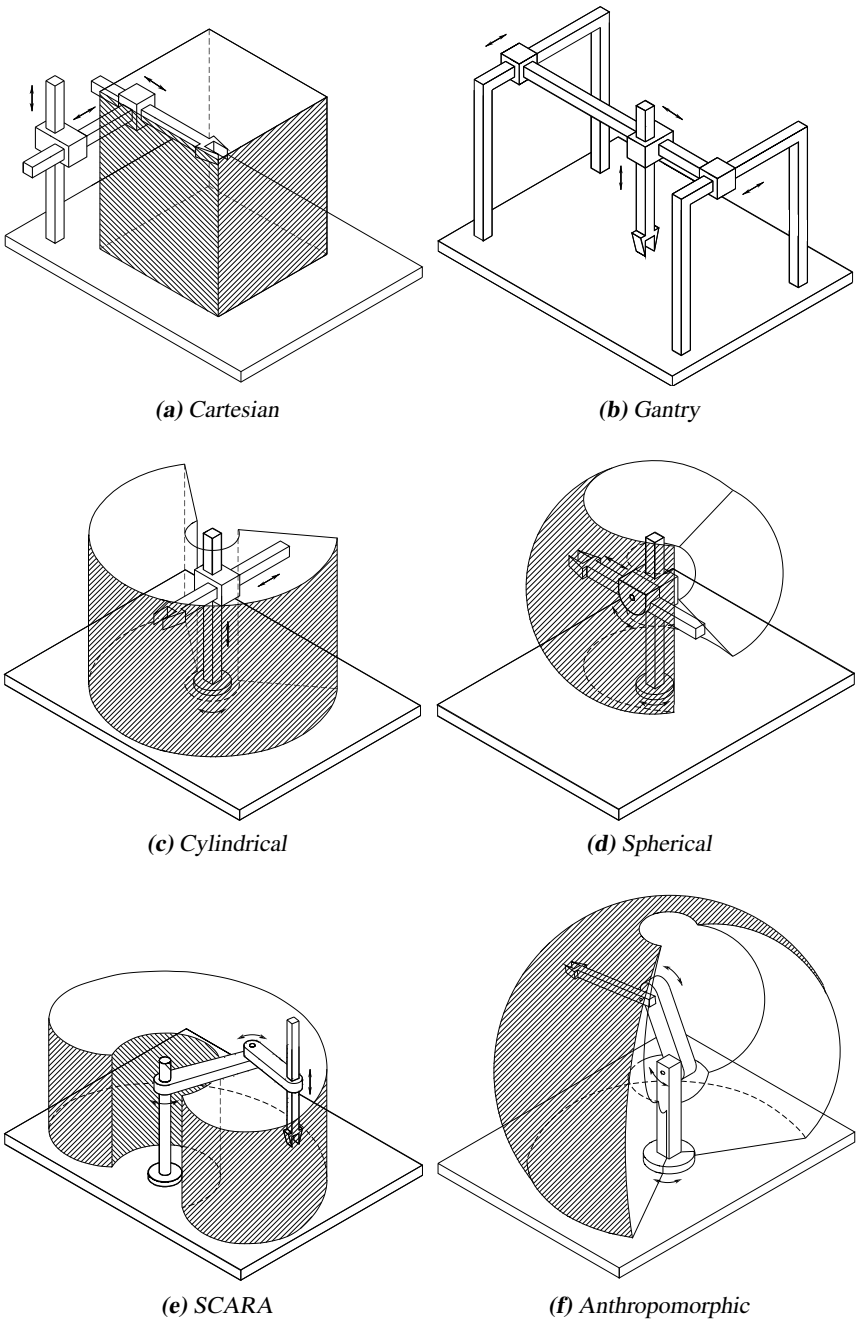
Figure 2.1: *The ABB IRB6600 robot with its six revolute joints. Axes 1-3 position the end effector (not shown in this figure) and axes 4-6 constitute the wrist.*

actuated by an electric motor combined with a gearbox[1]. Prismatic joints give relative translational motion between links, whereas the revolute joints give relative rotational motion between the links. The manipulator is characterized by an *arm* that ensures mobility, a *wrist* that confers dexterity, and an *end effector* (e.g., the spot-welding gun in Figure 1.1) that performs the required task. Usually the robot has six joints, giving six *degrees-of-freedom (DOF)* and the ability to control both the position of the end effector and its orientation in the workspace. See Figure 2.1 for an example. The portion of the environment that can be reached by the robot's end effector is called the robot *workspace*. Its shape and volume depend both on the manipulator structure and mechanical joint limits.

Many different manipulator structures exist, where the main differences are due to the two different types of joints and how these are combined. Some classifying manipulator structures can be found in Sciavicco and Siciliano (2000) and are also depicted in Figure 2.2. One can also distinguish between manipulators with an open or closed kinematic chain, which refers to how the links are connected. In this chapter, we will only consider manipulators with an open kinematic chain, also called serial type manipulators. In particular, we will restrict our treatment to the anthropomorphic manipulator, see Figures 2.1 and 2.2f, which is the most commonly used manipulator. This type of robot only uses revolute joints and is also called *elbow-type robot* due to its similarities with a human arm. See Figure 2.3 for an example of a robot with a closed kinematic chain. That type of robot is also called parallel kinematic machine (PKM).

The area in which a robot works is called a *robot cell* or *work cell*. For many applications it is common to use multiple robots in a work cell, like for the spot-welding line in

[1] The first robot generations also used hydraulic actuators. Pneumatic actuators are sometimes used in special applications, e.g., milking robots.



(a) *Cartesian* ***(b)*** *Gantry*

(c) *Cylindrical* ***(d)*** *Spherical*

(e) *SCARA* ***(f)*** *Anthropomorphic*

Figure 2.2: *Some classifying manipulator structures according to Sciavicco and Siciliano (2000). Photos courtesy of Prof. B. Siciliano.*

Figure 2.4. A dedicated device, called *positioner*, is often used to handle the work object. The positioner controls the work object and all the other devices are coordinated to move relative to the work object when it moves. The axes of the positioner are often referred to as *external axes* and are usually controlled by the robot controller.

2.1.2 Facts and Figures

The purpose of this section is to get a flavor of what industrial robots are used for, how the market is changing, and how that might affect the need for new technologies. The facts are mainly based on IFR (2006), Westerlund (2000), and Brogårdh (2007), as well as personal communication at ABB Robotics.

Industrial robots are essential components for the realization of automated manufacturing systems. The main reasons for investments in industrial robots are:

- to save costs,
- to increase productivity,
- to raise quality,
- to remain competitive in a global market, and
- to eliminate dangerous and laborious work.

Industrial robots are nowadays used in a wide range of applications, like spot welding (see Figure 2.4), arc welding, assembly, material handling, gluing and dispensing, painting, material cutting (e.g., using laser, plasma, and water jet), material removal (e.g., polishing, stub grinding, and deburring), and many more. Demands on productivity and cost cut in different types of industries further increase the use of robots and they are rapidly applied to new areas. One example is the food industry (see Figure 2.3) which is a growing market. According to IFR (2006), the world-wide installations in 2005 were distributed as:

- motor vehicles (25 %),
- automotive parts (24 %),
- electrical industries (24 %),
- chemical industries (12 %),
- metal products (5 %),
- machinery (5 %),
- food products (2 %),
- communication (2 %), and
- precision and optical products (1 %).

Figure 2.3: *Picking sausages with the ABB IRB340 FlexPicker.*

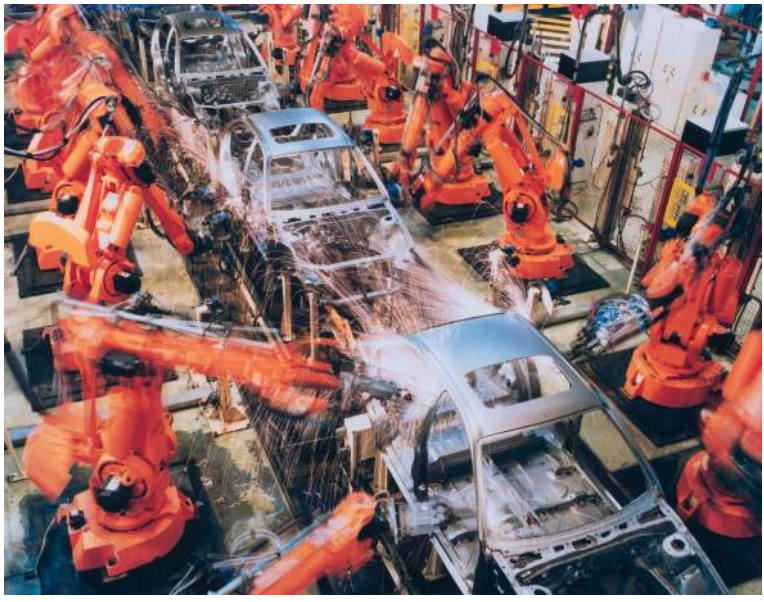


Figure 2.4: *A spot-welding line using ABB IRB6400 robots.*

The largest differences in Europe, compared to the whole world, are that the electrical industry only has 3 %, metal and food are twice as large (10 % and 4 %), and rubber and plastic (13 %) is a large market. In general, robot systems in industries other than automotive and electrical are more established in Europe than in other regions.

The density of industrial robots is, on average, about 1 robot per 10 workers in the motor vehicle industry, and 1 robot per 100 workers in the manufacturing industry, with large variations between different countries.

Since industrial robots started to be introduced in industry in the 1960s, more than 1 600 000 units have been installed (by the end of 2005). With an estimated average length of service life of 12 years, this gives a stock of about 923 000 operational industrial robots, approximately distributed as: $1/3$ in Europe, $1/3$ in Japan, less than $1/6$ in the US, and more than $1/6$ in the rest of the world.

In 2005, the world market peaked with about 126 700 new industrial robots installed. Out of these, 76 000 were supplied to Asian countries, with 50 500 in Japan only. The installed robots were distributed as: 59 % articulated (anthropomorphic), 20 % linear/Cartesian/gantry, 12 % cylindrical, and 8 % SCARA.

2.1.3 Historical Background

Industrial robots have become an essential component in the manufacturing industry of today. Still, in a historical perspective the industrial robot is a fairly new invention and a historical background could therefore be interesting. The historical background is by no means complete. To reduce the scope, the manufacturer ABB and the European market will be in focus, with some comments on other manufacturers and markets. The facts are mainly based on Westerlund (2000).

Technically, one could say that the industrial robot originates from hydraulic assembly machines that arrived in the 1950s and from the NC (numerically controlled) machines. The first industrial robot, a Unimate robot from the company Unimation, was installed in 1961 to serve a die casting machine in General Motor's factory in Trenton, New Jersey. The big breakthrough came in 1964 when General Motors ordered 66 Unimate robots. In 1969, Unimation installed its first large spot-welding line with 26 robots, used for spot-welding car bodies at General Motors.

In the 1970s, more and more companies started using robots. In 1973, there was a total of about 3 000 robots in operation around the world, of which a third were produced by Unimation. At that time, 71 companies worldwide manufactured industrial robots. Technically, the robots mainly used hydraulic actuators and usually combined revolute and prismatic joints.

The industries in Japan were very quick to apply the new technology of industrial robots, which meant that they could increase productivity and take new market shares. Pretty soon they developed their own robots and in 1980, 19 000 industrial robots were manufactured in Japan by approximately 150 different manufacturers, such as Kawasaki, Yaskawa, Mitsubishi Heavy Industries, Kobe Steel, and Fanuc. Between 1980 and 1988, the number of robots in operation worldwide increased tenfold and 1988, the figure was 256 000. Of these, 175 000, or 68 %, had been installed in Japan.

Europe, apart from Scandinavia, awoke quite late to industrial robots. One reason is that there was no shortage of labor there. The first industrial robot was installed in 1967

when Svenska Metallverken in Upplands Väsby, Sweden, bought a Unimate robot, but the big breakthrough in Europe came during the latter part of the 1970s. During the 1980s, the automotive industry in Europe and America discovered that the Japanese investment in robotized spot welding produced a more consistent quality and, therefore, welding became a prioritized area. At the end of the 1980s, there was also a rapid growth in the field of assembly.

By the end of the 1980s, most manufacturers with an annual production volume below 1 000 robots had to start looking for a partner. Usually large companies bought up small, specialized companies. During the last decade of the twentieth century, the largest companies on the European market were ABB, Fanuc, Yaskawa/Motoman, KUKA, Comau, and Renault Automation/ACMA. ABB was more than twice as big as its closest rival on the European market. Today, the three largest world-wide robot manufacturers are Yaskawa/Motoman, Fanuc, and ABB. The European companies KUKA, Comau, and Reis are also large on the European market.

In the mid 1980s, a big shift in technology took place among the robot manufacturers when AC motors replaced the DC motors. AC motors offered better cooling, which meant that the performance could be increased. Continuous improvements in the robot control system also took place. To give a flavor of the technical development within the robotics area over the last 30 years, some milestones for the robot manufacturer ABB Robotics (ABB, 2007) are given:

Some Milestones for ABB Robotics

1974: Deliver its first industrial robot, called IRB6, with a lifting capacity of 6 kg. This is the world's first microcomputer controlled electrical industrial robot and DC motors are used in combination with harmonic drive planetary gears.

1979: World's first electrical robot for spot welding (IRB60).

1983: New control system, called S2, which introduces the concepts of TCP (tool center point) and the joy stick.

1986: S3, the third generation of controllers, as well as IRB2000, a 10 kg robot which is the first robot using AC motors, and has backlash-free gearboxes, a large working range, and great accuracy.

1986-1994: Expands through acquisition of, e.g., Trallfa, Cincinatti Milacron, Graco, and ACMA.

1991: IRB6000, a large 200 kg robot, is introduced. This first modular robot (e.g., possible to change the wrist) becomes the fastest and most accurate spot-welding robot on the market.

1994: S4, the fourth generation of the control system. The main contributions are the new programming language, RAPID™, and a model-based motion control system with full dynamic models, which greatly improves the robot performance with respect to accuracy and cycle time.

1998: Launch of the IRB340 robot, see Figure 2.3, the world's fastest pick and place robot.

2002: ABB becomes the first company in the world to sell 100 000 robots.

2004: IRC5, ABB's latest controller with many new features. One powerful innovation is the MultiMove™, allowing fully synchronized control of up to four robots using the same controller. The IRC5 controller and its ABB robot family can be seen in Figure 2.5.

Figure 2.5: *ABB robot family with the IRC5 controller.*

One might wonder what the future holds for the industrial robot? According to Westerlund (2000), industrial robots are today employed in roughly 20 different fields, with around 900 potential fields of use in the future. It could be that a "robot revolution" lies ahead of us. However, entering these new fields will be a great challenge for robot manufacturers. Among several things, the robot must, to a greater extent, be able to perceive what is going on in the environment and then additional sensors will be needed. The robot must also be user friendly and the software much easier to program. Probably a lighter mechanical structure will be needed and, above all, the price must be reduced. See also Brogårdh (2007) for some scenarios on future robot control development.

2.2 Modeling

The robot system consists of a mechanical arm, actuators, sensors, and a robot controller. The robot motion control problem is treated in Section 2.3, but the remaining parts of the robot system will now be briefly described and modeled, starting with the actuators and sensors. Next, the modeling of the mechanical arm will be carried out at different levels of complexity, starting with rigid body kinematics and dynamics, then adding flexibilities to better mimic the behavior of a real robot. These different modeling levels have corresponding system identification levels, see Chapter 4.

The material in this section is mainly based on Sciavicco and Siciliano (2000), and Spong and Vidyasagar (1989).

2.2.1 Actuators and Sensors

Two basic component types in the robot system are *actuators* and *sensors*. It is, of course, outside the scope of this thesis to give an overview of all different actuators and sensors used in various types of robot systems, including a description of their pros and cons. Here, the focus will be on the actuators and sensors used for the particular industrial

robots used in the experiments, i.e., robots in the ABB IRB6600 series. See, for example, Sciavicco and Siciliano (2000, pp. 295-320) for a more detailed overview.

Actuators

To enable movements of the robot arm, each joint is actuated by a motor through a transmission gearbox. Each joint also has a brake that is used for emergency stops and when the robot is turned off. The robot actuating system has some power supply and a power amplifier for each motor.

The *power amplifier* (or power/torque controller) has the task of modulating the power flow (provided by the power supply) according to a control signal and transmit this power to the motor in terms of suitable force and flow quantities. For electric motors, it is common to use transistor amplifiers which are suitably switched by using pulse-width modulation techniques.

The joint motion typically demands low speeds with high torques. A motor usually provides the opposite, i.e., high speeds with low torques. To overcome this, a *transmission* (gearbox) is used. Various types are used depending on the robot structure, desired performance, etc. Today compact gearboxes/speed reducers, e.g., harmonic drives and different types of planetary gears (RV, Cyclo), are popular due to their low backlash, compact size, and large gear ratio (often between 50–450). The robots used in the experiments are equipped with RV gears. Neglecting inertia and power loss, a gearbox has the following relations between the high speed (hs) and low speed (ls) sides

$$q_{ls} = r_g q_{hs}, \quad \tau_{ls} = \frac{1}{r_g}\tau_{hs},$$

where q_{ls} and q_{hs} are the angular positions, τ_{ls} and τ_{hs} are the torques, and r_g is the (inverse) gear ratio ($r_g \ll 1$). Using a gearbox will reduce the nonlinear coupling terms in the dynamic model but at the same time introduce gearbox flexibilities, backlash and friction. In some rare cases the motor is directly connected to the joint, without the use of any transmission, which is called *direct drive*. Using direct drive will make the coupling effects significant and in addition the control of the motors will be more difficult. Direct drive has not made any impact in robotics due to cost, the size of the motors (as well as the brakes), and a more difficult control problem. In addition, more accurate and expensive sensors will be needed to measure the joint angular position, compared to measuring the motor angular position before the gearbox (resolution increased by the gear ratio).

As was mentioned in Section 2.1.1, the motors are exclusively electric, even though the first robot generations used hydraulic actuators. Pneumatic actuators are sometimes used in special applications, e.g., milking robots. Here, only electric motors will be considered and, in particular, AC permanent magnet motors since these are used as actuators for the industrial robots used in the experiments. *AC permanent magnet motors* are extremely fast, compact, and robust. A drawback, however, is that the generated torque changes periodically with the rotor position. The resulting torque ripple is, according to Holtz and Springob (1996), "caused by distortion of the stator flux linkage distribution, variable magnetic reluctance at the stator slots, and secondary phenomena" (e.g., the power amplifier). The ripple caused by the variable magnetic reluctance is proportional to the current, which usually can be approximated by the commanded torque, τ_c, from the

Table 2.1: *Numerical values for the disturbances in* (2.1) *and* (2.2)*. Values partly from identification carried out by Uddeholt (1998).*

n	a_n (Nm)	$\phi_{a,n}$ (rad)
3	0.06	1.45
36	0.11	0.9
72	0.08	0.2
n	b_n (-)	$\phi_{b,n}$ (rad)
18	0.04	0
n	c_n (mrad)	$\phi_{c,n}$ (rad)
2	0.40	-1.84
10	0.20	0.81
12	0.22	2.94
48	0.18	2.94

robot controller (neglecting the fast power controller). Since the torque ripple is periodic in the motor angular position q_m, it can be modeled as a sum of sinusoids like

$$v_\tau(t) = \sum_{n\in\mathbb{N}_a} a_n \sin(nq_m(t) + \phi_{a,n}) + \tau_c(t) \sum_{n\in\mathbb{N}_b} b_n \sin(nq_m(t) + \phi_{b,n}), \tag{2.1}$$

where the number of components in $\mathbb{N}_a$ and $\mathbb{N}_b$ depend on the specific motor type and the level of approximation (Gutt et al., 1996). The applied torque τ can therefore be seen as a sum of the commanded torque τ_c and a disturbance term v_τ. For an example of numerical values, see Table 2.1.

Sensors

Using sensors is of crucial importance to achieve high-performance robotic systems. There are various types of sensors available, often divided into sensors that measure the internal state of the robot (proprioceptive sensors) and sensors that provide knowledge about the surrounding environment (heteroceptive sensors). Examples of proprioceptive sensors are encoders and resolvers for joint position measurements and tachometers for joint velocity measurements. Heteroceptive sensors include, for example, force sensors for end effector force measurements and vision sensors for object image measurements when the manipulator interacts with the environment.

Due to cost and reliability requirements, sensors are only used to measure motor angular position in a standard robot today. In some special applications, vision systems and force sensors are also used (see e.g., Olsson, 2007). Due to the transmission and other sources of flexibilities, advanced dynamic models are then used to accurately estimate the movements of the robot arm.

If the motor angular position is measured by a resolver, the position values are normally obtained by using Tracking Resolver-to-Digital Converters (Hanselman, 1990). The position error, due to non-ideal resolver characteristics, can be described as a sum

of sinusoids like (Hanselman, 1990)

$$v_{q_m}(t) = \sum_{n \in \mathbb{N}_c} c_n \sin(n q_m(t) + \phi_{c,n}). \tag{2.2}$$

For an example of numerical values, see Table 2.1.

Remark 2.1. Due to the gearbox, each joint will have two angular positions, q_m (motor) and q_a (arm), which are related by the gear ratio as well as elastic effects. For simplicity, the kinematics and rigid body dynamics, described next, will be expressed in q_a. The gearbox dynamics should be added afterwards.

2.2.2 Kinematics

Kinematics of a robot refers to the geometric relationship between the joint variables and the end effector position and orientation in task space. The motion of the end effector in task space is usually defined in Cartesian coordinates with respect to a reference frame.

To be able to describe the robot kinematics in a convenient way, various coordinate systems are needed. Therefore a coordinate frame is attached to each link, to the base (base frame or reference frame), and to the end effector (end effector frame or tool frame).

Joint coordinates are given by the vector $q_a = (q_{a1}, q_{a2}, \ldots, q_{an})^T$, where n is the number of joints. A realization of the vector q_a is called a *configuration* of the robot.

The position of the tool frame, the *tool center point (TCP)*, can be expressed by a vector $x \in \mathbb{R}^3$. The orientation of the tool frame with respect to the base frame is represented by a *rotation matrix*, $R \in \mathbb{R}^{3\times3}$, with the properties $R^T R = I$, $\det R = +1$. Even though the matrix R has 9 elements, it is possible to parameterize it by a parameter vector γ of lower dimension. Many representations occur in the literature, like the Euler angles and roll-pitch-yaw angles (using 3 parameters), and axis/angle and unit quaternions (using 4 parameters). Using only three parameters will give singularities for certain orientations, which would be cumbersome. For details on different representations, see, for example, Spong and Vidyasagar (1989) and Funda et al. (1990).

Position Kinematics

The *forward kinematic problem* is to determine the mapping

$$X = \begin{bmatrix} x(q_a) \\ \gamma(q_a) \end{bmatrix} = f_{kin}(q_a), \tag{2.3}$$

from joint space to task space, where X also is called the robot *location*. The computation of this function is quite straightforward for a serial link robot and can be done iteratively from the base frame to the first link, then on to the second, and so on until the tool frame is reached. In each step, the relation is determined by geometric properties of the links and a single joint variable.

The *inverse kinematic problem* is to determine the inverse of the mapping, i.e., given a position and rotation of the tool frame calculate the corresponding robot joint configuration. This is in general a much harder problem since it is not always possible to find a closed-form solution and there may exist multiple or even infinitely many solutions.

Many of these problems can, however, be avoided by a careful design of the manipulator structure.

For a parallel kinematic machine (PKM), it is the other way around, i.e., the inverse problem is fairly easy and the forward kinematic problem is really hard to solve (Khalil and Dombre, 2002).

Velocity Kinematics

The *velocity kinematics* gives the relationship between the joint velocities and the corresponding end effector linear and angular velocities. Similar to (2.3) the velocity kinematics can be written as

$$V = \begin{bmatrix} v \\ \omega \end{bmatrix} = J(q_a)\dot{q}_a, \tag{2.4}$$

where $J(q_a) = \frac{\partial f_{kin}}{\partial q_a}(q_a) \in \mathbb{R}^{6\times n}$ is the *manipulator Jacobian*, and V represents the linear and angular velocities of the tool frame relative to the base frame. Equation (2.4) is obtained from (2.3) by differentiation with respect to time, see Spong and Vidyasagar (1989) for details.

The Jacobian is an important quantity in the analysis and control of robot motion. Since it is a function of the configuration q_a, those configurations for which it looses rank are of special interest. They are called *singularities* and can be interpreted as points in the workspace where a serial type robot looses one or more degrees of freedom. When planning the robot motion, singular points should be avoided. The Jacobian also describes the transformation from tool contact forces to corresponding joint torques.

2.2.3 Rigid Body Dynamics

A dynamic robot model describes the time evolution of the robot joints as a function of applied torques and forces. The main focus of this thesis is identification of unknown parameters in this type of model, and therefore the structure of dynamic robot models is of interest.

There are two common methods to obtain the dynamic model. The first is based on the *Euler-Lagrange* formulation and is systematic and conceptually simple. The second method is based on the *Newton-Euler* formulation and allows obtaining the model in a recursive form. See, for example, Spong and Vidyasagar (1989) for details. The resulting equations describing the rigid body dynamics are given by

$$M_a(q_a)\ddot{q}_a + c_a(q_a, \dot{q}_a) + g_a(q_a) + \tau_{fa}(\dot{q}_a) = \tau_a, \tag{2.5}$$

where $M_a(q_a)$ is the *inertia matrix*, $c_a(q_a, \dot{q}_a)$ is referred to as the velocity dependent term, containing the centrifugal and Coriolis effects, $g_a(q_a)$ is the gravitational term, $\tau_{fa}(\dot{q}_a)$ is the friction torque, and τ_a is the vector of applied torques.

The dynamics of (2.5) have a number of important properties that are helpful in the analysis and design of the control system (Sciavicco and Siciliano, 2000; Spong and Vidyasagar, 1989). Of particular interest here is the *linearity* with respect to the *dynamic parameters* (neglecting τ_{fa}), which are sometimes called the *standard inertial parameters*.

Each link gives ten inertial parameters: body mass, mass location, and the inertia matrix (only six elements due to symmetry). It is common to use a friction model

$$\tau_{fa}(\dot{q}_a) = F_v \dot{q}_a + F_c \operatorname{sign}(\dot{q}_a), \tag{2.6}$$

which gives two additional parameters for each link, but still an expression that is linear in the parameters. The robot dynamic model (2.5) can then be rewritten as

$$H_{rb}(q_a, \dot{q}_a, \ddot{q}_a)\theta_{rb} = \tau_a, \tag{2.7}$$

or as the energy difference model

$$\Delta h_{rb}(q_a, \dot{q}_a)\theta_{rb} = \Delta \mathcal{H}_{rb} = \mathcal{H}_{rb}(t_b) - \mathcal{H}_{rb}(t_a) = \int_{t_a}^{t_b} \tau_a^T \dot{q}_a dt, \tag{2.8}$$

where $\theta_{rb} \in \mathbb{R}^{12n}$ is the parameter vector and n is the number of links (Kozlowski, 1998). $\mathcal{H}_{rb}$ is the total energy of the system. Equations (2.7) and (2.8) are extensively used for identification of rigid body dynamics, see Section 4.3 for further details.

2.2.4 Flexible Body Dynamics

Traditionally, the dynamic models used for control are either entirely rigid (An et al., 1988), or only flexible joint models are considered, i.e., elastic gear transmission and rigid links (Albu-Schäffer and Hirzinger, 2000; Spong, 1987). This is usually sufficient (depends on the required performance) for a traditional robot structure.

A trend in robotics nowadays is to build lightweight robots, where the weight is reduced but with preserved payload capabilities. The load-to-mass ratio usually depends on the size of the robot, where, e.g., the large ABB IRB6600 robot can handle $225\,\text{kg}$ and weights about $1\,750\,\text{kg}$, whereas their smallest robot IRB140 can handle $5\,\text{kg}$ and weights $98\,\text{kg}$. The new ABB IRB6620 spot-welding robot illustrates this trend since it handles $150\,\text{kg}$ but only weights $900\,\text{kg}$. There are also examples of extreme lightweight robots, such as the 7 DOF DLR robot in Albu-Schäffer and Hirzinger (2001), which handles $7\,\text{kg}$ and weights $18\,\text{kg}$. This includes actuators as well as many additional sensors.

Lightweight robots are used today in a variety of applications, ranging from space robotics to less known tasks like exploration of hazardous environments or nuclear waste retrieval. Service robotics and health care are two other areas of application. A lightweight robot, such as the 7 DOF DLR robot, will of course not have the same performance as a traditional industrial robot regarding, for example, workspace, position accuracy, acceleration, and stiffness. All these properties are crucial for most industrial applications. There are however many reasons for building lightweight robots, such as cost reduction, reduced power consumption, and safety issues (reduced mass of moving parts). A lighter mechanical structure will be cheaper since less metal is needed, and smaller actuators can be used. Due to these reasons, it is also interesting to reduce the weight of industrial robots. However, a lighter robot, will result in a weaker (more compliant) mechanical structure and enhance the elastic effects of the materials. In addition to joint flexibilities, also elastic effects in the link structure then become important.

The flexible joint models and the elastic models will now be briefly described.

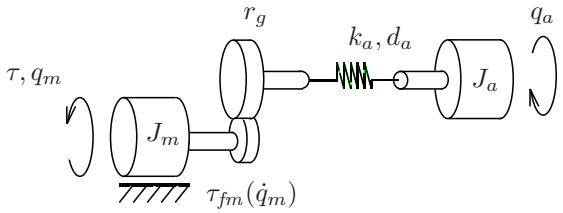


Figure 2.6: *The two-mass flexible joint model of the robot arm.*

Flexible Joint Models

Considering only one joint, a flexible joint model results in a *two-mass flexible model* according to Figure 2.6, where J_m and J_a are the moments of inertia of the motor and arm respectively, r_g is the (inverse) gear ratio (typically $r_g \ll 1$), k_g and d_g are the spring stiffness and damping (modeling the flexibility), $\tau_{fm}(\dot{q}_m)$ is the motor friction, and τ is the motor torque. The equations describing the dynamics are

$$J_m\ddot{q}_m + \tau_{fm}(\dot{q}_m) + r_g\tau_g = \tau, \tag{2.9a}$$
$$J_a\ddot{q}_a = \tau_g, \tag{2.9b}$$
$$k_g(r_g q_m - q_a) + d_g(r_g\dot{q}_m - \dot{q}_a) = \tau_g, \tag{2.9c}$$

The two-mass flexible model can be generalized to multivariable systems. A commonly adopted approximation then is to model the motors as simple rotating inertias, ignoring that the motors actually are moving due to the movements of the arm. This simplified model is obtained under the assumption that the kinetic energy of each rotor is mainly due to its own rotation, as discussed in Spong (1987). Gyroscopic forces between each rotor and the other links are then neglected. Such an approximation introduces minor errors for a traditional industrial robot which is fairly rigid. However, for lightweight robots where the masses of a flexible link and of its actuator could be comparable, neglecting or oversimplifying the dynamic effects of the motors could yield severe errors (Bascetta and Rocco, 2002).

For the generalized two-mass flexible model, the multivariable rigid body dynamics (2.5) is combined with a motor and a spring-damper pair for each joint, giving the dynamic equations

$$M_m\ddot{q}_m + \tau_{fm}(\dot{q}_m) + r_g\tau_g = \tau, \tag{2.10a}$$
$$M_a(q_a)\ddot{q}_a + c_a(q_a, \dot{q}_a) + g_a(q_a) = \tau_g, \tag{2.10b}$$
$$k_g(r_g q_m - q_a) + d_g(r_g\dot{q}_m - \dot{q}_a) = \tau_g, \tag{2.10c}$$

where τ now is the vector of applied motor torques, q_a is the vector of arm joint variables, and q_m is the vector of motor joint variables. Diagonal matrices describing the joint dynamics are defined as

$$M_m = \operatorname{diag}\{J_{m1}, \ldots, J_{mn}\}, \quad d_g = \operatorname{diag}\{d_{g1}, \ldots, d_{gn}\},$$
$$k_g = \operatorname{diag}\{k_{g1}, \ldots, k_{gn}\}, \quad r_g = \operatorname{diag}\{r_{g1}, \ldots, r_{gn}\},$$

and $\tau_{fm}(\dot{q}_m) = \begin{bmatrix}\tau_{fm1}(\dot{q}_{m1}) & \cdots & \tau_{fmn}(\dot{q}_{mn})\end{bmatrix}^T$. See also Spong (1987) for a derivation of the dynamic equations.

Remark 2.2. The arm friction $\tau_{fa}(\dot{q}_a)$ in (2.5) is located to the motor side and incorporated in (2.10) as $\tau_{fm}(\dot{q}_m) = r_g\tau_{fa}(r_g\dot{q}_m)$.

Elastic Models

An elastic link structure is actually described by partial differential equations, characterized by an infinite number of degrees of freedom. Obviously, dealing directly with infinite dimensional models is impractical both for estimation, simulation, and control design purposes. Hence it is necessary to introduce methods to describe the elasticity with a finite number of parameters. Three different approaches are generally used: *assumed modes*, *finite elements* and *lumped parameters*. See Theodore and Ghosal (1995) for a comparison of the first two approaches and Khalil and Gautier (2000) for an example of the last approach. Assumed modes are also treated in Bascetta and Rocco (2002), which, in addition, gives a good overview of the three different approaches. In this thesis, the lumped parameters approach will be used, as will be described next.

Extended Flexible Joint Models

Consider now the case where each elastic link is divided into a number of rigid bodies, connected by spring-damper pairs. The flexible joint model (2.10) is then extended as

$$M_m\ddot{q}_m + \tau_{fm}(\dot{q}_m) + r_g\tau_g = \tau, \tag{2.11a}$$

$$M_{ae}(q_a, q_e)\begin{bmatrix}\ddot{q}_a\\ \ddot{q}_e\end{bmatrix} + c_{ae}(q_a, q_e, \dot{q}_a, \dot{q}_e) + g_{ae}(q_a, q_e) = \begin{bmatrix}\tau_g\\ \tau_e\end{bmatrix}, \tag{2.11b}$$

$$k_g(r_g q_m - q_a) + d_g(r_g\dot{q}_m - \dot{q}_a) = \tau_g, \tag{2.11c}$$

$$-k_e q_e - d_e\dot{q}_e = \tau_e, \tag{2.11d}$$

where the additional variables q_e describe the angular motion between the rigid bodies due to elastic effects, k_e and d_e are defined similarly as k_g and d_g, and (2.11b) describes the dynamics of the arm, including elastic effects (with obvious changes in notation compared to (2.10b)). This model can easily be transformed to a nonlinear state-space model (cf. (1.1)) by using the fact that M_m and $M_{ae}(q_a, q_e)$ always are invertible (Spong and Vidyasagar, 1989). Introducing a state vector x and an input vector u as

$$x = \begin{bmatrix}q_m\\ q_a\\ q_e\\ \dot{q}_m\\ \dot{q}_a\\ \dot{q}_e\end{bmatrix}, \quad u = \tau, \tag{2.12}$$

gives

$$\dot{x} = f(x,u,\theta) = \begin{bmatrix} \dot{q}_m \\ \dot{q}_a \\ \dot{q}_e \\ M_m^{-1}(u - \tau_{fm}(\dot{q}_m) - r_g\tau_g) \\ M_{ae}^{-1}(q_a,q_e)\left[\begin{bmatrix}\tau_g \\ \tau_e\end{bmatrix} - c_{ae}(q_a,q_e,\dot{q}_a,\dot{q}_e) - g_{ae}(q_a,q_e)\right] \end{bmatrix}, \quad (2.13a)$$

$$y = h(x,u,\theta) = \dot{q}_m, \quad (2.13b)$$

where θ is a vector of unknown parameters in the model, usually the diagonal entries in k_g, d_g, k_e, and d_e. The measurement equation (2.13b) depends on what is considered as measurement signals, but usually $y = \dot{q}_m$ is used in the thesis.

This model, referred to as the *extended flexible joint model*, is described in Moberg and Hanssen (2007) and used in Paper A and Paper E. See also Section 4.4.4 for some notes on how (2.13) is linearized and used in the proposed identification procedure.

2.3 Control

The robot motion control problem is the problem of determining the time history of inputs to the actuators required to cause the end effector to execute a commanded motion. In general, the motion control problem is divided into three stages,

- motion planning,
- trajectory generation, and
- trajectory tracking.

See also Figure 2.7. Each stage will now be briefly described.

2.3.1 Motion Planning

Motion planning on the highest level involves finding a *path* in the operational space, which denotes the locus of Cartesian points or some other representation, that the end effector has to follow in the execution of the assigned motion. This involves meeting requirements from the process (e.g., gluing, dispensing, or arc welding) both on position accuracy and coordinated motion with the process (e.g., to make sure that the size of the bead remains constant during the planned motion).

In industrial applications this part is either done by using a teach pendant (see Figure 2.8) or by using a system separate from the robot control system, for example a CAD/CAM tool.

2.3.2 Trajectory Generation

The goal of the trajectory generation stage is to generate reference inputs to the control system which ensures that the manipulator executes the planned trajectories.

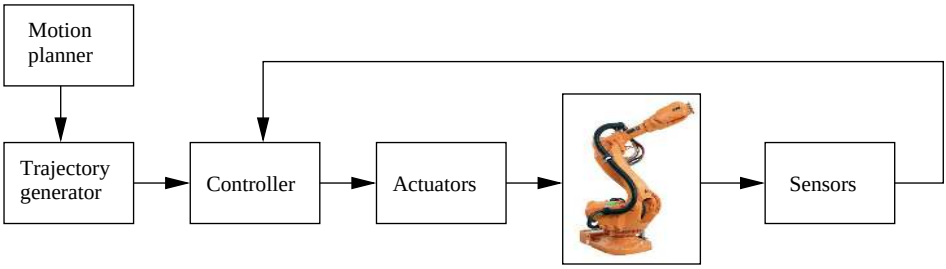


Figure 2.7: *Block diagram showing the components in the robot control problem.*

Figure 2.8: *An example of a teach pendant: The ABB IRC5 FlexPendant.*

Given the path representation in Cartesian space from the motion planning stage, this path first has to be transformed into a path representation in joint space. This is in general not possible to do analytically. Instead the path must be transformed using the *inverse kinematic model* of the robot manipulator at discrete points. These points are then interpolated, for example by using splines as in Nyström and Norrlöf (2003).

The generated path is a pure geometric description of the motion. A *trajectory*, on the other hand, is a path on which a time law is specified. Trajectory generation therefore deals with the problem of generating a trajectory with position, speed, acceleration, and jerk (derivative of acceleration) as functions of time. The trajectories must be *feasible*, i.e., the manipulator must be able to follow the generated trajectories. The actual robot dynamics and kinematics must be taken into account as well as various constraints on motor torques, speeds, accelerations, structural forces, and so on. The goal often is to find optimal paths where the maximum speed and acceleration are used, given demands from the process on accuracy and other process constraints. Therefore, the trajectory generation problem is important for the performance of the robot.

2.3.3 Trajectory Tracking

The trajectory tracking problem can be defined as the problem of controlling the robot joints according to the trajectory calculated by the trajectory generator.

There are many control techniques and methodologies that can be applied to the control of manipulators. The control method and its implementation can have significant impact on the performance of the manipulator and consequently on the range of possible applications. The mechanical design of the manipulator will also influence the type of controller needed. For example, the nonlinear coupling effects between different joints varies greatly between a Cartesian robot and an elbow-type robot. If a gearbox is used or not will also greatly affect the control problem, as was mentioned in Section 2.2.1. Using a gearbox will reduce the nonlinear coupling effects but at the same time introduce flexibilities, backlash and friction. Using a direct-drive robot, on the other hand, will make the coupling effects significant and in addition the control of the motors will be more difficult. Here, we will consider geared elbow-type robots. It is outside the scope of this thesis to give an overview of all available control methods. For more details on different aspects of robot control, there are numerous references, such as Spong and Vidyasagar (1989), Craig (1989), Sciavicco and Siciliano (2000), and Spong et al. (1993). However, to get a feeling for the problem, some words will be said about a commonly used control technique.

A standard procedure in robotics today is to measure only the motor angular position (to increase performance, additional sensors are sometimes added). Due to the gearbox flexibilities, there will be some dynamics between the motor angular position, q_m, and the arm angular position, q_a. Therefore, the generated joint trajectory, q_a^{ref}, must be transformed to a corresponding motor trajectory, q_m^{ref}, using a flexible dynamic model. A common architecture for the robot controller is shown in Figure 2.9, using both feedback and feedforward controllers. X is the end-effector Cartesian location which should follow the path created by the motion planner.

Since the trajectory is known beforehand, it can be used in a feedforward controller. The generated trajectories are feasible so the feedforward controller should ideally give

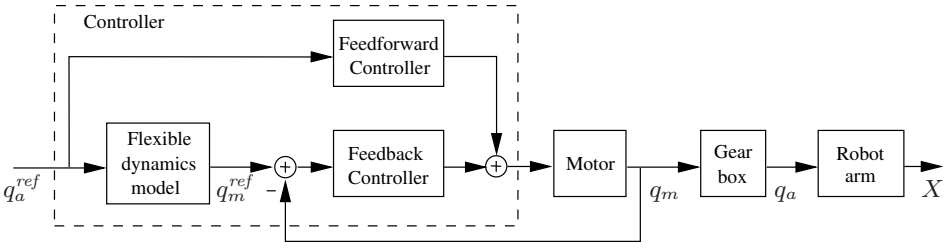


Figure 2.9: *Block diagram showing the controller and a robot with gearbox flexibilities. The dashed block corresponds to the Controller block in Figure 2.7.*

zero tracking error on the motor side, $e(t) = q_m^{ref}(t) - q_m(t)$. Due to model errors and disturbances, there will still be a non-zero tracking error and this is handled by the feedback controller. A common choice for the feedforward controller is to use *computed torque* (Spong and Vidyasagar, 1989) which in principle means the inverse of the robot dynamics. For a rigid robot described by (2.5), the feedforward controller would output

$$\tau_{ffw} = M_a(q_a^{ref})\ddot{q}_a^{ref} + c_a(q_a^{ref}, \dot{q}_a^{ref}) + g_a(q_a^{ref}) + \tau_{fa}(\dot{q}_a^{ref}), \tag{2.14}$$

where q_a^{ref} is the reference trajectory. This is a complicated expression that must be evaluated at a high sampling rate and is therefore often approximated by removing slowly varying parts. These are instead handled by the feedback controller. Adding flexibilities further complicates the feedforward controller.

For the feedback controller, the simplest control strategy is called *independent joint control* and means that each joint axis is controlled as a SISO system. Coupling effects are then treated as disturbances and what is left to control is a multi-mass flexible SISO system (for simplicity, the two-mass flexible system (2.9) is used as illustration). A fairly simple PID controller will usually give satisfactory behavior. Since the feedforward controller is expected to give the main performance, the feedback controller is mainly tuned for disturbance rejection. The controller must be robust to model errors as well as uncertainties in the robot load and variations in the workspace. The arm inertia (J_a in (2.9)) will, for example, vary as a function of the configuration, q_a. To improve the performance of the feedback controller, sometimes a gain scheduling technique is used with the varying arm inertia, J_a, as one of the scheduling variables. A second step would be to include some coupling effects of major influence. For example, axes two and three are for some configurations strongly coupled.

In some applications the disturbances come from the process, such as contact forces during e.g., stub grinding, deburring, and drilling. The feedback controller can then be used to modify the stiffness of the robot. This is e.g., treated in Olsson et al. (2007) where a force sensor is added to measure the contact forces and used in an additional feedback loop. See also Olsson (2007) for additional details and Brogårdh (2007) for other examples of using force sensors.

3

System Identification

System identification deals with the problem of estimating mathematical models of dynamic systems using measurement data. Section 3.1 gives a brief review of some basic properties of linear time-invariant dynamic systems. This is described in many textbooks (see, e.g., Ljung, 1999; Rugh, 1996) and can be skipped by readers already familiar with the subject. The system identification procedure is described in Section 3.2 and details on model structures and how to compute the estimates are given in Sections 3.3 and 3.4, respectively. Identification in closed loop puts additional requirements on the identification methods, which is briefly discussed in Section 3.5. The quality of the estimated models is usually measured by the statistical concepts bias and variance, which is dealt with in Section 3.6. Finally, the chapter is ended by some notes on experiment design in Section 3.7 and nonparametric FRF estimation of nonlinear systems in Section 3.8.

3.1 Preliminaries

Consider a *continuous-time* system with a scalar input $u(t)$ and a scalar output $y(t)$. The system is said to be *time-invariant* if its response to a certain input signal does not depend on absolute time. It is said to be *linear* if its output response to a linear combination of inputs is the same linear combination of the output responses of the individual inputs. In addition, it is said to be *causal* if the output at a certain time only depends on the input up to that time. In the following, only causal systems will be treated.

The output of a linear time-invariant (LTI) system can be completely described by its *impulse response* $g(\tau)$ as

$$y(t) = \int_0^\infty g(\tau)u(t-\tau)d\tau. \tag{3.1}$$

Introducing the Laplace transform

$$G(s) = \mathcal{L}(g) = \int_0^\infty g(t)e^{-st}dt, \tag{3.2}$$

$$U(s) = \mathcal{L}(u), \quad Y(s) = \mathcal{L}(y), \tag{3.3}$$

allows us to write (3.1) as

$$Y(s) = G(s)U(s). \tag{3.4}$$

Replacing s with the differential operator p, $pu(t) = \frac{d}{dt}u(t)$, gives the *transfer operator* $G(p)$. (With abuse of notation, $G(p)$ will sometimes be called transfer function as well.) This makes it possible to rewrite (3.1) as

$$y(t) = G(p)u(t). \tag{3.5}$$

We will exclusively work with observations of inputs and outputs in *discrete time*. We thus assume that $y(t)$ and $u(t)$ are observed at the *sampling instants* $t_k = kT_s,\ k = 0, 1, \ldots, \infty$ where the interval T_s will be called the *sampling period*. Often in computer controlled applications, the input signal $u(t)$ is kept constant between the sampling instants like

$$u(t) = u_k, \quad kT_s \le t < (k+1)T_s, \tag{3.6}$$

which is called *zero-order hold*. Inserting (3.6) in (3.1) gives a relation between y and u at the sampling instants

$$y(kT_s) = \sum_{l=0}^{\infty} g_{T_s}(l)u_{k-l}, \tag{3.7}$$

where $g_{T_s}(l)$ is the discrete-time *pulse response* defined as

$$g_{T_s}(l) = \int_{(l-1)T_s}^{lT_s} g(\tau)d\tau, \quad l \ge 1, \tag{3.8}$$

and $g_{T_s}(l) = 0,\ l \le 0$. For discrete-time signals, we use the z-transform

$$G_{T_s}(z) = \mathcal{Z}(g_{T_s}) = \sum_{k=0}^{\infty} g_{T_s}(k)z^{-k}, \tag{3.9}$$

$$U_{T_s}(z) = \mathcal{Z}(u) = \sum_{k=0}^{\infty} u(kT_s)z^{-k}, \quad Y_{T_s}(z) = \mathcal{Z}(y). \tag{3.10}$$

Replacing z with the shift operator q, $qu(t) = u(t+T_s)$, gives the corresponding discrete-time *transfer operator*.

Evaluation of the transfer function at the point $z = e^{j\omega T_s}$ ($s = j\omega$ for continuous time) will describe the system response to a sinusoidal input $u(t) = \cos(\omega t)$,

$$y(t) = |G_{T_s}(e^{j\omega T_s})| \cos(\omega t + \arg G_{T_s}(e^{j\omega T_s})). \tag{3.11}$$

The complex-valued function

$$G_{T_s}(e^{j\omega T_s}), \quad -\pi/T_s \leq \omega \leq \pi/T_s, \tag{3.12}$$

is therefore called the *frequency response function (FRF)* or just the *frequency function*. It is common to graphically display this function as $\log |G_{T_s}(e^{j\omega T_s})|$ and $\arg G_{T_s}(e^{j\omega T_s})$ plotted against $\log \omega$ (for $0 < \omega \leq \pi/T_s$) in a *Bode plot*. For continuous-time descriptions, the frequency function $G(j\omega)$ is defined for $-\infty < \omega < \infty$.

One might wonder how the continuous-time and discrete-time frequency functions are related. In Åström and Wittenmark (1984, Chap. 4) it is shown that

$$G_{T_s}(e^{j\omega T_s}) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} G_{zoh}(j(\omega + k\frac{2\pi}{T_s})), \tag{3.13}$$

with

$$G_{zoh}(s) = \frac{1 - e^{-sT_s}}{s} G(s), \tag{3.14}$$

the new continuous-time transfer function where the sampling effects due to (3.6) are added. If the sampling is fast enough, such that the $G(j\omega),\ |\omega| > \pi/T_s$, is small, then (3.13) can be approximated as

$$G_{T_s}(e^{j\omega T_s}) \approx \frac{1 - e^{-j\omega T_s}}{j\omega T_s} G(j\omega), \tag{3.15}$$

which shows that $G_{T_s}(e^{j\omega T_s})$ and $G(j\omega)$ agree quite well when $\omega \ll \pi/T_s$. It is also possible to show that

$$|G(j\omega) - G_{T_s}(e^{j\omega T_s})| \leq \omega \cdot T_s \cdot \int_0^{\infty} |g(\tau)| d\tau, \tag{3.16}$$

where $g(\tau)$ is the continuous-time impulse response (see, for example, Ljung and Glad, 1994, p. 75). A rule of thumb is that the agreement between the frequency response functions is good enough for frequencies below one tenth of the sampling frequency ($\omega < 2\pi/(10T_s)$).

Consider now the Fourier transform, defined as

$$Y(j\omega) = \int_0^{\infty} y(t) e^{-j\omega t} dt, \tag{3.17}$$

for a continuous-time signal $y(t)$ and

$$Y_{T_s}(e^{j\omega T_s}) = \sum_{k=0}^{\infty} y(kT_s) e^{-j\omega T_s k}, \tag{3.18}$$

for a discrete-time signal $y(kT_s),\ k = 0, 1, \ldots, \infty$. These transforms are related as

$$Y_{T_s}(e^{j\omega T_s}) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} Y(j(\omega + k\frac{2\pi}{T_s})). \tag{3.19}$$

If the bandwidth of the continuous-time signal is larger than half the sampling frequency, higher frequencies will be shifted in and appear as lower frequencies. This error is called *alias error* and can be avoided by using an anti-alias filter before sampling the signal. If the bandwidth is less than half the sampling frequency, no information will be lost and the continuous-time and discrete-time Fourier transforms will coincide for $-\pi/T_s \leq \omega \leq \pi/T_s$.

Of course, in a real situation, only a limited number of data points $y(kT_s),\ k = 1, 2, \ldots, N$, are collected. It is then common to consider the Discrete Fourier Transform (DFT), defined as

$$Y(\omega_k) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} y(nT_s) e^{-j\omega_k T_s n}, \tag{3.20}$$

with the DFT frequencies

$$\omega_k = k\frac{2\pi}{NT_s}, \quad k = 1, 2, \ldots, N. \tag{3.21}$$

Depending on the properties of the original infinite-length signal, the DFT (3.20) may differ from the discrete-time Fourier Transform (3.18). Assuming a periodic signal and measuring an integer number of periods, (3.20) and (3.18) will coincide for the DFT frequencies (3.21). In most other cases, there will be *leakage errors* due to the limited time window. These errors can be reduced by using a windowed signal (typically giving less attention to samples in the beginning and the end of the data record). See, for example, Ljung (1999) and Pintelon and Schoukens (2001) for suggestions on suitable windows.

In the sequel, T_s in $G_{T_s}(q)$ is usually left out for easier notation.

3.2 The System Identification Procedure

Consider the setting in Figure 3.1, where u is the plant input and y is the measured plant output, corrupted by measurement noise v. Assume now that we, for some reason, want to model this system. The model could be used both as a means for achieving deeper knowledge about the system and as a design tool, for example as a basis for simulations or for controller design. One common use of models is prediction, where we want to predict future output values. This is, for example, used in weather forecasts.

Two main routes to obtain these models are *physical modeling* and *system identification*. Modeling involves using basic physical laws and other well-established relationships. On the other hand, system identification uses experimental data to find a suitable model. In this chapter the focus, obviously, will be on the latter route.

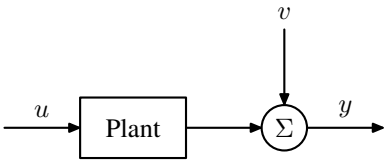


Figure 3.1: *Plant subject to system identification.*

The system identification procedure involves finding a model that describes the input-output data sufficiently well, according to some criterion. The search for this model will be carried out among a set of candidate models. System identification therefore includes several steps, like:

1. Design the experiments.
2. Collect experimental data from the system.
3. Select a set of models or the model structure to represent the system.
4. Determine the "best" model in the set, guided by the data.
5. Validate the model.

In each step, there are numerous options which inevitably cannot be described in detail here. Still, some particular options for steps 1, 3, and 4 with relevance to our studied problem will be pointed out in the following sections. For a detailed treatment, the reader is referred to, for example, Ljung (1999), Söderström and Stoica (1989), and Pintelon and Schoukens (2001).

3.3 Model Structures

As mentioned previously, the identification procedure involves a search for the best model among a set of candidate models, called the *model set*. The selection of this model set is the most important and, at the same time, the most difficult choice in the identification procedure. Physical insight and a priori knowledge about the system are often of great assistance. The model set is just a collection of models. In order to facilitate the numerical search, the model set can be parameterized by a parameter vector θ belonging to a parameter set $D_{\mathcal{M}}$. This parameterization of the model set is called a *model structure* and is denoted by $\mathcal{M}$.

Some options to consider in the selection of a suitable model structure are:

- linear versus nonlinear models,
- parametric versus nonparametric models,
- black-box versus gray-box models,
- discrete-time versus continuous-time models, and
- linear-in-the-parameters versus nonlinear-in-the-parameters.

Most of these options can be combined, giving a wide range of possible model structures to choose from. For each choice, there are a number of identification methods for the computation of the parameter estimate. The model structure selection will also greatly affect how hard the problem of computing the parameter estimate will be. Each of these options will now be briefly explained.

Consider once more the setting in Figure 3.1 and the aim to find a model describing the collected input-output data. The first choice to make is if to use a *linear* or *nonlinear*

model. Almost every real-life system is nonlinear, including the industrial robot that is studied in this thesis. The theory of nonlinear systems is very involved and nonlinear systems are therefore often approximated by linear models. This is done under the assumption that the behavior can be sufficiently well approximated by a linear model in the operating region. Whether this is a valid assumption or not depends on the application. For a detailed discussion on linear models of nonlinear systems, see Enqvist (2005) and Pintelon and Schoukens (2001).

The second choice is between a *parametric* or *nonparametric* model. A nonparametric model can, for example, be the impulse response (time domain) or the frequency response function (frequency domain) at a large number of points. The only assumption made then is that the input-output relation can be described by a linear model. For a parametric model, on the other hand, some additional knowledge and/or assumptions are used to reduce the model's degrees-of-freedom by parameterizing it by a limited number of parameters. One example could be a transfer function model, parameterized by its poles and zeros. For a parametric model, the physical insight will typically be larger and it can be seen as a concentration of information. This concentration will also reduce the variance in the estimated model, see Section 3.6. Usually a nonparametric model is simpler to create since less knowledge about the system is needed.

Assume now that the system in Figure 3.1 can be modeled by a parametric linear time-invariant model like

$$y(t) = G(p, \theta)u(t) + H(p, \theta)e(t), \tag{3.22}$$

in the *continuous-time* case and

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t), \tag{3.23}$$

in the *discrete-time* case. Note that (3.23) only is valid at the sampling instants $t = kT_s,\ k = 0, 1, \ldots, \infty$. The plant transfer function, $G(\cdot, \theta)$, describes the input-output relation and the term $H(\cdot, \theta)e(t)$, where $e(t)$ is zero mean white noise, is used to describe the measurement noise $v(t)$. $G(\cdot, \theta)$ and $H(\cdot, \theta)$ are parameterized by the parameter vector $\theta \in D_{\mathcal{M}}$. The model structure $\mathcal{M}$ can then be written more formally as

$$\mathcal{M}: \quad D_{\mathcal{M}} \ni \theta \rightarrow \mathcal{M}(\theta) = \{G(\cdot, \theta), H(\cdot, \theta)\}, \tag{3.24}$$

which means the mapping from the parameter vector θ, belonging to the parameter set $D_{\mathcal{M}}$, to a particular model $\mathcal{M}(\theta)$. For details on model sets and model structures, see Ljung (1999, pp. 107–108).

Sometimes the model structure $\mathcal{M}$ is obtained after careful modeling, with some unknown parts or parameters left to estimate. This type of model structure is called *gray box*. Similarly, a model structure whose parameters do not reflect physical considerations in the system is called *black box*.

The transfer functions $G(\cdot, \theta)$ and $H(\cdot, \theta)$ can be parameterized in many different ways, e.g., as rational functions where the parameter vector θ contains the coefficients in the numerator and denominator polynomials. In this thesis, we will mainly work with *state-space models*

$$\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) + K(\theta)e(t), \tag{3.25a}$$
$$y(t) = C(\theta)x(t) + D(\theta)u(t) + e(t), \tag{3.25b}$$

or

$$x(t+T_s) = A(\theta)x(t) + B(\theta)u(t) + K(\theta)e(t), \tag{3.26a}$$
$$y(t) = C(\theta)x(t) + D(\theta)u(t) + e(t), \tag{3.26b}$$

where here for simplicity the *innovations form* is considered (Ljung, 1999, p. 99). In this description, $x(t)$ is the state vector and $A(\theta)$, $B(\theta)$, $K(\theta)$, $C(\theta)$, and $D(\theta)$ are matrices of suitable dimensions. A nice property of state-space models is that they work equally well for multivariable systems as for scalar systems. The state-space models (3.25) and (3.26) are just choices of parameterization of the linear systems (3.22) and (3.23), respectively. Consider now the continuous-time case. Using the differential operator, (3.25) can be rewritten as

$$y(t) = G(p,\theta)u(t) + H(p,\theta)e(t), \tag{3.27a}$$
$$G(p,\theta) = C(\theta)(pI - A(\theta))^{-1}B(\theta) + D(\theta), \tag{3.27b}$$
$$H(p,\theta) = C(\theta)(pI - A(\theta))^{-1}K(\theta) + I, \tag{3.27c}$$

showing the relationship to the model (3.22). The discrete-time case is obtained by replacing p by q in the expressions. For a nonlinear system, the dynamics of a finite-dimensional system can be modeled using a nonlinear state-space model like

$$\dot{x}(t) = f(t, x(t), u(t), w(t), \theta), \tag{3.28a}$$
$$y(t) = h(t, x(t), u(t), v(t), \theta), \tag{3.28b}$$

in the continuous-time case, with obvious modifications for a discrete-time nonlinear state-space model. Here, f and h are nonlinear functions and $w(t)$ and $v(t)$ are independent random variables. Nonlinear state-space models will be used in various parts of the thesis, see, e.g., Paper A, Paper E, and Paper F.

3.4 Calculating the Estimate

Suppose now that a model structure has been selected, according to the previous section. Depending on the choice of model structure, a number of different methods exist. It is common to distinguish between parametric and nonparametric methods and also between time-domain and frequency-domain methods. Here, we will briefly discuss parametric time-domain methods, parametric frequency-domain methods, and nonparametric frequency-domain methods, which all are used in various parts of the thesis. For nonparametric time-domain methods like transient-response analysis and correlation analysis, see, e.g., Ljung (1999) and the references therein.

3.4.1 Parametric Time-Domain Methods

With time-domain methods for system identification, we here mean methods that compute an estimate of the unknown parameter vector θ using measurements of the input and output for a number of time instants, $Z^N = \{u(1), y(1), \ldots, u(N), y(N)\}$. (For simplicity, we here assume sampling period $T_s = 1$.) Several methods are available in the

literature but to limit the discussion, only the *prediction error method* (Ljung, 1999) will be considered. The idea behind the prediction error method is to find the parameters that will minimize the prediction errors,

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta), \quad t = 1, 2, \ldots, N, \tag{3.29}$$

where $\hat{y}(t|\theta)$ is the model's prediction of $y(t)$ given Z^{t-1}. For the discrete-time case (3.23), the standard predictor is given by

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t). \tag{3.30}$$

For the continuous-time case (3.22), a predictor can be calculated by, for example, first sampling the model. For the discrete-time state-space model (3.26), the predictor $\hat{y}(t|\theta)$ is given by

$$\begin{aligned}\hat{x}(t+1,\theta) =& A(\theta)\hat{x}(t,\theta) + B(\theta)u(t) + \\ &+ K(\theta)\left[y(t) - C(\theta)\hat{x}(t,\theta) - D(\theta)u(t)\right], \qquad (3.31a)\\ \hat{y}(t|\theta) =& C(\theta)\hat{x}(t,\theta) + D(\theta)u(t). \qquad (3.31b)\end{aligned}$$

For the minimization of the prediction errors (3.29), one could choose different norms. A common choice is the quadratic criterion, which gives the estimator

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta), \tag{3.32a}$$

$$V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N} \frac{1}{2}\varepsilon^2(t,\theta), \tag{3.32b}$$

with $\varepsilon^2(t,\theta)$ from (3.29). Other norms than ε^2 might be useful, for example to be robust against outliers (Ljung, 1999). The criterion is usually minimized by a numerical search method, for example the Gauss-Newton method (Ljung, 1999, Chap. 10).

A problem with many model structures is that the criterion $V_N(\theta)$ is a non-convex function with (possibly) many local minima. Good starting values for the numerical search are then needed in order to avoid getting trapped in a local minimum. In some special cases, there exist an analytical solution. If the predictor is *linear in the parameters*, it can be written as a linear regression,

$$\hat{y}(t,\theta) = \varphi^T(t)\theta, \tag{3.33}$$

where $\varphi(t)$ contains the measurement data. The minimizing parameter vector to (3.32) then is the solution to a standard least-squares problem

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta) = \left[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)y(t). \tag{3.34}$$

For state-space models, there is also the possibility to use *subspace methods* (Van Overschee and DeMoor, 1996), which basically estimate the matrices of the state-space model by solving a sequence of least-squares problems.

For a general nonlinear system as in (3.28), it is often very hard to determine a predictor on formal probabilistic grounds. In most cases there is no explicit form available for the optimal solution. This implies that a predictor must be constructed either by *ad hoc* approaches, or by some approximation of the optimal solution, e.g., by using sequential Monte Carlo techniques (Doucet et al., 2001). One simple way of deriving an *ad hoc* predictor is to disregard the process noise $w(t)$, which gives

$$\dot{x}(t) = f(t, x(t), u(t), 0, \theta), \tag{3.35a}$$

$$\hat{y}(t|\theta) = h(t, x(t), u(t), 0, \theta). \tag{3.35b}$$

This predictor is actually a simulation of the noise-free model, using the measured input. However, this only works if the model is stable since a stable predictor always is required. This type of predictor is used in Paper F.

3.4.2 Nonparametric Frequency-Domain Methods

Nonparametric frequency-domain identification methods means that the plant frequency response function (FRF) is estimated without using a certain model set described by a number of parameters.

Nonparametric estimates of the FRF give valuable information about the dynamics of a system. They are often used as an intermediate step in a parametric identification process to assess the quality of the measurements and visualize the complexity of the modeling problem. In a second step, a parametric model can be estimated by one of the following options:

1. Treat the estimated FRF as a measurement and minimize the discrepancy between the model and the estimated FRF. This is done in, e.g., experimental modal analysis (Verboven, 2002) and is also the adopted solution in this thesis, see Paper A.

2. Use the measured time-domain or frequency-domain data, as is described in Sections 3.4.1 and 3.4.3.

Most nonparametric estimation methods consider only scalar systems, like the ETFE described below. In this thesis, we will also consider multivariable FRF estimation methods.

The Empirical Transfer Function Estimate (ETFE)

Using the DFT of the input and output, $U(\omega_k)$ and $Y(\omega_k)$, the *empirical transfer function estimate (ETFE)* is defined as

$$\widehat{G}(e^{j\omega_k T_s}) = \frac{Y(\omega_k)}{U(\omega_k)}. \tag{3.36}$$

It is possible to show that the variance of this estimate does not approach zero as the number of data tends to infinity, unless the data is periodic (Ljung, 1999). Therefore, the estimate is often smoothed using a weighting function $\alpha_k(\omega)$ as

$$\widehat{G}_{smooth}(e^{j\omega T_s}) = \frac{\sum_k \alpha_k(\omega)\widehat{G}(e^{j\omega_k T_s})}{\sum_k \alpha_k(\omega)}. \tag{3.37}$$

The smoothing will reduce the variance but introduce a bias (see Section 3.6 for bias and variance expressions). The width of the smoothing, i.e., the variance/bias trade-off, is determined by the user. The ETFE is also closely related to spectral estimation.

The Multivariable Frequency Response Function Estimate

For multivariable systems, nonparametric estimation methods are not that common. The identification method that will be described here can be seen as a generalization of the ETFE to multivariable systems. Consider a measurement setup

$$y(t) = G(q)u(t) + v(t), \tag{3.38}$$

with input $u(t) \in \mathbb{R}^{n_u}$, output $y(t) \in \mathbb{R}^{n_y}$, measurement noise $v(t) \in \mathbb{R}^{n_y}$, and $G(q)$ the $n_y \times n_u$ multivariable discrete-time transfer operator. The input is assumed to be periodic and an integer number of periods of the steady-state response are collected. The following linear mapping then holds exactly in the noise-free case

$$Y(\omega_k) = G(e^{j\omega_k T_s})U(\omega_k), \tag{3.39}$$

where $Y(\omega_k)$ and $U(\omega_k)$ are the DFTs of the input and output signals and $G(e^{j\omega_k T_s}) \in \mathbb{C}^{n_y \times n_u}$ is the FRF. To be able to extract $G(e^{j\omega_k T_s})$ from data, at least n_u different experiments are needed[1]. The data vectors from $n_e \geq n_u$ different experiments will be collected into matrices (bold-face in the sequel) where each column corresponds to one experiment. The input-output relation can then be written as

$$\mathbf{Y}(\omega_k) = G(e^{j\omega_k T_s})\mathbf{U}(\omega_k), \tag{3.40}$$

where $\mathbf{U}(\omega_k) \in \mathbb{C}^{n_u \times n_e}$ and $\mathbf{Y}(\omega_k) \in \mathbb{C}^{n_y \times n_e}$. If $\mathbf{U}(\omega_k)$ has rank n_u, an estimate of $G(e^{j\omega_k T_s})$ can be formed by, e.g., using the H_1 *estimator* (Guillaume et al., 1996; Pintelon and Schoukens, 2001; Verboven, 2002)

$$\widehat{G}^{\mathrm{H}_1}(e^{j\omega_k T_s}) = \mathbf{Y}(\omega_k)\mathbf{U}^H(\omega_k)[\mathbf{U}(\omega_k)\mathbf{U}^H(\omega_k)]^{-1}, \tag{3.41}$$

where $(\cdot)^H$ denotes complex conjugate transpose.

When the measurements are corrupted by noise, (3.40) changes to

$$\mathbf{Y}(\omega_k) = G(e^{j\omega_k T_s})\mathbf{U}(\omega_k) + \mathbf{V}(\omega_k). \tag{3.42}$$

The H_1 estimator (3.41) can still be used, but the estimate will contain errors due to the noise. Assume now that n_e is a multiple of n_u such that the DFT matrices can be partitioned into M blocks of size $n_u \times n_u$ as

$$\mathbf{U}(\omega_k) = \begin{bmatrix} \mathbf{U}^{[1]}(\omega_k) & \dots & \mathbf{U}^{[M]}(\omega_k) \end{bmatrix}, \tag{3.43}$$

[1] There are some methods proposed in the literature that only use one experiment (Ljung, 2003; Smolders and Swevers, 2006). There is nothing magical about them, the same amount of steady-state data is needed in order to get the same signal-to-noise ratio and frequency resolution. The main improvement is that the total measurement time can be slightly reduced since only one transient part is needed, compared to a transient part in each of the experiments in our setup.

and similarly for $\mathbf{Y}(\omega_k)$. Another useful estimator then is the *arithmetic mean (ARI) estimator* (Guillaume, 1998; Pintelon and Schoukens, 2001)

$$\widehat{G}^{\text{ARI}} = \frac{1}{M}\sum_{m=1}^{M}\widehat{G}^{[m]}, \tag{3.44}$$

where

$$\widehat{G}^{[m]} = \mathbf{Y}^{[m]}[\mathbf{U}^{[m]}]^{-1}. \tag{3.45}$$

In case of measurement noise on the input, or closed-loop data, this estimator gives less bias than the H_1 estimator (Pintelon and Schoukens, 2001). A number of different non-parametric multivariable FRF estimation methods are described and analyzed in Paper B and Paper C. See also Section 3.8 for some notes on FRF estimation of nonlinear systems.

3.4.3 Parametric Frequency-Domain Methods

A parametric frequency-domain identification method estimates the unknown parameters θ from *frequency-domain data* Z^N which, in most cases, are obtained by a DFT from the raw time-domain data like

$$Z^N = \{U(\omega_1), Y(\omega_1), \ldots, U(\omega_N), Y(\omega_N)\}. \tag{3.46}$$

The data set could also be obtained directly from the system using a measurement device providing frequency-domain data.

If we assume that the system can be described by (3.22) or (3.23), we have the following (approximate) relations

$$Y(\omega_k) = G(j\omega_k, \theta)U(\omega_k) + H(j\omega_k, \theta)E(\omega_k), \tag{3.47}$$

in continuous time and

$$Y(\omega_k) = G(e^{j\omega_k T_s}, \theta)U(\omega_k) + H(e^{j\omega_k T_s}, \theta)E(\omega_k), \tag{3.48}$$

in discrete time. For non-periodic signals, the relations are only approximate since a transient term then should be added as well. For (3.47) to hold, we have also assumed that the bandwidth of the original continuous-time signals is less than half the sampling frequency such that the continuous-time and discrete-time Fourier transforms, (3.17) and (3.18), coincide.

A common criterion for the parameter estimate is the *(weighted) least squares* criterion

$$V_N(\theta, Z^N) = \sum_{k=1}^{N} \left|Y(\omega_k) - G(j\omega_k, \theta)U(\omega_k)\right|^2 W_k, \tag{3.49}$$

for continuous time and

$$V_N(\theta, Z^N) = \sum_{k=1}^{N} \left|Y(\omega_k) - G(e^{j\omega_k T_s}, \theta)U(\omega_k)\right|^2 W_k, \tag{3.50}$$

for discrete time, where the *weight functions* W_k can be selected using the noise model $H(\cdot, \theta)$. If the noise model depends on θ, an additional term should be added as well (Ljung, 1999; Pintelon and Schoukens, 2001). The parameters enter nonlinearly in this criterion and the minimization therefore requires a numerical search procedure (problems with local minima). The criterion can be made quadratic in θ by parameterizing the model $G(\cdot, \theta)$ as a rational function and multiply $Y(\omega_k) - G(\cdot, \theta)U(\omega_k)$ by the denominator. The estimated parameters will then, in general, be biased. For details, see Pintelon and Schoukens (2001, Chap. 7.8).

For systems with a large dynamic range, which is the case for the highly resonant industrial robot we are studying, the *(weighted) logarithmic least squares* criterion (Pintelon and Schoukens, 2001, pp. 206–207) has been suggested as an alternative to (3.50). $V_N(\theta, Z^N)$ is then given by

$$V_N(\theta, Z^N) = \sum_{k=1}^{N} \left| \log \widehat{G}(e^{j\omega_k T_s}) - \log G(e^{j\omega_k T_s}, \theta) \right|^2 W_k, \tag{3.51}$$

where $\widehat{G}(e^{j\omega_k T_s})$ is the ETFE (3.36) and $\log G = \log |G| + j \arg G$. This criterion has improved numerical stability as well as robustness with respect to outliers. From a theoretical point of view, this will give inconsistent estimates ($\lim_{N\to\infty} \hat{\theta}_N \neq \theta_0$), but for systems with a fairly good signal-to-noise ratio, this is of minor practical importance. This type of criterion is used in Paper A.

3.4.4 Time-Domain versus Frequency-Domain Methods

According to Ljung (2004), most people in the automatic control community basically work with data in the time domain and primarily estimate parametric time-domain models. In some other communities (e.g., "instrumentation and measurement"), frequency-domain data, periodic inputs, and so on, are commonly used to obtain both parametric models and refined nonparametric FRF estimates. In Schoukens et al. (2004), the authors explain that the frequency-domain approach got a bad reputation in the control community due to leakage errors when transforming data from time domain to frequency domain. Even noiseless data (using a random input signal) resulted in large errors in nonparametric FRF estimates.

Quite recently, the true duality between time- and frequency-domain methods have become clear (Ljung, 2004). First, one can note that there is a one-to-one relationship between time-domain and frequency-domain data, so no information is lost by taking the DFT of the time-domain data. The only difference is that some information is more easily accessible in one domain than in the other. It is also possible to reformulate a (linear) identification problem from time domain to frequency domain, and vice versa, so in that sense, they are equivalent (see, for example, Ljung, 1999). The key thing in order to get equivalent descriptions is to properly handle the transients in both domains (Schoukens et al., 2004). However, the complexity of the identification methods will vary depending on the selected domain and there can also be some numerical differences.

For time-domain methods, discrete-time models are the natural choice. For frequency-domain methods, on the other hand, the choice of models is more general. This is due to the fact that the differential or difference equations are replaced by algebraic equations in

the related frequency variable, and then there are no major differences in the calculation of the model estimates.

For identification of a nonlinear model, time-domain methods are usually the preferred choice (Ljung, 1999), even though interesting insights can be obtained also by considering frequency-domain methods (Bai, 2003a,b; Crama and Schoukens, 2001). A drawback with frequency-domain methods is that even a fairly simple nonlinear time-domain model becomes a complicated model in the frequency domain due to the mixing of frequencies ($Y(\omega_k)$ depends on, in principle, all $U(\omega_l),\ l = 1, \ldots, N$). By using particular input signals in the experiments, e.g., a single sinusoid as in Bai (2003a,b), the mixing can be avoided and the output Fourier coefficients contain information such that the nonlinearity can be estimated.

For a more detailed discussion on similarities and differences using time-domain or frequency-domain identification methods, see for example Pintelon and Schoukens (2001, pp. 368–373), Ljung (1999, pp. 227–233), and the included references in this section.

3.5 Closed-Loop Identification

It is sometimes necessary to perform identification experiments under the presence of feedback, i.e., in closed loop. The reason may be an unstable plant, or that the plant has to be controlled for production, economic, or safety reasons. Consider a closed-loop system

$$y(t) = G_0(q)u(t) + v(t), \tag{3.52}$$

$$u(t) = r(t) - F_y(q)y(t), \tag{3.53}$$

where $F_y(q)$ denotes the controller and $G_0(q)$ is the true system. In this setting, typically the reference signal $r(t)$ is considered as the excitation signal. Depending on whether $r(t)$ is known or not, different identification methods can be used.

A problem with closed-loop data is that the data set typically contains less information about the open-loop system. A more serious problem is that many of the identification methods that work well in open loop, will give erroneous results when applied directly to closed-loop data. The reason is that there will be a nonzero correlation between the input $u(t)$ and the noise $v(t)$ in the measured output $y(t)$. This is, for example, the case for the nonparametric methods mentioned in Section 3.4.2, even though there are methods that can handle closed-loop data, see Paper B. Prediction error methods work fine as long as the true system can be described by the model and the data is informative enough.

A common classification of closed-loop identification methods is:

Direct methods: Standard identification methods are directly applied to the measured input-output data without any assumptions on how the data was generated.

Indirect methods: The controller is assumed to be known and the idea is to first estimate the closed-loop system from the reference signal to the output signal. Using the known controller, the open-loop system can then be calculated.

Joint input-output methods: The controller and the open-loop system are estimated by considering the input and output jointly as outputs from an augmented system, driven by the reference signal. The controller must then be parameterized as well.

A drawback with the last two approaches is that the controller must be linear in order to use linear models for the estimation of the closed-loop system and the augmented system, respectively.

As previously mentioned, the direct approach gives problems for many identification methods, with prediction error methods as one exception. The estimates will typically get biased and the bias depends on the properties of the feedback as well as the signal-to-noise ratio. Bias and variance expressions will be given in Section 3.6 both for open-loop and closed-loop data. The indirect and joint input-output approaches convert the closed-loop problem to an open-loop problem, where the reference signal plays the role of an input. Since the reference signal is uncorrelated with the output noise, these approaches can be used together with all the open-loop methods. This is also used in some of the nonparametric methods treated in Paper B.

For a detailed treatment of the subject, the interested reader is referred to, for example, Ljung (1999), Forssell (1999), and Forssell and Ljung (1999).

3.6 Bias and Variance

The model quality is often measured by the statistical concepts *bias* and *variance*. The model bias, $\mathrm{E}(\widehat{G}) - G_0$, is the (average) difference between the true system and the estimated model, while the variance, $\mathrm{E}(|\widehat{G} - \mathrm{E}(\widehat{G})|^2)$, is a measure of how much the estimate will vary between different measurements. The bias term is mostly affected by the model set. Small bias is obtained for large, flexible model sets, and/or for model sets that are well-adapted to the given system.

Consider a true system given by

$$\mathcal{S}: \quad y(t) = G_0(q)u(t) + v(t) = G_0(q)u(t) + H_0(q)e(t), \tag{3.54}$$

where $u(t)$, for an open-loop system, is considered the excitation signal. For a closed-loop system, the input is selected like

$$u(t) = r(t) - F_y(q)y(t), \tag{3.55}$$

where $F_y(q)$ is the controller and $r(t)$ is the reference signal. Note that setting $F_y = 0$ gives us the open-loop case where $u(t) = r(t)$. The input spectrum $\Phi_u(\omega)$ can be written as

$$\Phi_u(\omega) = \underbrace{|S_0(e^{j\omega T_s})|^2\Phi_r(\omega)}_{\Phi_u^r(\omega)} + \underbrace{|F_y(e^{j\omega T_s})|^2|S_0(e^{j\omega T_s})|^2\Phi_v(\omega)}_{\Phi_u^e(\omega)}, \tag{3.56}$$

where $\Phi_v(\omega)$ is the noise spectrum, $\Phi_r(\omega)$ is the spectrum of the reference signal and $S_0(q) = (1 + F_y(q)G_0(q))^{-1}$ is the sensitivity function. $\Phi_u^r(\omega)$ is the part of the input spectrum originating from the reference signal and $\Phi_u^e(\omega)$ is the part originating from the output noise (due to output feedback). For open-loop data, $\Phi_u^e(\omega) = 0$.

It is outside the scope of this chapter to present bias and variance expressions for all methods presented in Section 3.4. However, to get a feeling for the problem, one prediction error method and the nonparametric ETFE will be considered.

3.6.1 Parametric Methods

Assume now that we want to identify a model of (3.54) using a prediction error method with the model structure

$$\mathcal{M}: \quad y(t) = G(q,\theta)u(t) + H(q,\theta)e(t). \tag{3.57}$$

True System in the Model Set ($\mathcal{S} \in \mathcal{M}$)

If we assume that the true system is in the model set ($\mathcal{S} \in \mathcal{M}$), i.e., there exists a parameter θ_0 such that $G_0(q) = G(q,\theta_0)$ and $H_0(q) = H(q,\theta_0)$. Then it is possible to show that the estimate $\hat{\theta}_N$ from (3.32) has the following asymptotic ($N \to \infty$) properties (Ljung, 1999, Chap. 9)

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \to \mathcal{N}(0, P_\theta), \tag{3.58a}$$

$$P_\theta = \lambda_0 \left[\mathrm{E}(\psi(t,\theta_0)\psi^T(t,\theta_0))\right]^{-1}, \tag{3.58b}$$

$$\psi(t,\theta) = \frac{\partial \hat{y}(t|\theta)}{\partial \theta}, \tag{3.58c}$$

$$\lambda_0 = \mathrm{E}(e^2(t)). \tag{3.58d}$$

Here $\mathcal{N}$ denotes the Normal distribution. This means that when the true system is in the model set, $\hat{\theta}_N$ will converge to the true parameter θ_0 and the covariance of the estimation error decays as $1/N$. These properties hold both for open-loop and closed-loop data.

True System Not in the Model Set ($\mathcal{S} \notin \mathcal{M}$)

In case the true system cannot be described by the model set ($\mathcal{S} \notin \mathcal{M}$), the estimated model obviously will deviate from the true system. The properties of the resulting models have been extensively studied in the system identification community. One example is the model bias when using a fixed noise model, $H(q,\theta) = H_*(q)$. The limiting model G_* ($N \to \infty$) is then given by (Ljung, 1999, Eq. (13.53))

$$G_* = \arg\min_G \int_{-\pi/T_s}^{\pi/T_s} \left|G_0(e^{j\omega T_s}) + B(e^{j\omega T_s}) - G(e^{j\omega T_s},\theta)\right|^2 \frac{\Phi_u(\omega)}{|H_*(e^{j\omega T_s})|^2} d\omega,$$

$$|B(e^{j\omega T_s})|^2 = \frac{\lambda_0}{\Phi_u(\omega)} \cdot \frac{\Phi_u^e(\omega)}{\Phi_u(\omega)} \cdot |H_0(e^{j\omega T_s}) - H_*(e^{j\omega T_s})|^2,$$

where $\lambda_0 = \mathrm{E}(e^2(t))$ is the noise variance. The bias term B will then be small in frequency ranges where either (or all) of the following holds:

- the noise model is good ($H_0 - H_*$ is small),
- the feedback contribution to the input spectrum (Φ_u^e/Φ_u) is small, and
- the signal-to-noise ratio is good (λ_0/Φ_u is small).

Note that the model bias, $G_* - G_0$, can be larger than B. The expression only says that the best G_* that can be achieved is $G_* = G_0 + B$, given that the model $G(q, \theta)$ is flexible enough to describe $G_0 + B$.

The variance of the estimated transfer function $\widehat{G}_N$ can, e.g., be obtained from asymptotic black-box theory (orders of both G and H tend to infinity, as well as N) like

$$\mathrm{Cov}(\widehat{G}_N) = \frac{n}{N}\frac{\Phi_v(\omega)}{\Phi_u^r(\omega)}, \tag{3.59}$$

(see Ljung, 1999, Eq. (13.55)). The noise-to-signal ratio determines how well the transfer function is estimated. The variance is also increased if more parameters n are used, which highlights the trade-off between bias and variance (the bias is typically reduced if more parameters are used). Note that it is only the part of the input spectrum originating from the reference signal that will reduce the variance.

Prediction error methods will give consistent estimates, even for closed-loop data, given that the true system can be described in the model set and that the data is informative enough. In particular, it is important to have a flexible noise model to reduce bias effects.

3.6.2 Nonparametric Methods

As an example of a nonparametric method, we will consider the ETFE. For open-loop data, we have (Ljung, 1999)

$$\mathrm{E}(\widehat{G}_N(e^{j\omega T_s})) = G_0(e^{j\omega T_s}) + \frac{\rho_1(N)}{U_N(\omega)}, \tag{3.60}$$

$$\mathrm{Cov}(\widehat{G}_N(e^{j\omega T_s})) = \frac{1}{|U_N(\omega)|^2}(\Phi_v(\omega) + \rho_2(N)), \tag{3.61}$$

where $|\rho_1| \leq C_1/\sqrt{N}$ and $|\rho_2| \leq C_2/N$ for some constants C_1 and C_2. (The dependence of the number of samples is here explicitly shown by adding the index N in the DFTs and the ETFE.) The properties of the ETFE highly depend on whether the data is periodic or not. If the data is periodic, $C_1 = 0$ and $|U_N(\omega)|^2$ increases like const $\cdot\, N$ for the excited frequencies. This gives an unbiased ETFE and the variance decays like $1/N$. If the data is non-periodic, the ETFE will be asymptotically unbiased and the variance will be equal to the noise-to-signal ratio. In that case, the ETFE is a very crude estimate. This is the reason why only periodic excitation signals are used in this thesis.

Remark 3.1. For a nonparametric method, the number of "parameters" usually equals the number of data points (magnitude and phase of $N/2$ frequencies), i.e., $n = N$. By noting that $\Phi_u(\omega) = |U_N(\omega)|^2$, (3.59) and (3.61) coincide (asymptotically, $\rho_2(N)$ disappears). Using periodic data, with N_P samples in each period, can be seen as a reduction of the number of "parameters" to $n = N_P$. The asymptotic covariance is then proportional to $n/N = 1/P$, where P is the number of periods that is used.

Nonparametric estimation methods usually work under the assumption that the input $u(t)$ is uncorrelated with the output noise $v(t)$. This is no longer true for closed-loop data and therefore the estimates usually are biased. The estimates using ETFE tend to (Ljung,

1999, p. 191)

$$\widehat{G}(e^{j\omega T_s}) = \frac{G_0(e^{j\omega T_s})\Phi_r(\omega) - F_y(e^{-j\omega T_s})\Phi_v(\omega)}{\Phi_r(\omega) + |F_y(e^{j\omega T_s})|^2\Phi_v(\omega)}, \tag{3.62}$$

as the number of data goes to infinity and where Φ_r and Φ_v denote the spectrum of $r(t)$ and $v(t)$. If there is no noise in the system, the estimate will tend to the true system. For systems with a large signal-to-noise ratio, the bias can also be negligible. If, on the other hand, the noise dominates over the reference signal, the estimate will tend to the inverse of the controller.

Bias and variance for the multivariable FRF estimation methods are studied in Paper B. See also Paper C and Paper D for experimental results.

3.7 Experiment Design

Most of the results in this thesis are based on experimental data. The choice of input signals has a substantial influence on the observed data, and therefore also on the estimated models. This choice will determine the operating point of the system and which parts and modes of the system that will be excited.

Experiment design has been treated in the literature for almost half a century in both statistics and engineering areas, where some basic references are Fedorov (1972) and Goodwin and Payne (1977). A recent survey can also be found in Gevers (2005). The experiment design problem aims at finding the optimal experiment conditions, χ, contained in some set $\boldsymbol{\chi}$, that minimize a criterion related to the expected outcome of the experiment under certain constraints. The experimental conditions in $\boldsymbol{\chi}$ can be described in terms of available signals to measure and manipulate, the sampling period, amplitudes, frequencies, and so on. Constraints are often related to the total measurement time and the power or amplitude of the excitation signals and/or some other variables. The expected outcome can, e.g., be to minimize the uncertainty in the estimated parameters or the estimated FRF in a certain frequency interval.

There are two different aspects often associated with the choice of excitation signal; the second-order properties, like its spectrum $\Phi(\omega)$, and the "shape" of the signal. For a linear system, asymptotically only the spectrum will have effect. For a limited number of data, and especially for a nonlinear system, the signal shape will have influence as well.

The selection of spectrum and excitation signal will be treated in Sections 3.7.1 and 3.7.2, respectively. Section 3.7.3 gives some comments on experiment design for nonlinear gray-box models, which is also treated in Paper E. Finally, Section 3.7.4 deals with transients.

3.7.1 Selection of Power Spectrum

When selecting an excitation signal, the first thing to make sure is that the estimated model will be acceptable in terms of bias. When the bias may be significant it is wise to let the experiment resemble the situation under which the model is going to be used. Once this is achieved, the signal spectrum can be further optimized with respect to variance.

A problem when talking about optimal excitation signals (with respect to bias and variance in the estimates) is that this often is in conflict with another important property, namely *validation power*. For an nth order SISO system (numerator and denominator polynomials of order n), the optimal signal might be a sum of n sinusoids, but this signal will not be able to reveal if the true system is of a higher order. The signal should therefore be rich enough to be able to validate (and invalidate) the estimated models.

Nonparametric Frequency-Domain Methods

In a nonparametric FRF estimate, there is (asymptotically) no relation between the estimates at different frequencies and the excitation can then be designed to achieve a predefined accuracy in the frequency band of interest (for example maximizing the absolute or relative accuracy). From (3.61) we get that the spectrum

$$\Phi_u(\omega) = \text{const} \cdot \Phi_v(\omega), \tag{3.63}$$

will give about the same absolute accuracy for all frequencies in the estimate. To instead get the same relative accuracy,

$$\Phi_u(\omega) = \text{const} \cdot \frac{\Phi_v(\omega)}{|G_0(e^{j\omega T_s})|^2}, \tag{3.64}$$

could be used. This corresponds to an output spectrum approximately proportional to the noise spectrum.

Parametric Methods

For the parametric case, loosely speaking, an optimal power spectrum is such that the available power is used at the frequencies where it contributes most to the knowledge of the system. Using the prediction error method (3.32) gives the asymptotic covariance matrix (3.58b), which shows that interesting parameters must have a clear effect on the output predictions. For the open-loop case, we also have the expression for the *average information matrix per sample* $\bar{M}(\Phi_u)$ (Ljung, 1999, pp. 416–417)

$$\bar{M}(\Phi_u) = \int_{-\pi}^{\pi} \tilde{M}(\omega)\Phi_u(\omega)d\omega + M_e, \tag{3.65}$$

$$\tilde{M}(\omega) = \frac{\lambda_0 G_\theta'(e^{j\omega T_s}, \theta_0)[G_\theta'(e^{j\omega T_s}, \theta_0)]^T}{2\pi\kappa_0\Phi_v(\omega)},$$

Here G_θ' is the gradient of G with respect to the parameters, κ_0 depends on the selected norm in the criterion, and M_e is a term independent of Φ_u. The information matrix should be large in order to get a small covariance matrix. To achieve this, the input power should be spent at frequencies where $\tilde{M}(\omega)$ is large, i.e., where the Bode plot is sensitive to parameter variations (G_θ' large). Often a sufficient guidance on where to spend the input power is found by varying the parameters in the model and checking how the Bode plot changes.

The classical approach for optimal experiment design is to minimize some scalar measure of the (asymptotic) covariance matrix $P_\theta(\chi, \theta_0)$ in (3.58b) with constraints on the input and/or the output power. Some commonly used measures are A-optimality $\mathrm{Tr}\, P_\theta(\chi, \theta_0)$, D-optimality $\det P_\theta(\chi, \theta_0)$, E-optimality $\lambda_{max}\left(P_\theta(\chi, \theta_0)\right)$, and L-optimality $\mathrm{Tr}\left(W P_\theta(\chi, \theta_0)\right)$.

An inherent problem in most experiment design problems is that the covariance matrix depends on the true system parameters θ_0. These parameters are, at least partly, unknown or uncertain, otherwise there would be no need for an experiment. This has been handled in different ways in the literature. One could either assume a good prior estimate of θ_0 and hope for the best, or use some robust experiment design methods. Various strategies have been suggested, including sequential design (iterate between parameter estimation and experiment design), Bayesian design (minimize the expected value over the prior parameter distribution) and min-max design (minimize the worst case when the parameters are contained in a given set). For references, see Goodwin et al. (2006).

In addition to the problem of knowing θ_0, experiment design problems are, in their original form, often intractable due to non-convex and infinite-dimensional constraints, as well as the problem of finding a signal realization which has the desired spectral properties. However, due to great advances in the optimization community, there exist today many useful methods to reformulate the experiment design problems into tractable convex optimization problems. Here, Jansson and Hjalmarsson (2005) is a good example, where a framework for this reformulation is presented, based on linear parameterizations of the signal spectrum. Constraints on these spectra can be included, as well as quality constraints and robustness constraints as long as they can be rewritten as convex functions of the inverse (asymptotic) covariance matrix.

3.7.2 Selection of Excitation Signal

Given an (optimal) input spectrum $\Phi_u(\omega)$ and limitations on the input amplitude and other constraints, we must now select a particular excitation signal that fulfills all the requirements.

One way of achieving a certain spectrum could be to use a stepped sine excitation, consisting of a series of single sine measurements at the specified frequencies. Advanced digital signal processing algorithms, especially the Fast Fourier Transform (FFT), have lead to the use of more complex input signals. Instead of exciting the system frequency by frequency, broadband spectrum signals are generated. This gives considerable reduction of measurement time. Many different kinds of signals exist, e.g., (filtered) random noise, pseudo-random binary sequence (PRBS), swept sine (chirp), and sum of sinusoids (multisine) (see Pintelon and Schoukens, 2001, for details).

In this thesis, we will exclusively use periodic signals (with N_P samples in each period) and, in particular, the *random phase multisine* signal, which can be written as

$$u_{ms}(t) = \sum_{k=1}^{N_f} A_k \cos(\omega_k t + \phi_k), \tag{3.66}$$

with amplitudes A_k, frequencies ω_k chosen from the grid $\{\frac{2\pi l}{N_P T_s},\ l = 1, \ldots, N_P/2 - 1\}$ (assuming N_P even), and random phases ϕ_k uniformly distributed on the interval $[0, 2\pi)$.

The signal is preferably calculated using FFT. A simple MATLAB implementation is given below, together with an example.

Code 3.1 (MATLAB code for multisine signal** (3.66)**)

```
function u = multisine(N,Np,k,Ak)
% Error check
k = k(:); Ak = Ak(:); % column vectors
if length(Ak) == 1,  % allow scalar amplitude
  Ak = Ak*ones(size(k));
end
if length(k)~=length(Ak),
  error('Ak should be scalar or have same size as k');
end
if rem(Np,2), error('Np should be even'); end
if any(~ismember(k,1:Np/2-1)),
  error('k should be in set 1:Np/2-1');
end
% One period
U = zeros(Np,1);
U(1+k) = Np*Ak.*exp(2*pi*i*rand(size(Ak)));
u = real(ifft(U));
% Make periodic
u = kron(ones(ceil(N/Np),1),u);
u = u(end-N+1:end);
end
```

Example 3.1: Odd random phase multisine signal

Assume that we want to generate 3 periods of a multisine signal with sampling period $T_s = 0.1\,\text{s}$ and $N_P = 100$ samples in each period. This gives us a period of $T_0 = N_P T_s = 10\,\text{s}$ and therefore a frequency resolution of $1/T_0 = 0.1\,\text{Hz}$. The available frequency grid (in Hz) is $\{0.1, 0.2, \ldots, 4.9\}$. In this example, we select the odd frequencies in the interval $[0.1\ 1]\,\text{Hz}$ and use $A_k = A = 1$ for those frequencies. The MATLAB code for this is given below.

```
Ts = 0.1; Np = 100; N = 3*Np; T0 = Ts*Np;
do_odd = 1; % all (0), odd (1)
fr = [0.1 1]; % frequency interval in Hz
% available frequencies:
k = (1:(1+do_odd):Np/2-1)'; f = k/T0;
% select frequencies:
k = k(find(fr(1)<=f & f<=fr(2))); f = k/T0;
Ak = ones(size(k)); % amplitude distribution
u = multisine(N,Np,k,Ak); % create multisine
% plot amplitudes:
U = abs(fft(u(1:Np))*2/Np);
```

```
ka = (1:Np/2-1)'; fa = ka/T0;
figure;
stem(fa,U(1+ka),'o'); hold on; stem(f,U(1+k),'*');
xlabel('Frequency (Hz)'); ylabel('Magnitude');
% plot signal:
figure;
plot(Ts*(0:N-1)',u);
xlabel('Time (s)'); ylabel('Signal');
```

Executing this code gives the result shown in Figure 3.2.

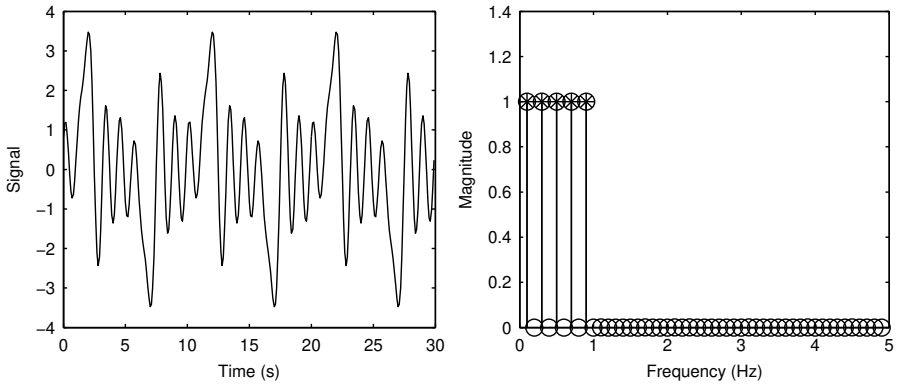


Figure 3.2: *Multisine signal for Example 3.1.*

The multisine signal is very flexible since the user can specify exactly which frequencies to use (from the available frequency grid due to a period signal). Given a desired power spectrum $\Phi_u(\omega)$, the amplitudes are set as $A_k = 2\sqrt{\Phi_u(\omega_k)/N_P}$. The random phases can also be optimized in order to minimize the *crest factor*, i.e., the ratio of the peak value and the root mean square value (see Pintelon and Schoukens, 2001). For a nonlinear system, this should however be done with some care since crest factor minimization gives an almost binary amplitude distribution of the signal $u_{ms}(t)$. Random phase multisine signals lead to a Normal behavior which is a better general assumption about the situation under which the model is going to be used. (Recall the first paragraph in Section 3.7.1.) This is also illustrated in the following example.

Example 3.2: Crest factor optimization

Consider a multisine signal $u_{ms}(t)$ with $N_P = 10^4$ samples and $A_k = A$ for all $N_P/2-1$ possible frequencies. This gives (for one particular realization of the phases) a crest factor of 4.02. The phases are then optimized by using $10\,000$ iterations of the algorithm in Van der Ouderaa et al. (1988), giving a crest factor of 1.12. The resulting amplitude distributions can be seen in Figure 3.3.

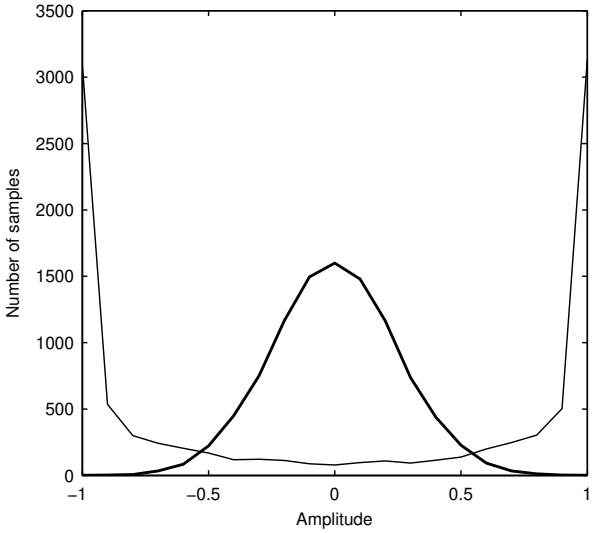


Figure 3.3: *Amplitude distribution of original multisine signal u_{ms} (thick line) and crest factor optimized signal (thin line). The amplitudes are scaled such that $\max_t |u_{ms}(t)| = 1$ for both signals.*

Periodic signals have certain advantages which can be used to avoid leakage in the DFT, to improve the signal-to-noise ratio by averaging over periods, and to estimate the noise level by comparing the output between different periods. An example of this is given in Paper D, where the random phase multisine signal is used to estimate the uncertainty in nonparametric FRF estimates, both due to measurement noise and nonlinearities.

For multivariable nonparametric FRF estimation, the *orthogonal random phase multisine* signal is suggested in Dobrowiecki and Schoukens (2007), given certain amplitude constraints for the input signal. The blocks in (3.43) are then given by

$$\mathbf{U}^{[m]}(\omega_k) = \mathbf{U}^{[m]}_{\text{diag}}(\omega_k)\mathbf{T}, \quad m = 1, \ldots, M, \tag{3.67}$$

with

$$\mathbf{U}^{[m]}_{\text{diag}}(\omega_k) = \operatorname{diag}\left\{U_l^{[m]}(\omega_k)\right\}_{l=1}^{n_u},$$

a diagonal matrix where each $U_l^{[m]}(\omega_k)$ is a scalar random phase multisine signal, and $\mathbf{T}$ an orthogonal matrix

$$\mathbf{T}_{il} = e^{\frac{2\pi j}{n_u}(i-1)(l-1)}.$$

The measurement procedure is thus to generate n_u random phase multisine signals that are used in the first experiment in a block and then shift them orthogonally in the following $n_u - 1$ experiments according to $\mathbf{U}^{[m]}_{\text{diag}}\mathbf{T}$. This procedure is then repeated for each of the M blocks. The orthogonal multisine signal guarantees that each $\mathbf{U}^{[m]}$ is invertible, compared to the random fluctuations that occur when using a separate random phase multisine in each element of $\mathbf{U}^{[m]}$ (see Dobrowiecki and Schoukens, 2007, for examples).

3.7.3 Nonlinear Gray-Box Models

Consider the nonlinear gray-box model (3.28) and assume some parameter estimator, e.g., (3.32) with predictor (3.35). The general problem of designing the optimal excitation signal that minimizes the parameter covariance matrix, given constraints on the excitation signal as well as some states, is a nonlinear optimal control problem which is extremely hard to solve.

In Paper E, a different approach is used where we assume that the excitation is a small perturbation around a certain operating point (x_0, u_0), called position in the paper. The perturbation is assumed to be small enough to justify linearizion of the nonlinear system (3.28) and the use of linear theory. For a nonlinear system, the information about the unknown parameters will vary depending on the selected position. Therefore, given a limited total measurement time, one should perform experiments in the position(s) that contribute the most to the information about the unknown parameters. From a set of candidate positions, the selection of the optimal position(s) is solved as a convex optimization problem. For details, see Paper E.

3.7.4 Dealing with Transients

When applying an input signal to a system, the system response will contain transient effects as well as the steady-state response. For parametric methods, usually transient effects are handled by estimating the initial conditions of the system using additional parameters (Ljung, 1999; Pintelon and Schoukens, 2001).

The nonparametric methods mentioned in Section 3.4.2, assume that (3.39) holds in order to avoid errors. This will hold as long as the difference between the initial and final conditions of the system is zero (Pintelon and Schoukens, 2001, Sec. 5.3.2.2). This is, for example, the case with periodic excitation in steady state (no difference between the output signal from one period to the next). The same is true for burst excitation if the whole system response is collected.

The waiting time to reach steady state mainly depends on the damping of the system. In a real situation, it is enough to wait until the transient effects are less than other sources of errors, such as measurement noise. See Schoukens et al. (2000) for estimates of the required waiting time based on the dominating time constant of the system.

3.8 Nonparametric FRF Estimation of Nonlinear Systems

The nonparametric FRF estimation in Section 3.4.2 inherently gives a linear nonparametric model. For a nonlinear system, we are looking for a linear approximation of the input-output behavior. The properties of such approximations have been extensively studied in, e.g., Schoukens et al. (1998), Schoukens et al. (2005), and Dobrowiecki and Schoukens (2007).

3.8.1 Analysis Using Volterra Series

For the analysis of the estimated FRFs, the random phase multisine signal (3.66) is rewritten by using $\cos x = \frac{1}{2}(e^{jx} + e^{-jx})$ as

$$u(t) = \sum_{\substack{k=-N_f \\ k\neq 0}}^{N_f} U_k e^{j(\omega_k t + \phi_k)}, \tag{3.68}$$

where $U_k = U_{-k} = A_k/2$ and $\phi_k = -\phi_{-k}$. This allows us to use both positive and negative frequencies. For the analysis *Volterra series* are used to model the nonlinear systems. In the frequency domain, a SISO Volterra series for periodic inputs can be described as (Schoukens et al., 1998)

$$Y(\omega_k) = \sum_{\alpha=1}^{\infty} \left[\sum_{\substack{k_1=-N_f \\ k_1\neq 0}}^{N_f} \cdots \sum_{\substack{k_{\alpha-1}=-N_f \\ k_{\alpha-1}\neq 0}}^{N_f} G_\alpha(\omega_{k_1}, \ldots, \omega_{k_\alpha}) \prod_{i=1}^{\alpha} U(\omega_{k_i}) \right], \tag{3.69}$$

with $\omega_{k_\alpha} = \omega_k - \sum_{i=1}^{\alpha-1} \omega_{k_i}$ and $G_\alpha(\omega_{k_1}, \ldots, \omega_{k_\alpha})$ the symmetrized frequency-domain Volterra kernel of order α.

Remark 3.2. In this section, as well as in most of the included papers, we will use a simplified notation $\widehat{G}(\omega_k)$ for the FRF $\widehat{G}(j\omega_k)$ or $\widehat{G}(e^{j\omega_k T_s})$.

The properties of the nonparametric FRF estimate

$$\widehat{G}(\omega_k) = Y(\omega_k)/U(\omega_k), \tag{3.70}$$

with $Y(\omega_k)$ from (3.69) are extensively studied in, e.g., Schoukens et al. (1998) and Schoukens et al. (2005). The fundamental observation is that using a random phase multisine (3.68) as input signal in an open-loop setting, with N_f sufficiently large, the FRF estimate (3.70) can be written as

$$\widehat{G}(\omega_k) = G_R(\omega_k) + G_S(\omega_k) + N_G(\omega_k), \tag{3.71}$$

with G_R the best linear approximation to the nonlinear system[2], G_S a zero mean stochastic nonlinear contribution, and N_G the measurement noise contribution. For different realizations of the random phases in (3.68), the stochastic nonlinear contribution, G_S, acts as circular complex noise. Therefore G_S cannot be distinguished from the measurement noise. Still, for a given realization, G_S is a deterministic component that will look the same for different periods. The best linear approximation, G_R, depends (asymptotically) only on the input spectrum and odd nonlinearities (odd α in (3.69)).

3.8.2 Excitation Signals

For a linear system, $G_S = 0$ and it is sufficient to use only one realization of the random phase multisine and average over multiple periods to reduce N_G. For a nonlinear system,

[2]For the class of Normal input signals (e.g., random phase multisine) with a given power spectrum.

averaging over periods will not reduce G_S. It is then important to use several realizations of the random phase multisine signal, which will enable a reduction of both G_S and N_G. By comparing FRF estimates from different periods and realizations, it also possible to detect and quantify the level of nonlinear distortions in the system. Such procedures are described in Schoukens et al. (2005) and are also used in Paper D for the estimation of nonlinear effects in FRF estimation of industrial robots.

It is usually recommended to use so-called odd random phase multisines, i.e., to only excite the odd harmonics (from the grid $\{\frac{2\pi}{N_P T_s}(2k+1),\ k = 0, 1, \ldots, N_P/4 - 1\}$). The motivation is that in that case, only odd nonlinearities contribute to the variance of G_S (at the odd frequencies). Exciting all harmonics increases the variance of G_S also at the odd frequencies due to even nonlinearities and a combination of odd and even frequencies.

To understand this, consider once more (3.69). To obtain a contribution to the output at frequency ω_k from a nonlinearity of degree α, the sum of α frequency components $\sum_{i=1}^{\alpha} \omega_{k_i}$ should equal ω_k (both positive and negative frequencies are allowed). These contributions can be divided into bias and variance, where expectation is with respect to the phases ϕ_k for different realizations of the signal. If $\mathrm{E}(\sum_{i=1}^{\alpha} \phi_{k_i}) = \phi_k$ (which is only true if ϕ_k is included in the sum and the remaining phases come in pairs $(\phi_l, -\phi_l)$), then the nonlinearity gives a bias contribution, otherwise a variance contribution. This is why only odd nonlinearities give a contribution to the bias.

Using an odd random phase multisine as excitation, then even nonlinearities will only show up at even frequencies and odd nonlinearities at odd frequencies (study $\omega_k = \sum_{i=1}^{\alpha} \omega_{k_i}$ with only odd ω_{k_i}). This is an important insight since the bias term only depends on the odd nonlinearities and the input spectrum.

If even frequencies are added, the variance will increase since there are additional combinations of frequency components that sum up to ω_k. Considering the odd frequencies, even nonlinearities now also contribute to the variance through a combination of odd and even frequencies (e.g., $2 + 2 + 2 + 1 = 7$ for $\alpha = 4$). Using a full random phase multisine signal will therefore force the user to average out (over different realizations) the effect of even nonlinearities in order to obtain the best linear approximation.

The only drawback from using only the odd frequencies is the frequency resolution, but the SNR is on the other hand increased. The frequency resolution can also be changed by simply increasing the period $T_0 = N_P T_s$ of the multisine signal.

For additional properties, as well as a more detailed treatment of linear approximations of nonlinear systems, see the included references in this section as well as, e.g., Enqvist (2005).

3.8.3 Example: Nonlinear Two-Mass Model

The chapter is ended by an example where we study the effects of nonlinear spring stiffness and Coulomb friction for nonparametric FRF estimation of a two-mass system.

Example 3.3: Nonlinear two-mass model

Consider the two-mass model in Figure 2.6. Here, a nonlinear spring stiffness and Coulomb friction will be added to study how the estimated FRF is affected by these nonlinearities.

With the states $x_1 = q_m - q_a$, $x_2 = \dot{q}_m$, and $x_3 = \dot{q}_a$, the dynamics is given by

$$\begin{aligned}
\dot{x}_1 &= x_2 - x_3, \\
\dot{x}_2 &= \frac{1}{J_m}\left(u - \tau_f(x_2) - \tau_s(x_1) - d(x_2 - x_3)\right), \\
\dot{x}_3 &= \frac{1}{J_a}\left(\tau_s(x_1) + d(x_2 - x_3)\right),
\end{aligned}$$

with a continuous friction model (see (5) in Paper F)

$$\tau_f(x_2) = F_c \tanh(\beta x_2),$$

and a piece-wise linear spring stiffness

$$\tau_s(x_1) = \begin{cases} k_1 x_1, & |x_1| \leq b, \\ \operatorname{sign}(x_1)(k_2(|x_1| - b) + k_1 b), & |x_1| > b. \end{cases}$$

The motor acceleration $\dot{x}_2$ is considered as output. Since the purpose of the example is to study nonlinear effects on the FRF estimate, no measurement noise is used in the simulations. Numerical values for the model parameters are: $J_m = 0.008$, $J_a = 0.02$, $d = 0.05$, $F_c = 1$, $\beta = 7$, $k_1 = 10$, $k_2 = 40$, and $b = 0.4$.

As excitation $u(t)$, an odd random phase multisine signal (3.66) is used with period $T_0 = N_P T_s = 10\,\text{s}$ and $A_k = A$ for all odd frequencies in the interval 1–40 Hz. Two cases of amplitudes A are used, such that $u_{\max} = \max_t |u(t)|$ is either 5 or 20. The amplitudes are chosen such that only the linear region of the spring is used for $u_{\max} = 5$, which enables us to see the effect of nonlinear spring stiffness. To study the effects of Coulomb friction, the simulations are also carried out with either $F_c = 0$ or $F_c = 1$.

100 experiments are performed with a new realization of the random phases in each experiment. The resulting FRF estimates can be seen in Figures 3.4 and 3.5. By comparing the FRF estimates for the two amplitudes in Figures 3.4, it is evident that the nonlinear spring introduces nonlinear distortions in the estimate for $u_{\max} = 20$. By averaging over 100 experiments, these distortions are reduced, revealing the linear approximation (right column in Figure 3.4), which (almost) corresponds to a linear two-mass system with an increased spring stiffness.

The effects of Coulomb friction are studied in Figure 3.5. The nonlinear distortions can be seen also here. It is evident that the Coulomb friction affects the FRF both at low and high frequencies. The resonance in the system has disappeared in the linear approximation!

Remark 3.3. In the closed-loop experiments in this thesis, the effects of Coulomb friction are reduced by avoiding zero velocity as much as possible. For details, see Paper C, Paper D, and Paper F.

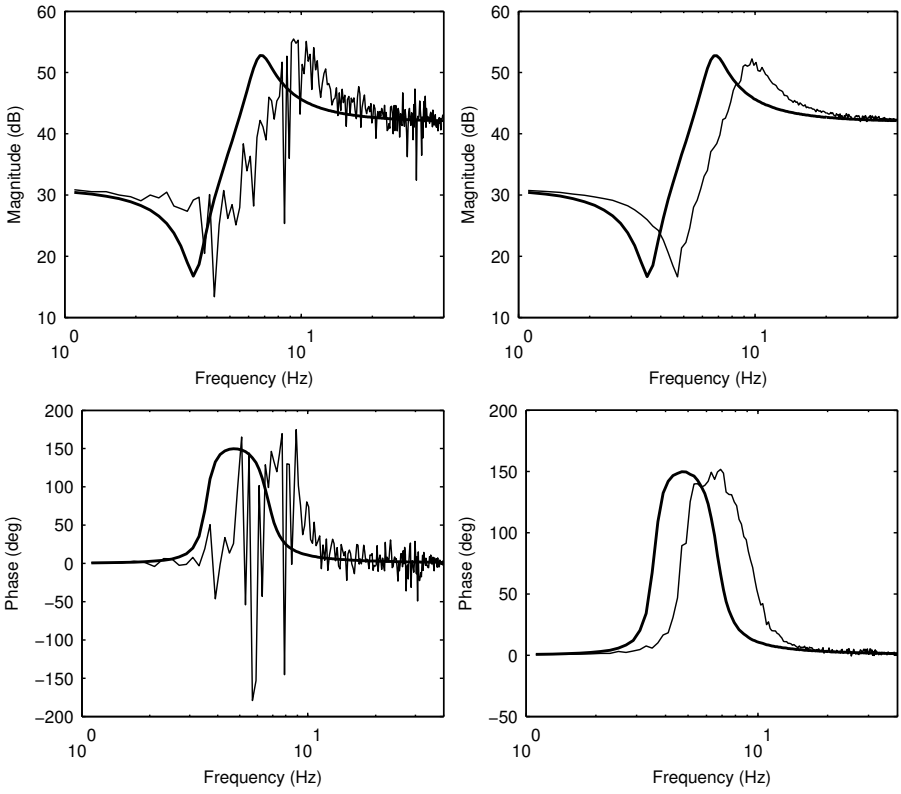


Figure 3.4: *Nonparametric FRF estimates in Example 3.3 with* $F_c = 0$*. Thick line:* $u_{max} = 5$*, thin line:* $u_{max} = 20$*. Left: 1 experiment, right: 100 experiments.*

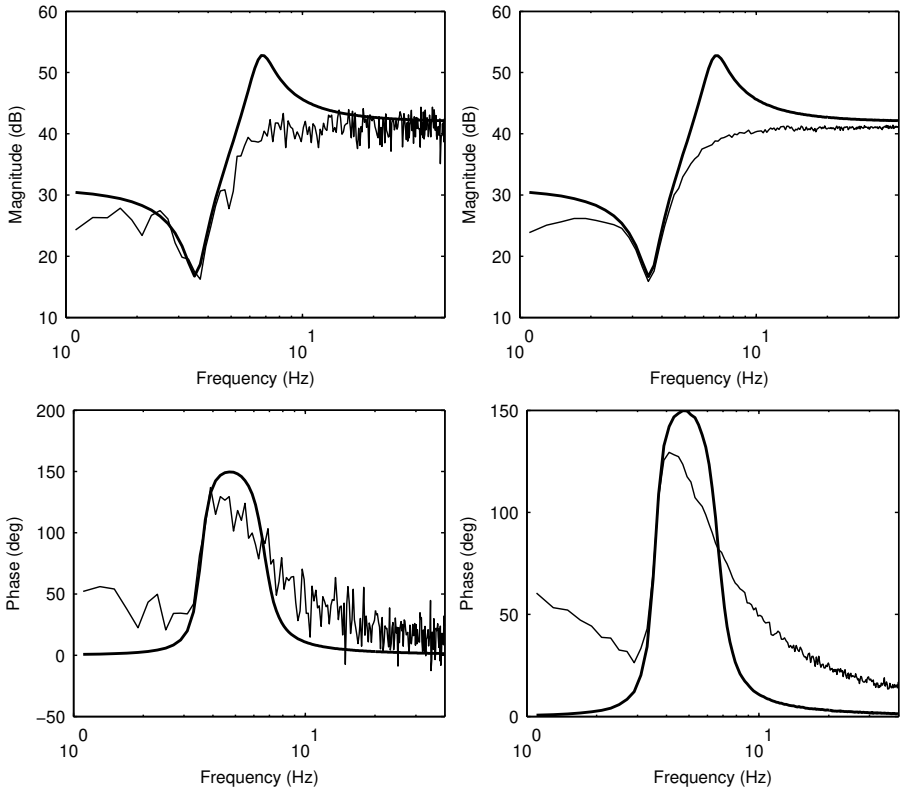


Figure 3.5: *Nonparametric FRF estimates in Example 3.3 with* $u_{max} = 5$. *Thick line:* $F_c = 0$, *thin line:* $F_c = 1$. *Left: 1 experiment, right: 100 experiments.*

4

System Identification in Robotics

There are many available system identification methods and the techniques of system identification have a wide application area. Therefore, one might ask the question: *Why dedicate a chapter for the topic of system identification in robotics*? The answer is that the robot application is associated with many challenging problems for system identification methods. The system is multivariable, nonlinear, unstable, and oscillatory. Having all these topics in the same problem is really challenging so usually, in the literature, at least one of the first three topics is left out. Moreover, the robot dynamics can be modeled quite accurately using known mechanical relations, as was pointed out in Chapter 2. These dynamic equations have particular features that can be exploited by the identification methods.

Since identification in robotics is a much studied problem, the purpose of this chapter is to give an overview of some earlier results related to the identification of kinematics, rigid body dynamics, flexibilities, and nonlinearities. In particular, Section 4.4.4 gives some details on the extended flexible joint model that is used throughout the thesis. The chapter is ended by a short description of EXPDES, a software tool written by the author of this thesis for experiment design, nonparametric multivariable FRF estimation, and identification of unknown parameters in nonlinear gray-box models.

4.1 Introduction

System identification in robotics can be divided into at least three different levels or application areas (see, for example, Albu-Schäffer and Hirzinger, 2001). These levels involve the estimation of the

- kinematic description (including elastostatic effects),
- dynamic model, often divided into

  - rigid body dynamics, and
  - flexible body dynamics, and

- joint model.

Flexible body dynamics here refers to elastic effects in the robot structure, for example in the links. The joint model involves motor inertia, gearbox elasticity and damping, backlash, motor characteristics, and friction parameters. In addition to these three levels, one can also consider the estimation of load dynamics (Kozlowski, 1998, Chap. 6) and adaptive control (Ortega and Spong, 1989) as parts of the system identification problem. Other important areas left out in this chapter are the identification of sensor and actuator disturbances (Jahns and Soong, 1996), identification of thermal models for motors, gears, and their surroundings (Ambrose and Ambrose, 1995; Brogårdh et al., 2006; Graeser and Klingel, 2002), and identification of mechanical stress in arm- and component structures for fatigue analysis (Brogårdh et al., 2001; Wang and Zhao, 2002).

Nominal parameter values for the kinematics and rigid body dynamics can often be obtained from CAD models on the basis of the geometries and materials of the robot arm and components. These values are often not accurate enough, due to simplifications in the modeling and complex dynamic effects, such as joint friction and balancing equipment forces, that cannot be sufficiently modeled. To obtain high accuracy, these parameters must be tuned by the use of experimental data. Some parts of the joint model, like the motor characteristics, can be measured before assembly whereas things like friction depend on how links, bearings, etc. are put together, and therefore must be estimated after assembly. Identification is also important for the verification of the rigid body model.

In a majority of the literature on robot identification (see, for example, Kozlowski (1998) and the references in this chapter), a rigid body model is assumed and flexibilities, if any, are assumed to be located at the joint level. This can partly be explained by a previous use of fairly rigid arm structures and bearing systems. The assumption has also been previously verified, for example by using modal analysis (Behi and Tesar, 1991). With the increasing use of new mechanical structures (bend-over-backwards capability, non-symmetric structure, lightweight robots) and new applications in industry where higher performance is demanded with respect to position accuracy and speed, elastic effects become even more evident and must be dealt with. Therefore, finding procedures for the identification of flexibilities is important.

In Section 4.2, the identification of kinematics will be briefly described. Section 4.3 gives an overview of rigid body identification. Finally, identification of (joint) flexibilities and nonlinearities are covered in Section 4.4.

4.2 Kinematics

The kinematic parameters describe the geometric relation between joint variables and the end effector location (position and orientation). These parameters also define the inverse kinematics (see Section 2.2.2), which is used during the path planning (see Section 2.3.1) to transform a Cartesian path to a corresponding path in joint space. The quality of this mapping will determine the robot's accuracy and it is therefore crucial to have a correct kinematic description.

Small errors in the kinematic model can be partly compensated for by modifying the Cartesian path when doing the motion planning in the robot cell. However, this might be a tedious job. Moreover, in order to fulfill increasing requirements on off-line programming, where the whole motion planning is carried out in software, methods for identification and compensation of kinematic errors must be continuously refined. Today, a volumetric accuracy of ± 0.5 mm can be obtained, even for a big robot (Brogårdh, 2007). Note that the repeatability of a standard industrial robot is much better (e.g., ± 0.1 mm for an ABB IRB6400 robot that handles 200 kg). The repeatability also sets a limit for the accuracy that can be achieved by identification and compensation of kinematic errors.

Identification of the kinematic parameters is also called *kinematic calibration* or *robot calibration* and many different techniques exist. For a survey, see, e.g., Hollerbach (1989). However, the general idea is to consider the forward kinematics (2.3) of the robot and perform a first order Taylor series expansion of the location X with respect to the kinematic parameters θ like

$$\Delta X = \frac{\partial f_{kin}}{\partial \theta}(q_m, \theta_{nom})\Delta\theta, \tag{4.1}$$

where $\Delta X = X - f_{kin}(q_m, \theta_{nom})$ is the deviation in end effector location from the predicted location and $\Delta\theta = \theta - \theta_{nom}$ is the corresponding parameter deviation from the nominal value θ_{nom}. This system of equations can be solved in a least-squares sense for $\Delta\theta$ by considering a number of locations (see corresponding solution below in (4.4) and (4.6)). In addition to the robot link parameters, θ should include additional parameters describing, for example, elastostatic effects (deflection due to gravity) and the mounting of the robot base and the end effector.

Remark 4.1. In the kinematic description (4.1), the motor angular position q_m is used, compared to q_a in (2.3), to highlight that the kinematic description also should include the gearbox. This is important since only q_m is measured in a standard industrial robot, *not* the angular position q_a after the gearbox.

To be able to calculate the deviation ΔX, measurements of the end effector location X are needed. These measurements can be either relative or absolute. Examples of relative measurements could be to approach the same point from different directions or measuring an object with known geometry (but unknown absolute position). Zhong et al. (1996) suggest the use of a trigger probe as end effector to touch constraint planes in the workspace (the location of the planes are not necessarily known exactly). Using relative measurements, many of the kinematic error parameters $\Delta\theta$ can be estimated. For absolute measurements, two main routes can be seen. The first one uses some mechanical apparatus that allows constraining the end effector at given locations with a priori known precision. The second route uses direct measurement systems of the end effector location in the Cartesian space, for example a laser tracking system. This is the kind of system used in Abderrahim and Whittaker (2000), where the Denavit-Hartenberg parameters[1] are estimated by combining measurements of the end effector caused by movements of each joint, one at a time. See also Khalil and Besnard (2002) for a treatment of elastostatic effects due to both joint and link flexibilities.

[1] The Denavit-Hartenberg parameters are sometimes used to describe the kinematics using 4 parameters for each link, see, e.g., Sciavicco and Siciliano (2000) for details.

To obtain accurate estimates of the kinematic parameters, it is important to excite the system as much as possible, which involves measurements for a large number of robot configurations throughout the workspace. To verify the updated kinematic model, a new set of measurements is used (in configurations that were not used for the estimation).

4.3 Rigid Body Dynamics

The identification of rigid body dynamics is based either on the robot dynamic model (2.7), also called the *differential model*[2]

$$H_{rb}(q_a, \dot{q}_a, \ddot{q}_a)\theta_{rb} = \tau_a, \tag{4.2}$$

or the energy (difference) model (2.8), also called the *integral model*

$$\Delta h_{rb}(q_a, \dot{q}_a)\theta_{rb} = \Delta\mathcal{H}_{rb} = \mathcal{H}_{rb}(t_b) - \mathcal{H}_{rb}(t_a) = \int_{t_a}^{t_b} \tau_a^T \dot{q}_a dt, \tag{4.3}$$

where $\theta_{rb} \in \mathbb{R}^{12n}$ is the parameter vector and n is the number of links (assuming 10 inertial parameters and 2 friction parameters for each link). The two models are both linear in the parameters, which are the same in the two models. The integral model (4.3) consists of only one equation, compared to the n equations in the differential model (4.2). The differential model is, in that respect, richer in information than the integral model. An advantage with the integral model is that the joint acceleration $\ddot{q}_a$ is not needed. For easier treatment, only the differential model will be considered from now on. See, for example, Gautier and Khalil (1988) and Kozlowski (1998) for treatments of both the differential and integral models.

4.3.1 Measurements

Traditionally, flexibilities are ignored during the identification of rigid body dynamics, which is motivated by the use of low-frequency trajectories that do not excite flexibilities. That allows us to use the measured motor joint variables, q_m, as estimates of the arm joint variables, q_a, by a simple scaling of the gearbox ratio, $q_a = r_g q_m$. Both motor angular position and velocity can be measured, but it is common to only measure the angular position. If the robot dynamic model (4.2) is used, joint acceleration (and velocity) must be reconstructed. Since the identification is done off-line, the reconstruction is typically done using non-causal filters (no phase shift) and central difference algorithms.

The joint torques τ_a can, in rare cases of torque sensors at the joints, be measured directly. Otherwise they can, for example, be estimated from τ as $\tau_a = \frac{1}{r_g}(\tau - J_m\ddot{q}_m)$ (assuming the motor inertia J_m to be known). The motor torque τ is in turn either estimated from current measurements (in the case of electric actuators) or obtained from the torque references generated by the speed/position controller.

[2]The names *differential model* and *integral model* are according to Kozlowski (1998).

4.3.2 Base Parameters

A problem with the model representation (4.2) is that θ_{rb} is not identifiable, i.e., there are infinitely many parameter values that give the same dynamic model. For a particular robot, a number of the parameters are zero or linearly dependent. To find the minimum number of parameters, called *base parameters*, that characterize the dynamic model, different approaches can be found in the literature. The most basic distinction is if the problem is solved using a numerical or analytical method. Consider now N samples of data, $Z^N = \{q_a(t_i), \dot{q}_a(t_i), \ddot{q}_a(t_i), \tau_a(t_i),\ i = 1, 2, \ldots, N\}$, which are related according to (4.2) as

$$\underbrace{\begin{bmatrix} H_{rb}(q_a(t_1), \dot{q}_a(t_1), \ddot{q}_a(t_1)) \\ H_{rb}(q_a(t_2), \dot{q}_a(t_2), \ddot{q}_a(t_2)) \\ \vdots \\ H_{rb}(q_a(t_N), \dot{q}_a(t_N), \ddot{q}_a(t_N)) \end{bmatrix}}_{\Phi} \theta_{rb} = \underbrace{\begin{bmatrix} \tau_a(t_1) \\ \tau_a(t_2) \\ \vdots \\ \tau_a(t_N) \end{bmatrix}}_{Y}. \tag{4.4}$$

Numerical methods to obtain the base parameters are mainly based on QR or SVD factorizations of the regressor matrix Φ. The set Z^N should then be sufficiently rich so that all parameters are excited. One option is to insert random numbers for $q_a(t_i)$, $\dot{q}_a(t_i)$, and $\ddot{q}_a(t_i)$. See, for example, Gautier (1990) and Sheu and Walker (1989) for details and Pfeiffer and Hölzl (1995) for an example. Many analytical methods have also been proposed, based on symbolical manipulations of the model equations. For closed form solutions, see for example, Mayeda et al. (1990) and Gautier and Khalil (1990).

4.3.3 Estimators

The robot dynamic model (4.2) is replaced by a minimal parameter-linear description

$$H_b(q_a, \dot{q}_a, \ddot{q}_a)\theta_b = \tau_a, \tag{4.5}$$

where the base parameters θ_b are nonlinear functions of θ. Measurements from a large number of time instants are needed to avoid an ill-conditioned regressor matrix Φ_b (similar to (4.4)). To obtain the measurements, the robot is moved along a trajectory and joint motion and torque are measured. Finally the base parameters can be estimated, for example, by a *weighted least squares (WLS)* method like

$$\hat{\theta}_b = \arg\min_{\theta_b} \frac{1}{2}(Y - \Phi_b\theta_b)^T W (Y - \Phi_b\theta_b) = (\Phi_b^T W \Phi_b)^{-1} \Phi_b^T W Y. \tag{4.6}$$

Many different weighting matrices W occur in the literature, see Gautier and Poignet (2001) and Swevers et al. (1997) for examples. Other estimation methods are, for example, treated in Gautier and Poignet (2001), which contains a comparison between WLS and the extended Kalman filter. In Gautier et al. (1994), total least squares estimation is used, and Presse and Gautier (1993) use Bayesian estimation techniques.

4.3.4 Experiment Design

To make the parameter estimation problem (4.6) well conditioned, the excitation trajectory is often generated using an optimization scheme to obtain maximum excitation of all

base parameters. This involves nonlinear optimization with motion constraints on joint positions, velocities, and accelerations. Different optimization criteria exist, for example to minimize the condition number of Φ_b (Gautier and Khalil, 1991; Pfeiffer and Hölzl, 1995) or $-\log\det(\Phi_b^T W \Phi_b)$ (Swevers et al., 1997). Often a scaled version of the matrix Φ_b is used in the criterion to obtain parameter estimates with approximately the same relative accuracy. The matrix is scaled like $\bar{\Phi}_b = \Phi_b \operatorname{diag}\left\{\bar{\theta}_b\right\}$ where $\bar{\theta}_b$ contains a priori known parameter values.

For the optimization, the excitation trajectory must be parameterized, which can be done in many different ways. The most general one is perhaps Armstrong (1989), where the optimization variables are a sequence of joint accelerations. Gautier and Khalil (1991) use a sequence of joint positions and velocities as variables. Afterward, a continuous trajectory is obtained by interpolating a 5th order polynomial between these points, assuming zero initial and final acceleration. A problem is to be sure that the resulting trajectory fulfills all constraints and there is also no guarantee of optimality. In Swevers et al. (1997), the trajectory is parameterized as a finite Fourier series and the optimization variables are the coefficients in this series. Pfeiffer and Hölzl (1995) instead optimize the trajectory such that the trajectory always follows the steepest descent of the optimization criterion (time is discretized).

Grotjahn et al. (2001) suggest that the base parameters are divided into three groups; gravitational parameters θ_g, diagonal parameters θ_{M_d}, and off-diagonal parameters $\theta_{M_{od}}$. The parameter groups are identified one at a time by simple trajectories with trapezoidal velocity profile and a weighted least-squares method. The method has been successfully applied to the 6 DOF industrial robot manutec-r15.

A different approach is taken in Chenut et al. (2000) where additional sensors are used to improve the parameter estimates. They present a parameter estimation scheme where they combine measurements of the joint motion and joint torque with torque/force measurements of the base platform. The method is illustrated in simulations and yields improved parameter estimates.

A recent contribution is given in Swevers et al. (2007) where they consider both identification of rigid body dynamics and payload identification. Experiments are carried out using a KUKA KR 15 robot with a payload capacity of 15 kg. A key feature of their method is the use of a periodic excitation, already introduced in Swevers et al. (1997).

4.4 Flexibilities and Nonlinearities

As was mentioned in Section 2.2.4, there are various sources of flexibility in an industrial robot, such as elastic deformation of bearings and gears and deflection of the links under load. For many robots, the joint flexibility is dominant. A coupled two-mass model (2.10) is then often sufficient to model the dynamics. As was pointed out in Section 2.2.4, a trend is to build more compliant robot arms which also introduce significant flexibilities in the links and their connections. Therefore higher order models are sometimes needed in order to get a sufficiently accurate description of the system.

Identification of flexibilities is more complex than the identification of rigid body dynamics. The main reason is that only a subset of the state variables now are measured and one can therefore not use linear regression. This could of course be solved by adding

sensors (Albu-Schäffer and Hirzinger, 2001; Pfeiffer and Hölzl, 1995). A problem is that such a solution is expensive and the experiments quite involved with additional signals to measure, mounting of sensors, and so on, which will make such a solution harder to use as a simple way of tuning numerous robots, as well as re-tuning robots at the customer site. Also, as was mentioned in the problem statement in Section 1.3, this thesis investigates how much that can be achieved by using only the standard sensors.

Nonlinearities here refer to friction, backlash, nonlinear stiffness, and so on. Identification of friction and backlash will be treated in Section 4.4.1, followed by identification of the flexible joint model in Section 4.4.2. Identification using additional sensors will be briefly covered in Section 4.4.3. Finally, Section 4.4.4 gives some details on identification of the extended flexible joint model.

4.4.1 Identification of Friction and Backlash

In addition to flexibilities, nonlinear effects like friction and backlash have to be taken into account during the identification of the joint model. These effects will significantly influence the identification of the dynamics if not properly handled. Backlash in a gearbox can be seen as a piecewise-linear spring, as in Example 3.3, where the stiffness is low (zero in theory) for low torques. In a nonparametric FRF estimate, backlash will therefore show up as a reduced stiffness. The effects of Coulomb friction were also shown in Example 3.3, where the resonances almost disappeared and the magnitude for low frequencies was reduced. Moreover, nonlinearities can have major effects on the control performance, for example when the joint rotation changes direction, and accurate models describing these phenomena are therefore of importance.

The research on modeling and control of friction has come quite far, see for example the survey article Armstrong-Hélouvry et al. (1994). However, a problem is to be able to identify all unknown parameters in these complex friction models. Hence, the traditional model with the three components; viscous, Coulomb and static friction, is still often used. This model describes the stationary velocity-force relation fairly well, whereas important dynamic effects are completely missed. For control purposes and analysis of stick-slip, limit-cycles, etc., a better friction model is often needed. A fairly simple dynamic model for friction that captures most of the friction behavior is presented in Canudas de Wit et al. (1995) and is often called the LuGre model. This includes the Striebeck effect, hysteresis, spring-like characteristics for stiction, and varying break-away force.

Still, most of the friction models considered in the literature on identification in robotics are fairly simple. One reason might be that the model is sufficient during the identification of the rigid body dynamics, even though there are already some results on biased estimates due to a too simple friction model (Grotjahn et al., 2001). Another reason might be lack of measurements or that many friction problems are solved by better hardware (for example, lubricants). From now on, when talking about friction, we mean more than just viscous friction, which gives minor difficulties during identification and can be described by a linear model.

Backlash in mechanical systems is, for example, described in the survey article Nordin and Gutman (2002), where they use the definition: “the play between adjacent movable parts (as in a series of gears)”. The width of the backlash is the most essential parameter. Controlled systems with backlash often exhibit steady-state errors or, even worse,

limit cycles whereby the system oscillates. The transmission of industrial robots is often designed to minimize backlash. The reason is that since only the motor position is measured, the transmission must be predictable such that the end effector can be accurately positioned. Reducing the backlash often gives a more complicated mechanical design of the transmission and increased friction.

For the identification of nonlinearities such as friction and backlash, a number of references could be given. One interesting method is presented in Chen et al. (2002) which is based on open-loop experiments with binary input signals. By using the difference of the input-output signals from two experiments with different magnitudes, the friction can be ignored while doing the identification of the linear system. In a second step, the friction is estimated using the dc level of the linear system and a constant input. Isaksson et al. (2003) consider gray-box identification of a two-mass model with backlash, where black-box modeling is used to find initial parameter values, and Hovland et al. (2002) estimate backlash and spring stiffness using an extended Kalman filter (EKF). Identification of backlash and friction are also dealt with in Angerer et al. (2004) and Kara and Eker (2004). In this thesis, Paper F considers identification of a nonlinear gray-box model that includes friction and nonlinear spring stiffness.

4.4.2 Identification of Flexible Joint Models

There is a vast amount of literature available concerning the identification of joint parameters, strongly differing in both the model's level of detail and identification methods. The identification is done in time domain or frequency domain, based on physical or unstructured models, using linear or nonlinear optimization techniques. There are also major differences in how the nonlinearities are handled. Some estimate friction separately in a first step, and then in a second step estimate the joint dynamics. In this second step, they either compensate for the estimated friction, or use special excitation that will give minor frictional contribution. Using this route, linear identification methods can be applied. The other route is to use a nonlinear model structure describing the joint dynamics as well as friction and other nonlinearities.

In Pham et al. (2001) the estimation of physical parameters in a two-mass flexible model is treated through a least-squares technique similar to the one used in the estimation of rigid body dynamics. The linear-in-parameters formulation with only motor measurements is obtained through a nonlinear transformation and certain approximations that are justified by using a special trajectory that do not excite certain parameters. The identification experiments are carried out by moving one axis at a time.

In Östring et al. (2003), a method is applied where a gray-box model describing both inertial parameters and flexibilities can be identified directly in the time domain. This is done by utilizing a user-defined model structure in the MATLAB System Identification Toolbox and experimental data from closed-loop experiments for axis 1 of an ABB IRB1400 robot. Both black-box models and physically parameterized models are identified. Östring et al. suggest that a three-mass model is sufficient to describe the dynamics. See also Nissing and Polzer (2000), where the identification of a physically parameterized two-mass model in state-space form is treated. The three unknown physical parameters are estimated by using an iterative Gauss-Newton method and the experimental data comes from a 1 DOF hydraulic flexible arm.

In Berglund and Hovland (2000), a general method is described for the identification of masses, springs and dampers. The identification is based on an estimated nonparametric FRF in combination with the solution of an inverse eigenvalue problem. This procedure is used in Paper F to obtain initial estimates for the nonlinear prediction error method. See also Hovland et al. (2001) for an extension of Berglund and Hovland (2000) to systems containing coupled inertia terms, which is the case for multivariable systems. Another frequency-domain method is used in Khorrami et al. (1995), where recursive estimation of a two-mass model (for each joint) is treated. The parameters are estimated in the frequency domain using the ETFE or a time-varying version of the ETFE. The model is used for input pre-shaping (for example a notch filter for the reference signal).

In Johansson et al. (2000) they apply and evaluate different subspace identification methods (N4SID and MOESP) for identification of axes 1 and 4 of an ABB IRB2000 robot. They suggest the use of the MOESP algorithm for the identification of a black-box state-space model combined with a friction model.

Black-box identification is also treated in Ferretti et al. (1994b), where a 6 DOF industrial robot is considered. A third-order black-box transfer function is estimated for each joint. The excitation signal, applied as velocity reference for the controller, is a constant (to avoid effects of static friction) plus white noise. To eliminate, or at least minimize, the dynamic coupling among the links, different configurations are used for the identification of each joint dynamics. This is done by using a CAD model and symbolic manipulations of the robot dynamic equations. They also compute a physically parameterized expression of the transfer functions obtained from symbolic manipulations of the dynamic equations. They outline that a comparison of the estimated and physically parameterized transfer functions yields a system of nonlinear equations which could be solved for the physical parameters. No physical parameters are presented, but a similar procedure is used in Isaksson et al. (2003) for the estimation of initial values.

Discussion

One thing that most of the mentioned papers have in common is that they consider the robot joints one by one and estimate a SISO model for each joint. Coupling effects are then ignored. This thesis instead presents an identification procedure where the robot is treated as a MIMO system. This is an important contribution since coupling effects in a modern industrial robot cannot be ignored. Moreover, elastic effects in the bearings and the arm structure, which are handled by our MIMO approach, will also have a negative effect on a SISO procedure. Both since the additional elasticity will affect the estimated joint model, and since a joint model hardly can capture such complicated effects. Finally, our procedure immediately results in a global nonlinear dynamic model. This is hardly the case when estimating SISO models for each joint.

4.4.3 Identification Using Additional Sensors

As was previously mentioned, there are solutions where additional sensors are used for the identification. In Pfeiffer and Hölzl (1995), the joint parameters (stiffness, damping, and motor inertia) are estimated by fixation of the links. Between each link and link fixation a force sensor is applied. A slightly different procedure is used in Ferretti et al.

(1994a), where joint stiffness is estimated by constraining the robot end effector and using a force sensor to measure the constraint forces. Another method, requiring additional sensors, is *experimental modal analysis*. It is a widespread method in the mechanical engineering society used to determine a structure's dynamic characteristics; namely, resonant frequencies, damping values, and the associated pattern of structural deformation called mode shapes. Parametric identification of industrial robots using experimental modal analysis is, for example, treated in Behi and Tesar (1991). Using accelerometers in 11 points and an impact hammer, transfer functions are estimated and modal parameters and mode shapes are obtained by curve fitting the system transfer functions. The oscillation of the system is produced by the deformations of the shoulder and elbow joints (joints 2 and 3) about their axes of rotation and the deflection of the base plate relative to the foundation (giving two DOF). Therefore, four spring-damper pairs are used and springs, dampers and masses are then derived from the estimated modal matrix. See also Avitabile (2001) and Verboven (2002) for a good introduction to modal analysis.

4.4.4 Identification of Extended Flexible Joint Models

The identification of general elastic robot models depend on the selected modeling approach. If , e.g., a modal description is used, the parameters are usually identified using experimental modal analysis, as was described in the previous section.

The adopted modeling approach in this thesis is explained in Section 2.2.4, resulting in the *extended flexible joint model* (2.13), repeated below as (4.7) for convenience.

$$\dot{x} = f(x,u,\theta) = \begin{bmatrix} \dot{q}_m \\ \dot{q}_a \\ \dot{q}_e \\ M_m^{-1}(u - \tau_{fm}(\dot{q}_m) - r_g\tau_g) \\ M_{ae}^{-1}(q_a,q_e)\left(\begin{bmatrix}\tau_g \\ \tau_e\end{bmatrix} - c_{ae}(q_a,q_e,\dot{q}_a,\dot{q}_e) - g_{ae}(q_a,q_e)\right) \end{bmatrix}, \tag{4.7a}$$

$$y = h(x,u,\theta) = \dot{q}_m, \tag{4.7b}$$

with

$$\tau_g = k_g(r_g q_m - q_a) + d_g(r_g\dot{q}_m - \dot{q}_a), \tag{4.7c}$$

$$\tau_e = -k_e q_e - d_e\dot{q}_e, \tag{4.7d}$$

$$x = \begin{bmatrix} q_m^T & q_a^T & q_e^T & \dot{q}_m^T & \dot{q}_a^T & \dot{q}_e^T \end{bmatrix}^T. \tag{4.7e}$$

Linearizing the Extended Flexible Joint Model

An important part of the identification procedure presented in the thesis is the linearization of the nonlinear state-space model (4.7) in a stationary point (x_0, u_0), which is defined as the solution to

$$0 = f(x_0,u_0,\theta), \tag{4.8a}$$

$$y_0 = h(x_0,u_0,\theta). \tag{4.8b}$$

For a given robot configuration $q_m^{(i)}$, we have $q_{m0} = q_m^{(i)}$ and (4.8a) immediately gives that

$$\dot{q}_{m0} = \dot{q}_{a0} = \dot{q}_{e0} = 0.$$

The remaining parts of x_0 (q_{a0} and q_{e0}) and u_0 are obtained by solving

$$\begin{bmatrix} \tau_{g0} \\ \tau_{e0} \end{bmatrix} = g_{ae}(q_{a0}, q_{e0}), \quad \tau_{g0} = k_g(r_g q_{m0} - q_{a0}),$$
$$\tau_{e0} = -k_e q_{e0}, \quad u_0 = \tau_{fm}(0) + r_g \tau_{g0},$$

where the velocity-dependent term $c_{ae}(\cdot)$ in (4.7a) vanishes for zero velocity.

The linearized dynamics in a neighborhood of (x_0, u_0) is obtained from a first-order Taylor series expansion

$$\Delta\dot{x} = A(\theta)\Delta x + B(\theta)\Delta u, \tag{4.9a}$$
$$\Delta y = C(\theta)\Delta x + D(\theta)\Delta u, \tag{4.9b}$$

where $\Delta x = x - x_0$, $\Delta u = u - u_0$, $\Delta y = y - y_0$, and the matrices $A(\theta)$, $B(\theta)$, $C(\theta)$, and $D(\theta)$ are given by

$$A(\theta) = \frac{\partial f}{\partial x}(x_0, u_0, \theta), \quad B(\theta) = \frac{\partial f}{\partial u}(x_0, u_0, \theta),$$
$$C(\theta) = \frac{\partial h}{\partial x}(x_0, u_0, \theta), \quad D(\theta) = \frac{\partial h}{\partial u}(x_0, u_0, \theta).$$

Note that the stationary point and the linearized dynamics depend on the parameter vector θ. Therefore, given experiments in a robot configuration $q_m^{(i)}$, the corresponding stationary point should actually be denoted $(x_0^{(i)}(\theta), u_0^{(i)}(\theta))$.

FRF of Extended Flexible Joint Model

The FRF $G^{(i)}(\omega, \theta)$ of the parametric robot model (4.7) in a robot configuration $q_m^{(i)}$ is obtained by the following steps:

1. Calculate the stationary point $(x_0^{(i)}(\theta), u_0^{(i)}(\theta))$ by solving (4.8).
2. Linearize the nonlinear system (4.7) according to (4.9).
3. Obtain the continuous-time transfer function as (cf. (3.27b))
$$G_c^{(i)}(s, \theta) = C^{(i)}(\theta)(sI - A^{(i)}(\theta))^{-1}B^{(i)}(\theta) + D^{(i)}(\theta).$$
4. Convert the continuous-time FRF $G_c^{(i)}(j\omega, \theta)$ to discrete time
$$G_{T_s}^{(i)}(e^{j\omega T_s}, \theta) = \frac{1 - e^{-j\omega T_s}}{j\omega T_s} G_c^{(i)}(j\omega, \theta),$$
by assuming zero-order hold and using the approximation (3.15). The parametric FRF $G^{(i)}(\omega, \theta)$ is a short notation for $G_{T_s}^{(i)}(e^{j\omega T_s}, \theta)$.

The parametric FRF $G^{(i)}(\omega, \theta)$ is used together with the estimated nonparametric FRF $\widehat{G}^{(i)}(\omega_k)$ to identify the unknown parameter vector θ by minimizing the discrepancy between $G^{(i)}(\omega, \theta)$ and $\widehat{G}^{(i)}(\omega_k)$ for a number of robot configurations and frequencies. For details, see Paper A.

4.5 EXPDES—An Identification Toolbox for Industrial Robots

The software tool EXPDES stands for Experiment Design and is a MATLAB toolbox written by the author of this thesis. It is used for:

- Creating offline-generated trajectories of multisine type for the robot controller.
- Importing and preprocessing experimental data.
- Estimating nonparametric FRFs from experimental data.
- Plotting the FRF together with its uncertainty.
- Estimating parametric robot models.

With the software tool, most of the research work in this thesis is implemented and made easily available to facilitate accurate estimation of nonparametric FRFs as well as parametric models.

The main GUI window for experiment design can be seen in Figure 4.1. The experimental data can be seen in Figure 4.2. The user selects the periodic part of the experimental data by clicking in the figure.

Most of the code is written in an object-oriented fashion, similarly as the MATLAB System Identification Toolbox. An example can be seen below.

Code 4.1 (MATLAB code using EXPDES for nonlinear gray-box identification)

```
% Load FRF data:
frf1 = frfdata('frf_data1.mat');
% Set weights for identification:
frf1 = setweight(frf1);
% Load initial nonlinear gray-box model:
m0 = elast('rob_rigid.m','rob_load.m','rob_elast.m');
% Set which parameters to estimate
% (some parameters can be locked):
m0 = setpar(m0,1:26,'Locked',0);
% Select interval for random perturbation:
v = getpar(m0,[1:26],'Value');
m0 = setpar(m0,[1:26],'Min',v*0.1,'Max',v*10);
% Estimate 100 models with randomly perturbed
% initial parameters:
copt = Inf;
```

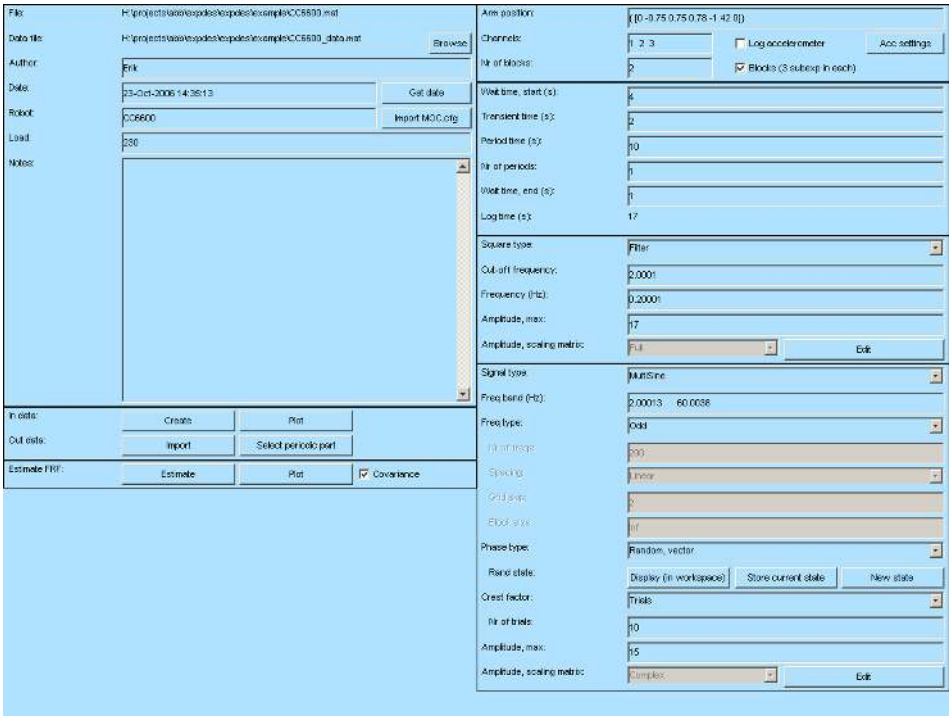


Figure 4.1: EXPDES *GUI main window.*

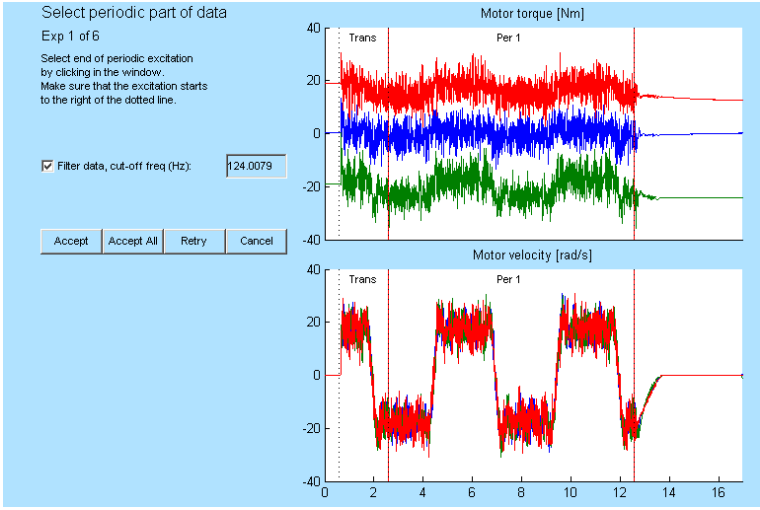


Figure 4.2: EXPDES *GUI window for selection of periodic data.*

```
for ii=1:100
  mi_init{ii} = randpar(m0);
  mi{ii} = est(frf1,mi_init{ii});
  ci(ii) = cost(frf1,mi{ii});
  if ci(ii) < copt
    copt = ci(ii);
    mopt = mi{ii};
  end
end
% Compare nonparametric FRF and linearized model
% for experiment nr 8:
compare(getexp(frf1,8),mopt);
```

This code is used in Paper A to assess the sensitivity to the initial parameters when estimating parameters in a nonlinear dynamic gray-box model. The estimated nonparametric FRFs in a number of positions are imported to the object `frf1` of type `frfdata`. Similarly, the object type `elast` is an example of a nonlinear gray-box model structure that is used. These objects have a number of methods defined to facilitate easy modification of weights, criterion, initial parameters, and so on. The last line, using `compare`, generates a figure, similarly to Figure 8 in Paper A, where the nonparametric FRF `frf1` with its uncertainty is plotted together with the FRF of the (linearized) parametric model `mopt`.

5

Concluding Remarks

This first part of the thesis has served as an introduction to the robotics area and system identification methods with relevance to our studied problem have been presented. The aim has been to show how the papers in Part II relate to each other and to the existing methods. In Section 5.1 the conclusions are given. A number of areas are still subject to future research. Some ideas are discussed in Section 5.2. See also the included papers for details.

5.1 Conclusion

The work presented in this thesis has dealt with dynamic model identification of industrial robots. More specifically, the models are physically parameterized nonlinear dynamic models that also capture resonances due to various elastic effects in, e.g., gearboxes, bearings, foundations, and the robot links. The elastic effects are modeled through a lumped parameter approach where each rigid body is connected by spring-damper pairs.

Paper A describes the adopted solution, where the unknown parameters (mainly the spring-damper pairs) are identified in the frequency domain, using estimates of the nonparametric frequency response function (FRF) in different robot configurations/positions. Different parameter estimators are compared and experimental results show the usefulness of the proposed identification procedure. The weighted logarithmic least squares estimator achieves the best result and the identified model gives a good global description of the dynamics in the frequency range of interest.

Different methods for estimating the nonparametric multivariable FRF are proposed in the literature. In Paper B, some classical estimators (H_1, arithmetic mean, joint input-output, and errors-in-variables) are theoretically analyzed, both in open loop and closed loop, and expressions for the bias and covariance are derived. In Paper C, the classical estimators as well as two estimators based on nonlinear averaging techniques (logarithmic mean, harmonic mean) are compared in an experimental study. The conclusion is that

the estimators based on nonlinear averaging give the best results for the industrial robot. Paper D next deals with how the nonparametric FRF estimate is affected by nonlinearities in the system. The results show that nonlinear distortions are indeed present and cause larger variability in the FRF than the measurement noise contributions. To improve the nonparametric FRF estimate, one should therefore use the available measurement time to average estimates from several experiments with the orthogonal random phase multisine. The importance of reducing the effects of static friction is also stressed.

As was mentioned earlier, the adopted solution is based on nonparametric FRF estimates in a number of positions. The selection of these positions in an optimal way is treated in Paper E. A convex optimization problem is posed, which is efficiently solved by considering the dual problem. In the numerical illustration, the 15 optimal positions reduce the parameter uncertainty by a factor of six, compared to using the optimal single position in all experiments.

The main approach in the thesis is to identify the robot model in the frequency domain, which has been explained in Papers A–E. In addition is time-domain nonlinear gray-box identification treated in Paper F, where experimental results are given for a stable scalar system (axis one of the robot). Extending the results of Paper F to a multivariable and unstable system is left as future research.

The research work is also implemented and made easily available in a software tool for accurate estimation of nonparametric FRFs as well as parametric robot models.

5.2 Future Research

For the nonparametric FRF estimation, different ways of reducing the effects of friction would be interesting to study. The selection of frequencies, as well as the amplitude spectrum can be further improved. This is important both to reduce bias in the nonparametric FRF estimate and to improve the estimated parametric robot model.

For the parametric estimators in Paper A, the selection of weights can certainly be improved by combining the user choices and the estimated FRF uncertainty.

The model structure in Paper A, where all elastic effects in the arm structure are lumped into a few spring-damper pairs, can of course be modified and refined, both regarding the location of these spring-damper pairs, as well as how many that are needed in order to properly model the system. Identifiability of these added parameters, can also be discussed. Maybe additional sensors are needed, e.g., accelerometers attached to the structure, as is the case in experimental modal analysis (Behi and Tesar, 1991). Additional sensors is therefore a topic that deserves more attention, for example to investigate how much that can be gained by using a simple 3D-accelerometer attached to the end effector.

In case the identification procedure is used for the tuning of individual robots, or at the customer site, the measurement time and excitation amplitudes become an issue. Parameter accuracy and problems with local minima, versus measurement time and excitation energy are then also interesting problems to study.

Using a frequency-domain method for identification of a nonlinear system has some problems, as was pointed out both in Section 1.3 and in Paper A. Therefore, it would be interesting to apply a time-domain prediction error method, as in Paper F, for the model used in Paper A. One big problem to tackle then is that the model is unstable, which

requires some changes to ensure a stable predictor. There can also be numerical problems since the dimension of the problem is larger, both in terms of states and parameters (compared to the scalar system considered in Paper F).

A thorough simulation study would also probably yield many interesting results, both for the nonparametric FRF estimation and for the parametric identification. The simulation setup should then include a realistic nonlinear robot model where friction as well as other nonlinear effects in actuators and sensors are added, such as torque ripple (Gutt et al., 1996) and resolver ripple (Hanselman, 1990).

Finally, to conclude this thesis: Identification of industrial robots is a challenging task. Using a general purpose method by pressing a button will almost surely fail. The problem instead requires a combination of tailored identification methods, experiment design, and a skilled user, using all available knowledge about the robot system.

Bibliography

ABB (2007). More than 30 years with ABB robotics. `http://www.abb.com/product/ap/seitp327/583a073bb0bb1922c12570c1004d3e6b.aspx`.

Abderrahim, M. and Whittaker, A. R. (2000). Kinematic model identification of industrial manipulators. *Robotics and Computer-Integrated Manufacturing*, 16(1):1–8.

Albu-Schäffer, A. and Hirzinger, G. (2001). Parameter identification and passivity based joint control for a 7DOF torque controlled light weight robot. In *Proc. 2001 IEEE International Conference on Robotics and Automation*, pages 2852–2858, Seoul, Korea.

Albu-Schäffer, A. and Hirzinger, G. (2000). State feedback controller for flexible joint robots: A globally stable approach implemented on DLR's light-weight robots. In *Proc. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1087–1093, Takamatsu, Japan.

Ambrose, R. O. and Ambrose, C. G. (1995). Robot models for space environments. In *Proc. 1995 IEEE International Conference on Robotics and Automation*, volume 2, pages 2113–2118, Nagoya, Japan.

An, C. H., Atkeson, C. G., and Hollerbach, J. M. (1988). *Model-Based Control of a Robot Manipulator*. MIT Press, Cambridge, Massachusetts.

Angerer, B. T., Hintz, C., and Schröder, D. (2004). Online identification of a nonlinear mechatronic system. *Control Engineering Practice*, 12(11):1465–1478.

Armstrong, B. (1989). On finding exciting trajectories for identification experiments involving systems with nonlinear dynamics. *International journal of robotics research*, 8(6):28–48.

Armstrong-Hélouvry, B., Dupont, P., and Canudas de Wit, C. (1994). A survey of models, analysis tools and compensation methods for the control of machines with friction. *Automatica*, 30(7):1083–1138.

Åström, K. and Wittenmark, B. (1984). *Computer Controlled Systems – Theory and Design*. Prentice-Hall, Englewood Cliffs, New Jersey.

Avitabile, P. (2001). Experimental modal analysis - a simple non-mathematical presentation. *Sound and vibration*, 35(1):20–31.

Bai, E.-W. (2003a). Frequency domain identification of hammerstein models. *IEEE Transactions on Automatic Control*, 48(4):530–542.

Bai, E.-W. (2003b). Frequency domain identification of wiener models. *Automatica*, 39(9):1521–1530.

Bascetta, L. and Rocco, P. (2002). Modelling flexible manipulators with motors at the joints. *Mathematical and Computer Modelling of Dynamical Systems*, 8(2):157–183.

Behi, F. and Tesar, D. (1991). Parametric identification for industrial manipulators using experimental modal analysis. *IEEE Transactions on Robotics and Automation*, 7(5):642–652.

Berglund, E. and Hovland, G. E. (2000). Automatic elasticity tuning of industrial robot manipulators. In *Proc. 39th IEEE Conference on Decision and Control*, pages 5091–5096, Sydney, Australia.

Brogårdh, T. (2007). Present and future robot control development—an industrial perspective. *Annual Reviews in Control*, 31(1):69–79.

Brogårdh, T., Ahlbäck, M., Bergsjö, J., Elfving, S., Lager, A., Moberg, S., Myhr, M., and Rylander, D. (2006). Method for thermal optimization. US Patent 7084595. Available from `http://www.patentstorm.us/patents/7084595.html`.

Brogårdh, T., Dahlquist, H., Elfving, S., and Henriksson, T. (2001). Method for control of an industrial robot. US Patent 6230079. Available from `http://www.patentstorm.us/patents/6230079.html`.

Canudas de Wit, C., Olsson, H., Åström, K. J., and Lischinsky, P. (1995). A new model for control of systems with friction. *IEEE Transactions on Automatic Control*, 40(3):419–425.

Chen, Y.-Y., Huang, P.-Y., and Yen, J.-Y. (2002). Frequency-domain identification algorithms for servo systems with friction. *IEEE Transactions on Control Systems Technology*, 10(5):654–665.

Chenut, X., Samin, J. C., Swevers, J., and Ganseman, C. (2000). Combining internal and external robot models for improved model parameter estimation. *Mechanical Systems and Signal Processing*, 14(5):691–704.

Craig, J. J. (1989). *Introduction to Robotics Mechanics and Control*. Addison Wesley.

Crama, P. and Schoukens, J. (2001). Initial estimates of wiener and hammerstein systems using multisine excitation. *IEEE Transactions on Instrumentation and Measurement*, 50(6):1791–1795.

Dobrowiecki, T. and Schoukens, J. (2007). Measuring a linear approximation to weakly nonlinear MIMO systems. *Automatica*, 43(10):1737–1751.

Doucet, A., de Freitas, N., and Gordon, N., editors (2001). *Sequential Monto Carlo Methods in Practice*. Springer Verlag.

Enqvist, M. (2005). *Linear Models of Nonlinear Systems*. Dissertations. no. 985, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden.

Fedorov, V. V. (1972). *Theory of optimal experiments*. Academic Press, New York.

Ferretti, G., Maffezzoni, C., Magnani, G., and Rocco, P. (1994a). Joint stiffness estimation based on force sensor measurements in industrial manipulators. *Journal of Dynamic Systems Measurement and Control, Transactions of the ASME*, 116:163–167.

Ferretti, G., Magnani, G., and Rocco, P. (1994b). Estimation of resonant transfer functions in the joints of an industrial robot. In *2nd IFAC Symposium on Intelligent Control and Control Applications, SICICA 94*, volume 1, pages 371–376, Budapest, Hungary.

Forssell, U. (1999). *Closed-loop Identification: Methods, Theory, and Applications*. Dissertations. no. 566, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden.

Forssell, U. and Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, 35:1215–1241.

Funda, J., Taylor, R. H., and Paul, R. P. (1990). On homogeneous transforms, quaternions, and computational efficiency. *IEEE Transactions on Robotics and Automation*, 6(3):382–388.

Gautier, M. (1990). Numerical calculation of the base inertial parameters of robots. In *Proc. 1990 IEEE International Conference on Robotics and Automation*, volume 2, pages 1020–1025, Cincinnati, Ohio.

Gautier, M. and Khalil, W. (1988). On the identification of the inertial parameters of robots. In *Proc. 27th IEEE Conference on Decision and Control*, pages 2264–2269, Austin, Texas.

Gautier, M. and Khalil, W. (1990). Direct calculation of minimum set of inertial parameters of serial robots. *IEEE Transactions on Robotics and Automation*, 6(3):368–373.

Gautier, M. and Khalil, W. (1991). Exciting trajectories for the identification of base inertial parameters of robots. In *Proc. 30th IEEE Conference on Decision and Control*, pages 494–499, Brighton, England.

Gautier, M. and Poignet, P. (2001). Extended kalman filtering and weighted least squares dynamic identification of robot. *Control Engineering Practice*, 9(12):1361–1372.

Gautier, M., Vandanjon, P. O., and Presse, C. (1994). Identification of inertial and drive gain parameters of robots. In *Proc. 33rd IEEE Conference on Decision and Control*, pages 3764–3769, Lake Buena Vista, Florida.

Gevers, M. (2005). Identification for control: From the early achievements to the revival of experiment design. *European Journal of Control*, 11(4-5):335–352.

Goodwin, G. C. and Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press, New York.

Goodwin, G. C., Rojas, J. S., and Welsh, J. S. (2006). Good, bad and optimal experiments for identification. In Glad, T. and Hendeby, G., editors, *Forever Ljung in System Identification*, pages 103–125. Studentlitteratur, Lund.

Graeser, R.-G. and Klingel, R. (2002). Control method for an industrial robot. US Patent 6345213.

Grotjahn, M., Daemi, M., and Heimann, B. (2001). Friction and rigid body identification of robot dynamics. *International Journal of Solids and Structures*, 38:1889–1902.

Guillaume, P. (1998). Frequency response measurements of multivariable systems using nonlinear averaging techniques. *IEEE Transactions on Instrumentation and Measurement*, 47(3):796–800.

Guillaume, P., Pintelon, R., and Schoukens, J. (1996). Accurate estimation of multivariable frequency response functions. In *Proc. 13th IFAC World Congress*, pages 423–428, San Francisco, California.

Gunnar, J., Wernholt, E., Hovland, G., and Brogårdh, T. (2006). Nonlinear grey-box identification of linear actuators containing hysteresis. In *Proc. 2006 IEEE International Conference on Robotics and Automation*, pages 1818–1823, Orlando, Florida.

Gutt, H.-J., Scholl, F. D., and Blattner, J. (1996). High precision servo drives with DSP-based torque ripple reduction. In *Proc. IEEE AFRICON, 1996*, volume 2, pages 632–637.

Hanselman, D. (1990). Resolver signal requirements for high accuracy resolver-to-digital conversion. *IEEE Transactions on Industrial Electronics*, 37(6):556–561.

Hollerbach, J. M. (1989). A survey of kinematic calibration. In Khatib, O., Craig, J. J., and Lozano-Perez, T., editors, *The robotics review. Vol. 1*, pages 207–242. MIT Press, Cambridge, MA.

Holtz, J. and Springob, L. (1996). Identification and compensation of torque ripple in high-precision permanent magnet motor drives. *IEEE Transactions on Industrial Electronics*, 43(2):309–320.

Hovland, G. E., Berglund, E., and Hanssen, S. (2001). Identification of coupled elastic dynamics using inverse eigenvalue theory. In *Proc. 32nd International Symposium on Robotics (ISR)*, pages 1392–1397, Seoul, Korea.

Hovland, G. E., Hanssen, S., Gallestey, E., Moberg, S., Brogårdh, T., Gunnarsson, S., and Isaksson, M. (2002). Nonlinear identification of backlash in robot transmissions. In *Proc. 33rd International Symposium on Robotics (ISR)*, Stockholm, Sweden.

IFR (2006). *World Robotics 2006*, chapter Executive summary. IFR Statistical Department. Available from `http://www.ifrstat.org/downloads/2006_Executive_Summary(1).pdf`.

Isaksson, A., Lindkvist, R., Zhang, X., Nordin, M., and Tallfors, M. (2003). Identification of mechanical parameters in drive train systems. In *Proc. 13th IFAC Symposium on System Identification*, pages 1542–1547, Rotterdam, The Netherlands.

Jahns, T. M. and Soong, W. L. (1996). Pulsating torque minimization techniques for permanent magnet AC motor drives – a review. *IEEE Transactions on Industrial Electronics*, 43(2):321–330.

Jansson, H. and Hjalmarsson, H. (2005). Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, 50(10):1534–1549.

Johansson, R., Robertsson, A., Nilsson, K., and Verhaegen, M. (2000). State-space system identification of robot manipulator dynamics. *Mechatronics*, 10(3):403–418.

Kara, T. and Eker, I. (2004). Nonlinear modeling and identification of a DC motor for bidirectional operation with real time experiments. *Energy Conversion and Management*, 45(7-8):1087–1106.

Khalil, W. and Besnard, S. (2002). Geometric calibration of robots with flexible joints and links. *Journal of Intelligent and Robotic Systems*, 34:357–379.

Khalil, W. and Dombre, E. (2002). *Modeling, Identification & Control of Robots*. Taylor & Francis Group, New York, 3rd edition.

Khalil, W. and Gautier, M. (2000). Modeling of mechanical systems with lumped elasticity. In *Proc. 2000 IEEE International Conference on Robotics and Automation*, pages 3964–3969, San Francisco, California.

Khorrami, F., Jain, S., and Tzes, A. (1995). Experimental results on adaptive nonlinear control and input preshaping for multi-link flexible manipulators. *Automatica*, 31(1):83–97.

Kozlowski, K. (1998). *Modelling and identification in robotics*. Advances in Industrial Control. Springer, London.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Ljung, L. (2003). Linear system identification as curve fitting. In Rantzer, A. and Byrnes, C. I., editors, *Directions in Mathematical Systems Theory and Optimization (Lecture Notes in Control and Information Sciences)*, volume 286, pages 203–215. Springer Verlag.

Ljung, L. (2004). State of the art in linear system identification: Time and frequency domain methods. In *Proc. 2004 American Control Conference*, pages 650–660, Boston, Massachusetts.

Ljung, L. and Glad, S. T. (1994). *Modeling of Dynamic Systems*. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Löfberg, J. (2004). YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proc. CACSD Conference*, Taipei, Taiwan. Available from `http://control.ee.ethz.ch/~joloef/yalmip.php`.

Mayeda, H., Yoshida, K., and Osuka, K. (1990). Base parameters of manipulator dynamic models. *IEEE Transactions on Robotics and Automation*, 6(3):312–321.

Moberg, S. and Hanssen, S. (2007). A DAE approach to feedforward control of flexible manipulators. In *Proc. 2007 IEEE International Conference on Robotics and Automation*, pages 3439–3444, Roma, Italy.

Nissing, D. and Polzer, J. (2000). Parameter identification of a substitution model for a flexible link. In *Proc. 12th IFAC Symposium on System Identification*, pages 1715–1719, Santa Barbara, California.

Nordin, M. and Gutman, P.-O. (2002). Controlling mechanical systems with backlash - a survey. *Automatica*, 38(10):1633–1649.

Nyström, M. and Norrlöf, M. (2003). Path generation for industrial robots. Technical Report LiTH-ISY-R-2529, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden.

Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.

Olsson, T. (2007). *High-Speed Vision and Force Feedback for Motion-Controlled Industrial Manipulators*. ISRN LUTFD2/TFRT--1078--SE, LTH, Lund University.

Olsson, T., Robertsson, A., and Johansson, R. (2007). Flexible force control for accurate low-cost robot drilling. In *Proc. 2007 IEEE International Conference on Robotics and Automation*, pages 4770–4775, Roma, Italy.

Ortega, R. and Spong, M. W. (1989). Adaptive motion control of rigid robots: A tutorial. *Automatica*, 25(6):877–888.

Östring, M., Gunnarsson, S., and Norrlöf, M. (2003). Closed-loop identification of an industrial robot containing flexibilities. *Control Engineering Practice*, 11:291–300.

Pfeiffer, F. and Hölzl, J. (1995). Parameter identification for industrial robots. In *Proc. 1995 IEEE International Conference on Robotics and Automation*, volume 2, pages 1468–1476, Nagoya, Japan.

Pham, M. T., Gautier, M., and Poignet, P. (2001). Identification of joint stiffness with bandpass filtering. In *Proc. 2001 IEEE International Conference on Robotics and Automation*, pages 2867–2872, Seoul, Korea.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach*. IEEE Press, New York.

Pintelon, R. and Schoukens, J. (2001). Measurement of frequency response functions using periodic excitations, corrupted by correlated input/output errors. *IEEE Transactions on Instrumentation and Measurement*, 50(6):1753–1760.

Presse, C. and Gautier, M. (1993). New criteria of exciting trajectories for robot identification. In *Proc. 1993 IEEE International Conference on Robotics and Automation*, volume 3, pages 907–912, Atlanta, Georgia.

Rugh, W. J. (1996). *Linear system theory*. Prentice Hall, Upper Saddle River, New Jersey, 2nd edition.

Schoukens, J., Dobrowiecki, T., and Pintelon, R. (1998). Parametric and nonparametric identification of linear systems in the presence of nonlinear distortions—a frequency domain approach. *IEEE Transactions on Automatic Control*, 43(2):176–190.

Schoukens, J., Pintelon, R., Dobrowiecki, T., and Rolain, Y. (2005). Identification of linear systems with nonlinear distortions. *Automatica*, 41(3):491–504.

Schoukens, J., Pintelon, R., and Rolain, Y. (2004). Time domain identification, frequency domain identification. Equivalencies! Differences? In *Proc. 2004 American Control Conference*, pages 661–666, Boston, Massachusetts.

Schoukens, J., Pintelon, R. M., and Rolain, Y. J. (2000). Broadband versus stepped sine FRF measurements. *IEEE Transactions on Instrumentation and Measurement*, 49(2):275–278.

Sciavicco, L. and Siciliano, B. (2000). *Modeling and Control of Robotic Manipulators*. Springer.

Sheu, S.-Y. and Walker, M. W. (1989). Basis sets for manipulator inertial parameters. In *Proc. 1989 IEEE International Conference on Robotics and Automation*, volume 3, pages 1517–1522, Scottsdale, Arizona.

Smolders, K. and Swevers, J. (2006). Nonparametric MIMO FRF matrix estimation using a single periodic broadband excitation. In *Proc. 2006 IEEE International Conference on Control Applications*, pages 2583–2588, Munich, Germany.

Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice-Hall Int., London.

Spong, M. W. (1987). Modeling and control of elastic joint robots. *Journal of Dynamic Systems Measurement and Control*, 109(4):310–319.

Spong, M. W., Lewis, F. L., and Abdallah, C. T., editors (1993). *Robot control : dynamics, motion planning and analysis*. IEEE Press.

Spong, M. W. and Vidyasagar, M. (1989). *Robot Dynamics and Control*. Wiley.

Swevers, J., Ganseman, C., Tükel, D. B., De Schutter, J., and Van Brussel, H. (1997). Optimal robot excitation and identification. *IEEE Transactions on Robotics and Automation*, 13(5):730–740.

Swevers, J., Verdonck, W., and De Schutter, J. (2007). Dynamic model identification for industrial robots. *IEEE Control Systems Magazine*, 27(5):58–71.

Theodore, R. J. and Ghosal, A. (1995). Comparison of the assumed modes and finite element models for flexible multilink manipulators. *International Journal of Robotics Research*, 14:91–111.

Uddeholt, E. (1998). *Identifiering och kompensering av motor- och resolverrippel*. Master thesis IR-RT-EX-9806, Dept. of Signals, Sensors & Systems, Royal Institute of Technology, SE-100 44 Stockholm, Sweden.

Van der Ouderaa, E., Schoukens, J., and Renneboog, J. (1988). Peak factor minimization using a time-frequency domain swapping algorithm. *IEEE Transactions on Instrumentation and Measurement*, 37(1):145–147.

Van Overschee, P. and DeMoor, B. (1996). *Subspace Identification of Linear Systems: Theory, Implementation, Applications*. Kluwer Academic Publishers.

Verboven, P. (2002). *Frequency-domain system identification for modal analysis*. PhD thesis, Vrije Universiteit Brussel, Belgium.

Wang, S. and Zhao, J. (2002). FEM optimization for robot structure. In *Proc. 2002 IEEE International Conference on Industrial Technology*, volume 1, pages 510–513.

Wernholt, E. (2004). *On Multivariable and Nonlinear Identification of Industrial Robots*. Licentiate thesis no. 1131, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden.

Wernholt, E. and Gunnarsson, S. (2004). On the use of a multivariable frequency response estimation method for closed loop identification. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 827–832, Nassau, Bahamas.

Wernholt, E. and Gunnarsson, S. (2005). Nonlinear grey-box identification of industrial robots containing flexibilities. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Wernholt, E. and Gunnarsson, S. (2006a). Detection and estimation of nonlinear distortions in industrial robots. In *Proc. 23rd IEEE Instumentation and Measurement Technology Conference*, pages 1913–1918, Sorrento, Italy.

Wernholt, E. and Gunnarsson, S. (2006b). Nonlinear identification of a physically parameterized robot model. In *Proc. 14th IFAC Symposium on System Identification*, pages 143–148, Newcastle, Australia.

Wernholt, E. and Gunnarsson, S. (2007a). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Wernholt, E. and Gunnarsson, S. (2007b). Estimation of nonlinear effects in frequency-domain identification of industrial robots. *Accepted for publication in IEEE Transactions on Instrumentation and Measurement*.

Wernholt, E. and Gunnarsson, S. (2007c). Nonlinear gray-box identification of a flexible manipulator. *Submitted to Control Engineering Practice*.

Wernholt, E. and Löfberg, J. (2007). Experiment design for identification of nonlinear gray-box models with application to industrial robots. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Wernholt, E. and Moberg, S. (2007a). Experimental comparison of methods for multivariable frequency response function estimation. Technical Report LiTH-ISY-R-2827, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Wernholt, E. and Moberg, S. (2007b). Frequency-domain gray-box identification of industrial robots. Technical Report LiTH-ISY-R-2826, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Westerlund, L. (2000). *The Extended Arm of Man – A History of the Industrial Robot*. Informationsförlaget, Stockholm, Sweden.

Zhong, X.-L., Lewis, J. M., and L.N.-Nagy, F. (1996). Autonomous robot calibration using a trigger probe. *Robotics and Autonomous Systems*, 18(4):395–410.

Part II

Publications

Paper A

Frequency-Domain Gray-Box Identification of Industrial Robots

Edited version of the paper:

Wernholt, E. and Moberg, S. (2007b). Frequency-domain gray-box identification of industrial robots. Technical Report LiTH-ISY-R-2826, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Frequency-Domain Gray-Box Identification of Industrial Robots

Erik Wernholt[1] and Stig Moberg[1,2]

[1]Dept. of Electrical Engineering,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: {erikw,stig}@isy.liu.se.

[2]ABB AB – Robotics,
SE–721 68 Västerås, Sweden.

Abstract

This paper considers identification of unknown parameters in elastic dynamic models of industrial robots. Identifying such models is a challenging task since an industrial robot is a multivariable, nonlinear, resonant, and unstable system. Unknown parameters (mainly spring-damper pairs) in a physically parameterized nonlinear dynamic model are identified in the frequency domain, using estimates of the nonparametric frequency response function (FRF) in different robot configurations/positions. The nonlinear parametric robot model is linearized in the same positions and the optimal parameters are obtained by minimizing the discrepancy between the nonparametric FRFs and the parametric FRFs (the FRFs of the linearized parametric robot model). In order to accurately estimate the nonparametric FRFs, the experiments must be carefully designed. The selection of optimal robot configurations for the experiments is also part of the design. Different parameter estimators are compared and experimental results show the usefulness of the proposed identification procedure. The weighted logarithmic least squares estimator achieves the best result and the identified model gives a good global description of the dynamics in the frequency range of interest.

Keywords: System identification, multivariable systems, nonlinear systems, closed-loop identification, frequency response methods, industrial robots

1 Introduction

Accurate dynamic models of industrial robots are needed for mechanical design, performance simulation, control, supervision, diagnosis, and so on. The industrial robot poses a

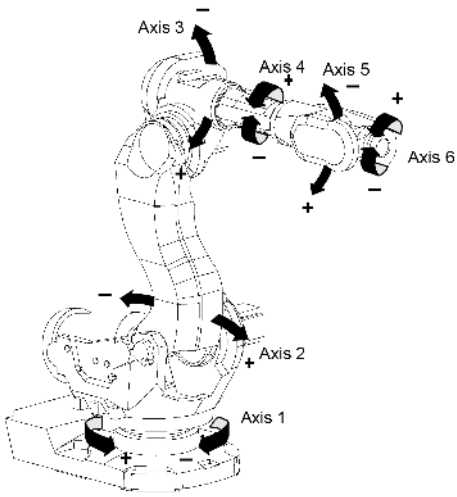


Figure 1: *The ABB manipulator IRB6600.*

challenging modeling problem both due to the system complexity and the required model accuracy. Usually a robot has six joints (also called axes), see Figure 1, with coupled dynamics, giving a truly multivariable system. The dynamics is nonlinear, both with respect to the rigid body dynamics and other things such as non-ideal motors and sensors, and a transmission with friction, backlash, hysteresis, and nonlinear stiffness. The system is resonant due to elastic effects and, in addition, experimental data must usually be collected while the robot controller is operating in closed loop since the system is unstable.

Historically, the dynamic models used for control are either entirely rigid (An et al., 1988), or only flexible joint models are considered, i.e., elastic gear transmission and rigid links (Albu-Schäffer and Hirzinger, 2000; Spong, 1987). The trend in industrial robots is toward lightweight robot structures with a reduced mass but with preserved payload capabilities. This is motivated by cost reduction as well as safety issues, but results in lower mechanical resonance frequencies inside the controller bandwidth. The sources of elasticity in such a manipulator are, e.g., gearboxes, bearings, elastic foundations, elastic payloads, as well as bending and torsion of the links. In Öhr et al. (2006) it is shown that there are cases when these other sources of flexibilities can be of the same order as the gearbox flexibilities for a modern industrial robot. Accurate dynamic models that also describe these elastic effects are therefore needed in order to obtain high performance. These models are, however, very difficult to use for robot control, where, e.g., feedforward control involves solving a DAE, but could in the future improve the performance (Moberg and Hanssen, 2007).

2 Problem Description

The main problem considered in this paper is about identification of unknown parameters in a nonlinear dynamic model of an industrial robot. The model must be *global*, i.e., valid

throughout the whole workspace (all robot configurations/positions), as well as *elastic*, which here means that resonances due to elastic effects are captured by the model. The elastic effects are modeled through a lumped parameter approach (Khalil and Gautier, 2000) where each rigid body is connected by spring-damper pairs (see also Section 3). The model is of gray-box type (a physically parameterized model) and the rigid body parameters of the model are usually assumed to be known from a CAD model or prior rigid body identification. The main objective is identification of elasticity parameters (spring-damper pairs) but other parameters can be added, such as the location in the robot structure of the spring-damper pairs and a few unknown rigid body parameters. It is also possible to include nonlinear descriptions of selected quantities (e.g., the gearbox stiffness) and identify those by a linearization for each position (the terms *robot configuration* and *position* are used interchangeably in this paper).

The real challenge for system identification methods is that the industrial robot is multivariable, nonlinear, unstable, and resonant at the same time. Usually, in the literature, at least one of the first three topics is left out. Identification of such a complex system is therefore a huge task, both in finding suitable model structures and efficient identification methods.

One solution could be to apply a nonlinear prediction error method (Ljung, 1999, pp. 146–147), where measured input-output data are fed to the model and the predicted output from the model is compared with the measured output. This has been treated in Wernholt and Gunnarsson (2006b) for axis one of the industrial robot, which means a stable scalar system (axis one is not affected by gravity). Extending these results to a multivariable and unstable system would involve, for example: finding a stable predictor, numerical problems, and handling large data sets. The last two problems stem from the fact that the system is resonant and numerically stiff, as well as large in dimension both with respect to the number of states and parameters. In addition comes also the choice of model structure (parameters) and handling local minima in the optimization. Apart from all these problems, such a solution would really tackle our main problem.

Due to the complexity of the industrial robot, it is common practice to estimate approximate models for various purposes. By, for example, using a low-frequency excitation, elastic effects have a minor influence and a nonlinear model of the rigid body dynamics can be estimated using least squares techniques. This is a much studied problem in the literature, see, e.g., Kozlowski (1998) for an overview. Taking elastic effects into account makes the identification problem much harder. The main reason is that only a subset of the state variables now are measured such that linear regression cannot be used. One option could be to add sensors during the data collection to measure all states, even though accurate measurements of all states are not at all easy to obtain (if even possible) and such sensors are probably very expensive (for example laser trackers).

It is common to study the local dynamic behavior around certain operating points (also called positions in the paper) and there estimate parametric or nonparametric linear models (see, e.g., Albu-Schäffer and Hirzinger, 2001; Behi and Tesar, 1991; Johansson et al., 2000; Öhr et al., 2006). One application area for these linear models is control design, where a global (feedback) controller is achieved through gain scheduling. The linear models can also be used for the tuning of elastic parameters in a global nonlinear robot model, which is the adopted solution in this paper:

- The local behavior is considered by estimating the nonparametric frequency re-

sponse function (FRF) of the system in a number of positions.

- Next, the nonlinear parametric robot model is linearized in each of these positions.
- Finally, the parameters are optimized such that the parametric FRFs (the FRFs of the linearized parametric robot model) match the estimated nonparametric FRFs.

This identification procedure, first suggested in Öhr et al. (2006), will here be described in more detail. Various aspects of the procedure are also treated in Wernholt and Gunnarsson (2006a), Wernholt and Löfberg (2007), Wernholt and Gunnarsson (2007) and Wernholt and Moberg (2007). Using an FRF-based procedure allows for data compression, unstable systems are handled without problems, it is easy to validate the model such that all important resonances are captured, and model requirements in the frequency domain are also easily handled.

The proposed procedure also has some possible problems. The choice of model structure (parameters) and handling local minima in the optimization are problems here as well. In addition comes some difficulties with biased nonparametric FRF estimates due to closed-loop data and nonlinearities. There are also cases when even a small perturbation around a working point can give large variations due to the nonlinearities, which makes a linear approximation inaccurate, e.g., passing through Coulomb friction, backlash, or different parts of a nonlinear stiffness. This can be partly handled by the choice of excitation (e.g., avoid zero velocity to reduce Coulomb friction). Using multiple positions is good for the parameter accuracy as well as for identifiability issues. It will, however, make it harder to use a linear approximation of certain quantities. Consider, for example, the problem of nonlinear stiffness, where a linear approximation will vary between different positions due to gravity and the amplitude of the excitation. It is then impossible to find a linear stiffness that perfectly matches the resonances for all positions. Still, if the nonlinearity can be parameterized and properly linearized in the different positions, those parameters could possibly be identified as well.

The procedure will now be described, starting with the robot model in Section 3, carrying on by describing the nonparametric FRF estimation and the parameter estimation in Sections 4 and 5, respectively. Experimental results are shown in Section 6, and finally some conclusions are drawn in Section 7.

3 Robot Model

The robot model described in this section comes from Moberg and Hanssen (2007). A general serial link industrial robot, as in Figure 1, is then modeled by a kinematic chain of rigid bodies, where each rigid body is connected to the preceding body by three torsional spring-damper pairs, giving three degrees-of-freedom (DOF) to each rigid body. At most one of these DOFs can be actuated, corresponding to a connection of the two rigid bodies by a motor and a gearbox. In this representation, a robot link (always actuated) can consist of one or more rigid bodies. The model equations, described in Moberg and Hanssen (2007), can be written as a nonlinear gray-box model

$$\dot{x}(t) = f(x(t), u(t), \theta), \tag{1a}$$

$$y(t) = h(x(t), u(t), \theta), \tag{1b}$$

with state vector $x(t)$, input vector $u(t)$, output vector $y(t)$, and nonlinear functions $f(\cdot)$ and $h(\cdot)$ that describe the dynamics. The rigid body parameters are assumed to be known and θ is a vector of unknown parameters for (mainly) springs and dampers. See the previous section for examples of other unknown parameters to include.

4 FRF Estimation

As a first step toward the parameter identification, estimates of the nonparametric FRF in a number of positions are needed. These are obtained by performing experiments where the robot is moved into a position and a speed reference signal is fed to the robot controller. The resulting motor torques (actually the torque reference to the torque controller) and angular positions are sampled and stored. The measured angular positions are then filtered and differentiated to obtain estimates of the motor angular speeds, which are here considered as the output signals.

The open-loop system to be identified is unstable, which makes it necessary to collect data while the robot controller is running in closed loop. Consider therefore the setup in Figure 2, where the controller takes as input the difference between the reference signal r and the measured and sampled output y, and u is the input. The disturbance, v, contains various sources of noise and disturbances. An experimental control system is used, which enables the use of off-line computed reference signals for each motor controller.

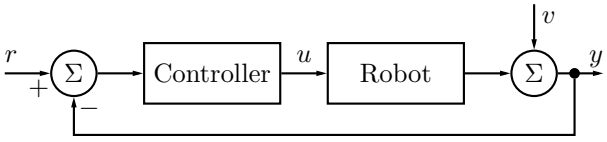


Figure 2: *Closed-loop measurement setup.*

To avoid leakage effects in the discrete Fourier transform (DFT), which is used by the estimation method, the excitation signal, r, is assumed to be periodic, with N_P samples in each period, and an integer number of periods of the steady-state response are collected. The nonparametric FRF estimate $\widehat{G}(\omega_k) \in \mathbb{C}^{n\times n}$ (assuming n inputs and outputs) is calculated from a block of n experiments like (Pintelon and Schoukens, 2001, p. 61)

$$\widehat{G}(\omega_k) = \mathbf{Y}(\omega_k)\mathbf{U}^{-1}(\omega_k), \tag{2}$$

where the n columns of $\mathbf{Y}(\omega_k)$ and $\mathbf{U}(\omega_k)$ contain the DFT of the sampled data from the n experiments. See also, e.g., Wernholt and Gunnarsson (2007) and Wernholt and Moberg (2007) for other FRF estimators for multivariable systems.

As excitation, an *orthogonal random phase multisine* signal (Dobrowiecki and Schoukens, 2007) is used, which here gives

$$\mathbf{R}(\omega_k) = \mathbf{R}_{\text{diag}}(\omega_k)\mathbf{T},$$

where $\mathbf{R}_{\text{diag}}(\omega_k)$ is a diagonal matrix

$$\mathbf{R}_{\text{diag}}(\omega_k) = \text{diag}\left\{R_1(\omega_k), \ldots, R_n(\omega_k)\right\},$$

and $\mathbf{T}$ is an orthogonal matrix

$$\mathbf{T}_{il} = e^{\frac{2\pi j}{n}(i-1)(l-1)},$$

with $\mathbf{T}\mathbf{T}^H = nI$. Each $R_l(\omega_k)$ is the DFT of a random phase multisine signal, which in the time domain can be written as

$$r(t) = \sum_{k=1}^{N_f} A_k \cos(\omega_k t + \phi_k), \tag{3}$$

with amplitudes A_k, frequencies ω_k chosen from the grid $\{\omega_k = \frac{2\pi k}{N_P T_s},\ k = 1, \ldots, \frac{N_p}{2} - 1\}$ (N_P even) with T_s the sampling period, and random phases ϕ_k uniformly distributed on the interval $[0, 2\pi)$. Using the orthogonal multisine signal in closed loop corresponds to an optimal experiment design given output amplitude constraints.

The selection of frequencies as well as the amplitude spectrum will affect the parameter estimation in the next step. Using too many frequencies will give a low signal-to-noise ratio, which increases both the bias and the variance in the nonparametric FRF estimate. The amplitude spectrum should also reflect the sensitivity for the unknown parameters (cf. $\Psi_0^{(i)}(k)$ in (5)), at least such that the unknown parameters influence the parametric FRF for the selected frequencies.

The nonparametric FRF estimate can be improved by averaging over multiple blocks and/or periods. The covariance matrix can then also be estimated. For a linear system, averaging over different periods is sufficient, whereas for a nonlinear system, it is essential to average over blocks where $\mathbf{R}_{\text{diag}}$ in each block should have different realizations of the random phases. The reason is that nonlinearities otherwise will distort the estimate and give a too low uncertainty estimate, see Pintelon and Schoukens (2001, Chap. 3) and Schoukens et al. (2005).

For the industrial robot, the nonlinearities cause large distortions and averaging over multiple blocks is therefore important (Wernholt and Gunnarsson, 2006a). For the same reason, one should only excite odd frequencies (only $\omega_k = 2\pi(2k+1)/(N_P T_s)$ in (3)), see Schoukens et al. (2005).

5 Parameter Estimation

When the FRFs have been estimated from data, the next step is to linearize the nonlinear model (1) in the same positions and calculate the parametric FRFs, $G^{(i)}(\omega_k, \theta)$, $i = 1, \ldots, Q$. A cost function $V(\theta)$ is then formed, measuring the (weighted) discrepancy between the parametric FRF and the estimated nonparametric FRF for all the Q positions. This cost function is finally minimized to identify the unknown parameters.

First, two different parameter estimators will be analyzed and compared. Next, the selection of optimal positions for the experiments is treated, and finally, the solution of the optimization problem is discussed.

Remark A.1. Note that the parametric FRF, $G^{(i)}(\omega_k, \theta)$, is a function of the nonlinear gray-box model (1) such that for each parameter vector θ during the minimization, (1) is linearized in $(x_0^{(i)}, u_0^{(i)})$ before calculating the FRF.

5.1 Estimators

Weighted Nonlinear Least Squares (NLS) Estimator

The NLS estimator is given by

$$\hat{\theta}_{N_f}^{\text{NLS}} = \arg\min_{\theta\in\Theta} V_{N_f}^{\text{NLS}}(\theta), \tag{4a}$$

$$V_{N_f}^{\text{NLS}}(\theta) = \sum_{i=1}^{Q}\sum_{k=1}^{N_f} [\mathcal{E}^{(i)}(k,\theta)]^H [\Lambda^{(i)}(k)]^{-1} \mathcal{E}^{(i)}(k,\theta), \tag{4b}$$

$$\mathcal{E}^{(i)}(k,\theta) = \text{vec}(\widehat{G}^{(i)}(\omega_k)) - \text{vec}(G^{(i)}(\omega_k,\theta)), \tag{4c}$$

with $\Lambda^{(i)}(k)$ a Hermitian ($\Lambda = \Lambda^H$) weighting matrix, and $(\cdot)^H$ denoting complex conjugate transpose. The asymptotic properties ($N_f \to \infty$) of this estimator will be derived in the following theorem.

Theorem A.1
Consider the NLS estimator (4) *and assume that:*

1. $\widehat{G}^{(i)}(\omega_k) = G^{(i)}(\omega_k,\theta_0) + \eta^{(i)}(\omega_k)$ *with* $\text{vec}(\eta^{(i)}(\omega_k))$ *a zero mean circular complex random vector, independent over* i *and* ω_k*, with covariance matrix* $\Lambda_0^{(i)}(\omega_k)$*.*
2. Θ *is a compact set where* $V_{N_f}^{NLS}(\theta)$ *and its first- and second-order derivatives are continuous for any value of* N_f*.*
3. *For* N_f *large enough, the expected value of* $V_{N_f}^{NLS}(\theta)$ *has a unique global minimum in* Θ*.*

The estimator $\hat{\theta}_{N_f}^{NLS}$ *will then converge to* θ_0 *as* $N_f \to \infty$ *and* $\sqrt{N_f}(\hat{\theta}_{N_f}^{NLS} - \theta_0)$ *is asymptotically Normal distributed with covariance matrix* P_θ*,*

$$\begin{aligned} P_\theta = \frac{1}{2}\Bigg[\frac{1}{N_f}\sum_{i=1}^{Q}\sum_{k=1}^{N_f} \Re\left\{\overline{\Psi_0^{(i)}(k)}\Xi^{(i)}(k)[\Psi_0^{(i)}(k)]^T\right\}\Bigg]^{-1} \\ \times\Bigg[\frac{1}{N_f}\sum_{i=1}^{Q}\sum_{k=1}^{N_f} \Re\left\{\overline{\Psi_0^{(i)}(k)}\Sigma^{(i)}(k)[\Psi_0^{(i)}(k)]^T\right\}\Bigg] \\ \times\Bigg[\frac{1}{N_f}\sum_{i=1}^{Q}\sum_{k=1}^{N_f} \Re\left\{\overline{\Psi_0^{(i)}(k)}\Xi^{(i)}(k)[\Psi_0^{(i)}(k)]^T\right\}\Bigg]^{-1}, \end{aligned} \tag{5}$$

with the Jacobian matrix $[\Psi_0^{(i)}(k)]^T = \left.\frac{\partial\,\text{vec}(G^{(i)}(\omega_k,\theta))}{\partial\theta}\right|_{\theta=\theta_0}$*,* $\overline{(\cdot)}$ *denoting complex conjugate, and*

$$\begin{aligned} \Xi^{(i)}(k) &= [\Lambda^{(i)}(\omega_k)]^{-1}, \\ \Sigma^{(i)}(k) &= [\Lambda^{(i)}(\omega_k)]^{-1}\Lambda_0^{(i)}(\omega_k)[\Lambda^{(i)}(\omega_k)]^{-1}. \end{aligned}$$

The covariance is minimized by using the optimal weights

$$\Lambda^{(i)}(\omega_k) = \Lambda_0^{(i)}(\omega_k), \tag{6}$$

which also simplifies (5) *to*

$$P_\theta = \frac{1}{2}\left[\frac{1}{N_f}\sum_{i=1}^{Q}\sum_{k=1}^{N_f} \Re\left\{\overline{\Psi_0^{(i)}(k)}[\Lambda^{(i)}(\omega_k)]^{-1}[\Psi_0^{(i)}(k)]^T\right\}\right]^{-1}.$$

Proof: Follows from fairly straightforward calculations using Theorem 7.21 in Pintelon and Schoukens (2001). □

Note that in addition to the mentioned assumptions, there are some technical details for the asymptotic normality that $\eta^{(i)}(\omega_k)$ has uniformly bounded absolute moments of order $4+\epsilon$ with $\epsilon > 0$. See Pintelon and Schoukens (2001, Theorem 7.21) for details.

Weighted Logarithmic Least Squares (LLS) Estimator

For systems with a large dynamic range, the NLS estimator may become ill-conditioned. The weighted logarithmic least squares (LLS) estimator has been suggested as an alternative (Pintelon and Schoukens, 2001, pp. 206–207)

$$\hat{\theta}_{N_f}^{\text{LLS}} = \arg\min_\theta V_{N_f}^{\text{LLS}}(\theta), \tag{7a}$$

$$V_{N_f}^{\text{LLS}}(\theta) = \sum_{i=1}^{Q}\sum_{k=1}^{N_f}[\mathcal{E}^{(i)}(k,\theta)]^H[\Lambda^{(i)}(k)]^{-1}\mathcal{E}^{(i)}(k,\theta), \tag{7b}$$

$$\mathcal{E}^{(i)}(k,\theta) = \log \operatorname{vec}(\widehat{G}^{(i)}(\omega_k)) - \log \operatorname{vec}(G^{(i)}(\omega_k,\theta)), \tag{7c}$$

where $\log G = \log|G| + j\arg G$. This estimator has improved numerical stability and is particularly robust to outliers in the measurements. However, from a theoretical point of view, the estimator is inconsistent ($\lim_{N_f\to\infty}\hat{\theta}_{N_f}^{\text{LLS}} \neq \theta_0$). The bias can be neglected if the signal-to-noise ratio ($\operatorname{vec}(\widehat{G})$ vs. $\sqrt{\operatorname{diag}\{\Lambda_0\}}$) is large enough (at least 10 dB according to Pintelon and Schoukens, 2001, p. 207).

Similarly to Theorem A.1, one can show that the covariance matrix, using the LLS estimator (7), is approximately given by (5) with

$$\Xi^{(i)}(k) = \left[G_d^{(i)}(\omega_k,\theta_0)\Lambda^{(i)}(\omega_k)[G_d^{(i)}(\omega_k,\theta_0)]^H\right]^{-1},$$
$$\Sigma^{(i)}(k) = \Xi^{(i)}(k)\Lambda_0^{(i)}(\omega_k)\Xi^{(i)}(k),$$

and $G_d^{(i)}(\omega_k,\theta_0) = \operatorname{diag}\left\{\operatorname{vec}(G^{(i)}(\omega_k,\theta_0))\right\}$. Using the optimal weights

$$\Lambda^{(i)}(\omega_k) = \left[G_d^{(i)}(\omega_k,\theta_0)\right]^{-1}\Lambda_0^{(i)}(\omega_k)\left[G_d^{(i)}(\omega_k,\theta_0)\right]^{-H}, \tag{8}$$

gives approximately the same covariance as for the NLS estimator.

Selection of Weights

Even if the covariance is minimized by using the optimal weights, the choice of weights will in general deviate from the optimal ones for a number of reasons. Firstly, the true covariance matrix $\Lambda_0^{(i)}(\omega_k)$ is usually not known so the user must instead be content with an estimated covariance matrix $\hat{\Lambda}_0^{(i)}(\omega_k)$. Secondly, the weights also reflect where the user requires the best model fit. This is important in case the model is unable to describe every detail in the measurements. The bias-inclination will then be small for frequencies, elements, and positions where the weights $[\Lambda^{(i)}(\omega_k)]^{-1}$ are large.

For a resonant system, it is often easier to use the LLS estimator in the way that even constant weights will make sure that both resonances and anti-resonances are matched by the model. This is due to the fact that the logarithm in the LLS estimator inherently gives the relative error, compared to the absolute error when using the NLS estimator. With the NLS estimator, the anti-resonances are easily missed if not choosing large weights at those frequencies.

5.2 Optimal Positions

Given a nonlinear gray-box model (1), the information about the unknown parameters will differ between nonparametric FRF estimates in different positions. Therefore, given a limited total measurement time, one should perform experiments in the position(s) that contribute the most to the information about the unknown parameters. In Wernholt and Löfberg (2007), this problem is formulated as follows: Assume a set of Q_c candidate positions. Determine the number of experiments to be performed in each position (m_i experiments in position i) such that the parameter uncertainty is minimized, given a total of $M = \sum_{i=1}^{Q_c} m_i$ experiments. Determining the values m_i, $i = 1, \ldots, Q_c$, is a combinatorial experiment design problem which relatively quickly will become intractable when Q_c is large. If M is not too small, a good approximate solution can be found by relaxing the constraint that each m_i should be an integer. This relaxed problem is convex, which enables the global optimum to be found. In the paper Wernholt and Löfberg (2007), it is also shown that the experiment design is efficiently solved by considering the dual problem. The candidate positions are obtained by griding the workspace. Given thousands of candidate positions, only a few positions typically have a nonzero m_i in the optimum. See Wernholt and Löfberg (2007) for details and examples.

5.3 Solving the Optimization Problem

The minimization problem to be solved, (4) or (7), is unfortunately non-convex. Here, the problem is solved using *fminunc* in MATLAB, which is a gradient-based method which only returns a local optimum. Due to the existence of local minima, a good initial parameter vector, θ_{init}, is important. The problem is solved for a number of random perturbations around θ_{init} in order to avoid local minima. Or, alternatively stated, to obtain a local minimum which is good enough for the purpose of the model. The quality of the resulting model, as well as problems with local minima and identifiability properties, depend on the choices of estimator, weights, and position(s) for the experiments. This will be illustrated in the next section.

6 Experimental Results

The described identification procedure from the previous sections will here be used for the identification of an industrial robot from the ABB IRB6600 series using an experimental controller. A nonlinear gray-box model with 26 unknown parameters is used. The nonparametric FRFs are estimated in the 15 optimal positions from Wernholt and Löfberg (2007) by using an odd orthogonal random phase multisine signal with a flat amplitude spectrum as excitation and averaging over a number of blocks. The parameters are then estimated using the following estimators:

`LLSM15U`: LLS estimator, $Q = 15$, only magnitude ($\log|G|$), user-defined weights.

`LLS15U`: LLS estimator, $Q = 15$, user-defined weights.

`LLS15O`: LLS estimator, $Q = 15$, optimal weights.

`NLS15U`: NLS estimator, $Q = 15$, user-defined weights.

`NLS15O`: NLS estimator, $Q = 15$, optimal weights.

`LLS1U`: LLS estimator, $Q = 1$, user-defined weights.

For simplicity only diagonal weights $[\Lambda^{(i)}(\omega_k)]^{-1}$ are considered[1]. The user-defined weights are constant for each element in the FRF, the same for all positions, zero for low frequencies where the nonparametric FRF is uncertain, and lower for the non-diagonal elements in the FRF. The optimal weights are calculated from (6) and (8), using the estimated covariance $\hat{\Lambda}_0^{(i)}(\omega_k)$ and $G^{(i)}(\omega_k, \theta_0) \approx \widehat{G}^{(i)}(\omega_k)$. The optimal weights often turn out to be small at the resonances and anti-resonances due to a larger relative error in the nonparametric FRF estimate at those frequencies. For the `LLS1U` estimator, the single position with the smallest theoretical parameter covariance is used.

To assess the sensitivity to the initial parameter vector, θ_{init}, 100 optimizations are performed for each of the 6 estimators, using randomly perturbed initial parameters, $\theta_{init}^{[l]},\ l = 1, \ldots, 100$ (the same for all estimators). Each element in $\theta_{init}^{[l]}$ is obtained by multiplying the corresponding element in θ_{init} by 10^{φ}, where φ is a random number from a uniform distribution on the interval $[-1,\ 1]$.

To evaluate the resulting 600 models, the same cost function is used for all models. The LLS cost, $V^{\text{LLS}}(\theta)$, is calculated with optimal weights and user-defined weights, which can be seen in Figures 3 and 4, respectively. The cost varies quite much between the different estimators. What is more important is the trend over the different optimizations. The first three estimators tend to be much more robust to varying initial parameters.

To compare the number of reasonable models, some measure is needed. Since the optimal weights are small at the resonances and anti-resonances, Figure 3 is not so well suited for judging if resonances (and anti-resonances) are accurately modeled or not. Consider therefore Figure 4. When comparing the FRFs of the parametric models with the estimated nonparametric FRFs, the models usually miss important resonances when the

[1]This means that, e.g., (4b) can be rewritten as
$V_{N_f}^{\text{NLS}}(\theta) = \sum_{i=1}^{Q}\sum_{k=1}^{N_f}\sum_{m=1}^{n_y}\sum_{n=1}^{n_u} |\widehat{G}_{mn}^{(i)}(\omega_k) - G_{mn}^{(i)}(\omega_k, \theta)|^2 W_{m+(n-1)n_y}^{(i)}(\omega_k)$,
where $W^{(i)}(\omega_k)$ is the diagonal of $[\Lambda^{(i)}(\omega_k)]^{-1}$.

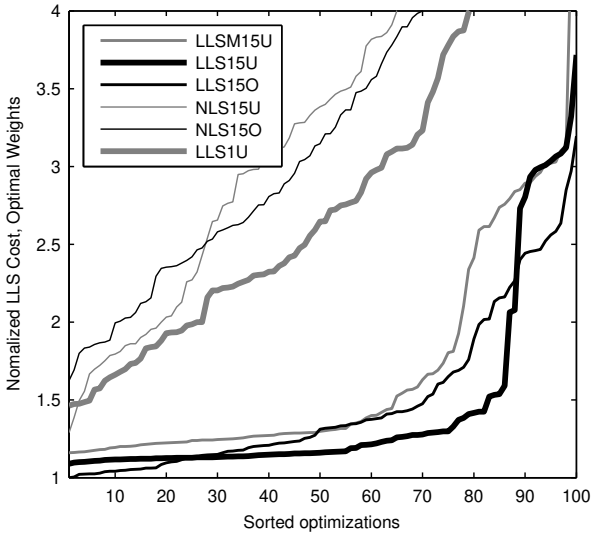


Figure 3: *Normalized LLS cost with optimal weights for all initial parameters and all estimators.*

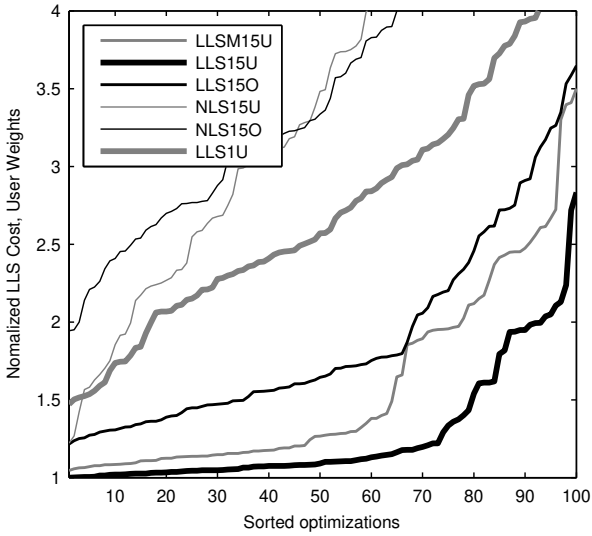


Figure 4: *Normalized LLS cost with user-defined weights for all initial parameters and all estimators.*

Table 1: *Statistics for the* `LLS15U` *estimator, where the first five columns are obtained using the* 31 *best models in Figure 4, and the last column is calculated for the best model, using* (5).

θ_i	min	median	max	mean	std	std_{th}
k_1	0.97	1.01	1.05	1.01	0.020	0.0069
k_2	0.99	1.02	1.04	1.02	0.012	0.0081
k_3	0.061	1.00	3.67	1.10	0.66	0.103
k_4	0.72	0.99	1.26	0.97	0.13	0.042
c_1	0.84	0.99	1.45	1.00	0.12	0.037
c_2	0.24	1.04	1.51	0.99	0.30	0.056
c_3	0.16	0.55	7.05	1.57	1.93	0.865
c_4	0.048	0.44	34.1	2.27	6.14	0.186

normalized cost in Figure 4 exceeds approximately 1.5. That gives the following percentage of reasonable models (out of the 100 models): `LLS15U`, 80 %, `LLSM15U`, 64 %, `LLS15O`, 35 %, `NLS15U`, 3 %, `LLS1U`, 2 %, and `NLS15O`, 0 %. These numbers are only approximate since the same cost can be achieved if one resonance is missed completely but all others are accurate, and if many resonances are only modeled with moderate accuracy. The latter is often the case for the `LLS15O` estimator since the exact location of the resonances, as well as their damping, are not so important when using the optimal weights.

To further evaluate the estimated models, the parameter variation is studied. The spring parameters of the gearboxes and the arm structure as a function of the sorted optimizations can be seen in Figures 5 and 6, respectively. These parameters are normalized by the best `LLS15U` model. One immediately notes that the arm structure springs are harder to estimate, in particular one of them. This parameter does not influence the FRF that much in the selected frequency interval and is therefore hard to estimate. Considering the number of optimizations with estimated parameters inside the interval $[0.5,\ 2]$ (black lines in the figures) gives the following percentage of reasonable models (out of the 100 models): `LLS15U`, 62 %, `LLSM15U`, 43 %, `LLS15O`, 21 %, `NLS15U`, 2 %, `LLS1U`, 2 %, and `NLS15O`, 0 %. One arm structure parameter is excluded in these numbers, but the `LLS15U` estimator actually manages to accurately estimate the 12 spring parameters in 12 % of the optimizations. The dampers are unfortunately much harder to accurately estimate, as can be seen in Figure 7. Some of the damping parameters fluctuate quite much even among the best models.

The `LLS15U` estimator is further analyzed by computing the theoretical parameter uncertainty from (5) for the model with the lowest V^{LLS} cost, as well as the statistics for the best 31 models, i.e., all models with a cost less than 1.05 in Figure 4. Statistics for 8 representative springs k_i and dampers d_i are shown in Table 1, where the parameters with both the smallest and the largest uncertainties are included. The conclusions are that the springs are more accurately estimated than the dampers and that the theoretical uncertainty gives a good indication of the quality of the estimated parameters.

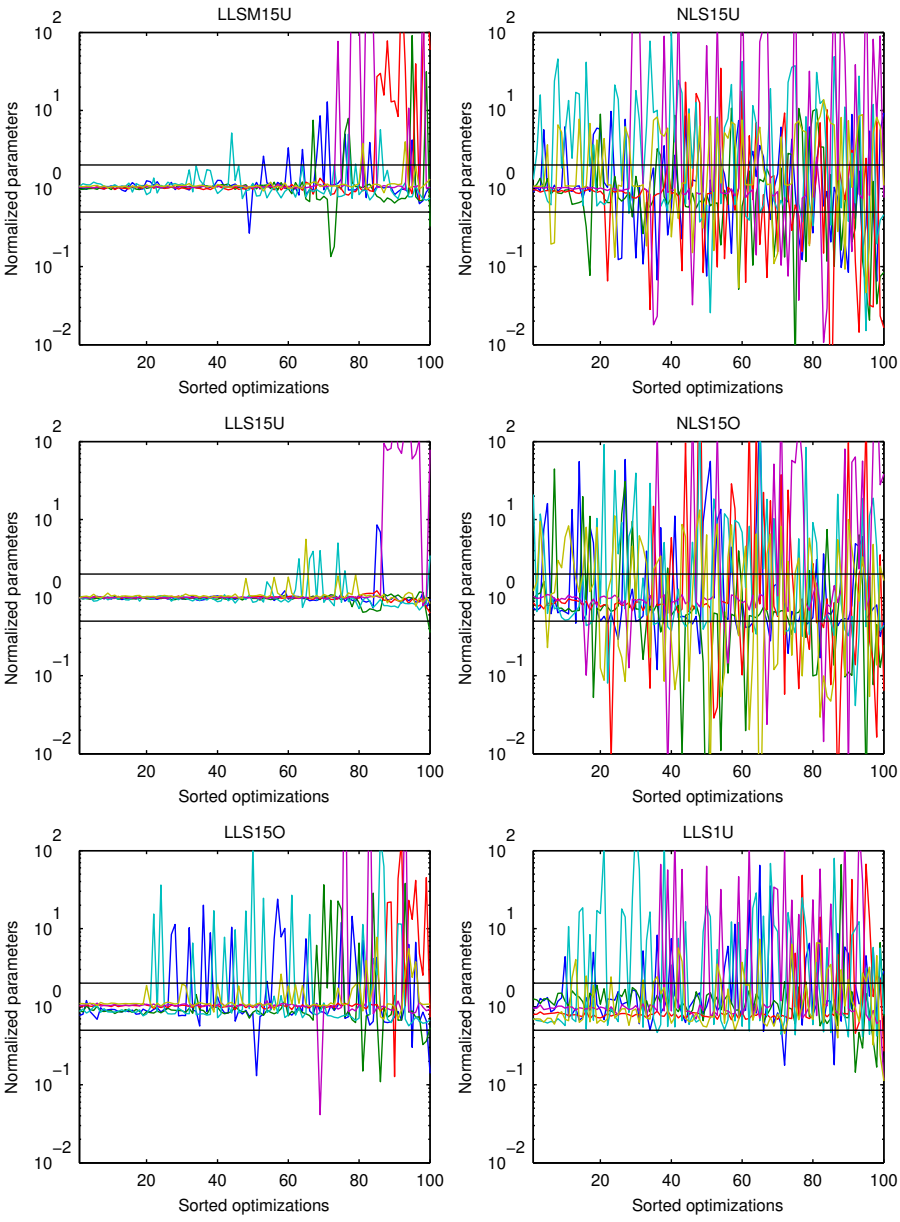


Figure 5: *Normalized spring parameters of the gearboxes for the 6 different estimators, sorted according to Figure 4.*

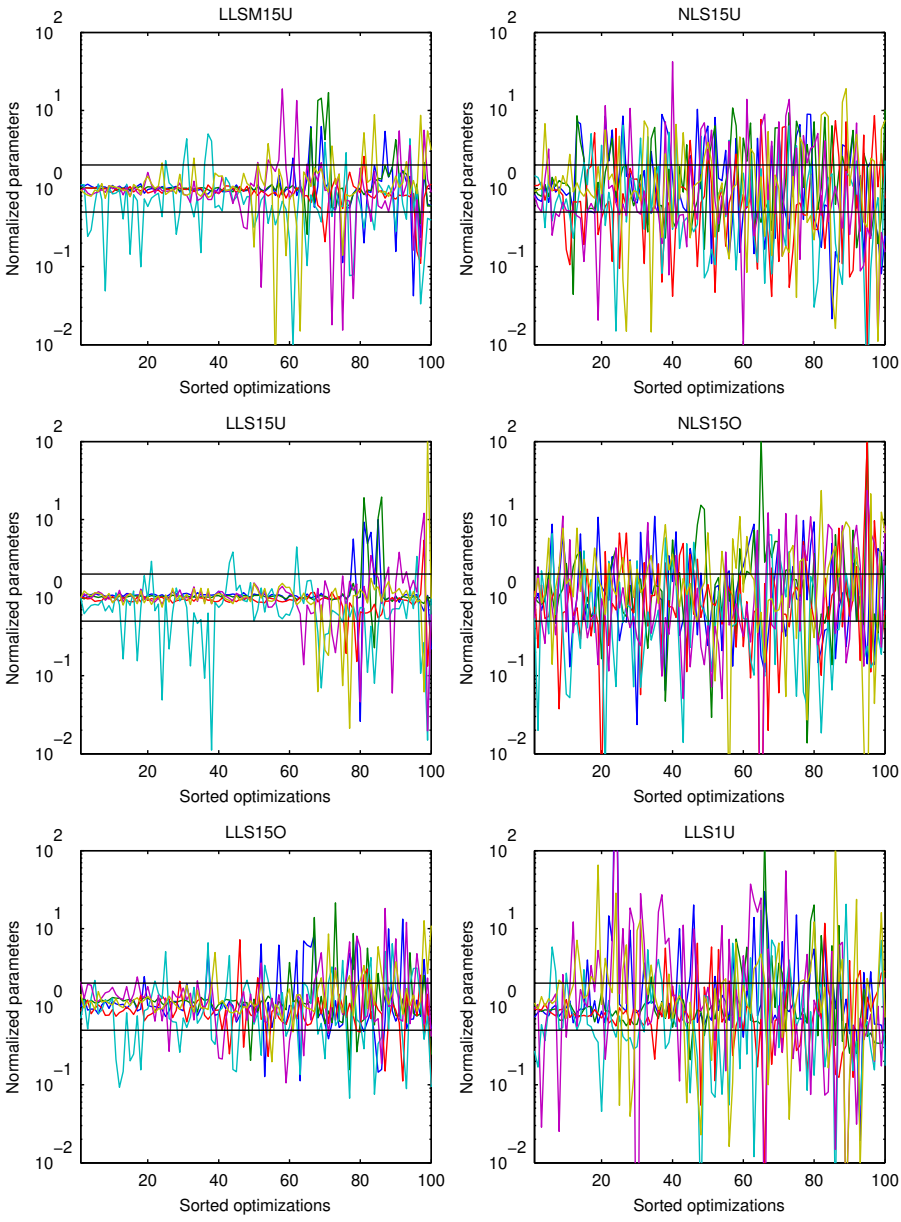


Figure 6: *Normalized spring parameters of the arm structure for the 6 different estimators, sorted according to Figure 4.*

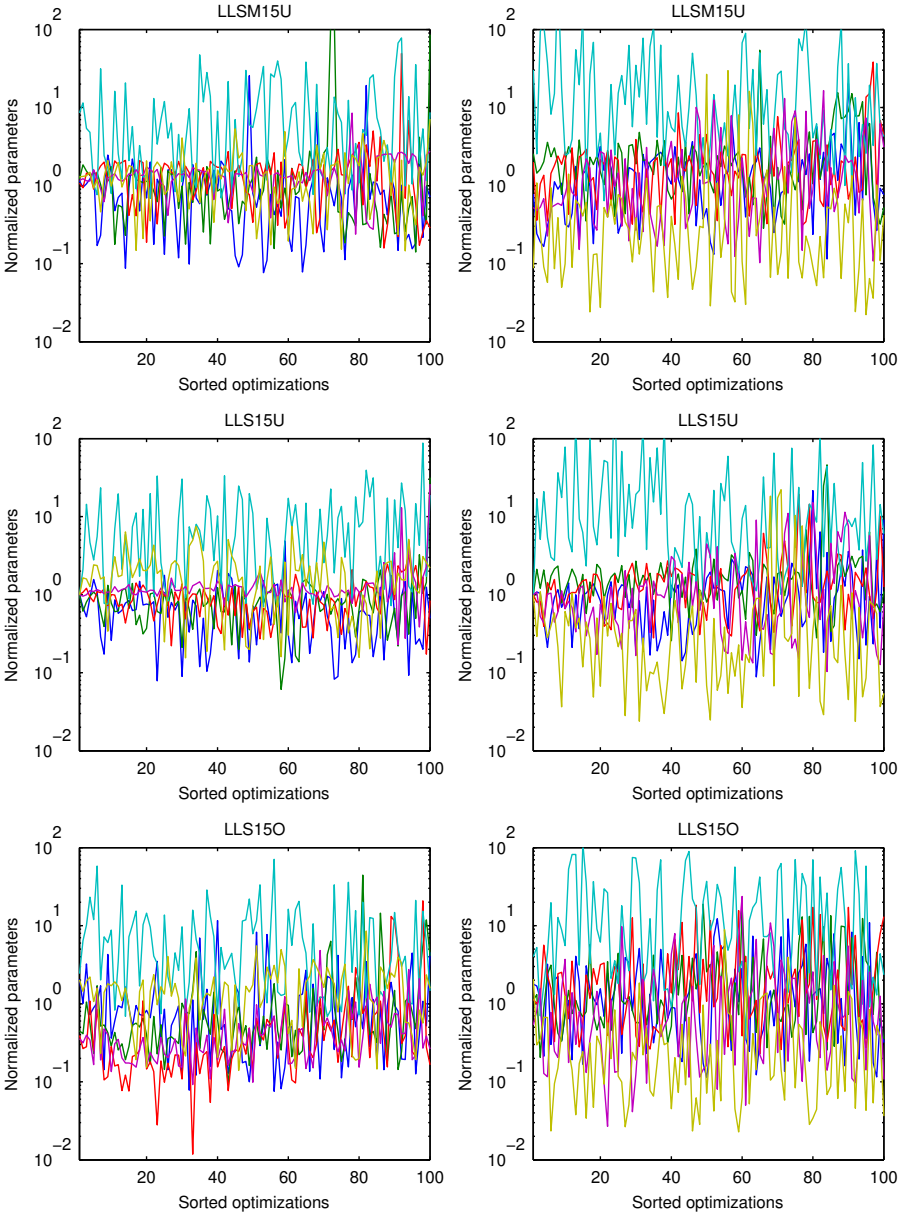


Figure 7: *Normalized damping parameters of the gearboxes (left column) and the arm structure (right column) for the estimators* `LLSM15U`, `LLS15U` *and* `LLS15O`, *sorted according to Figure 4.*

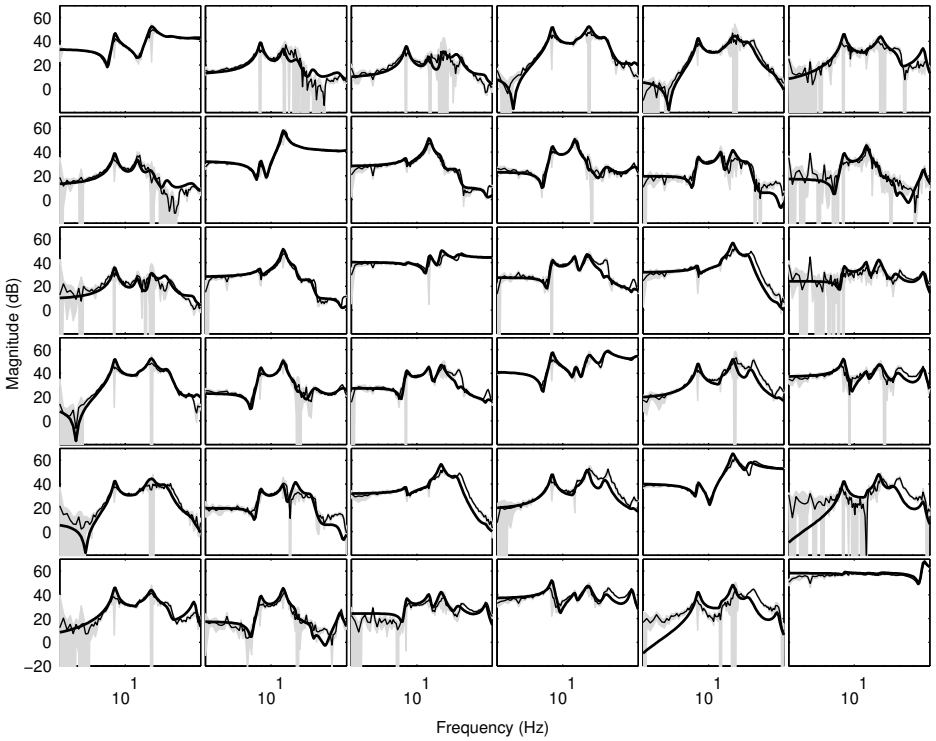


Figure 8: *Estimated nonparametric FRF $\widehat{G}^{(i)}(\omega_k)$ and parametric FRF $G^{(i)}(\omega_k, \theta)$ (the best* `LLS15U` *model) in one of the positions. Input: 6 motor torques, output: 6 motor accelerations. Thin line: $\widehat{G}^{(i)}(\omega_k)$, shaded: one standard deviation for $\widehat{G}^{(i)}(\omega_k)$, and thick line: $G^{(i)}(\omega_k, \theta)$.*

Figure 8 finally shows the magnitude of the estimated nonparametric FRF and the best parametric model for one of the positions. The identified model gives a good global description of the dynamics in the frequency range of interest.

7 Concluding Discussion

This paper has dealt with the problem of estimating unknown elasticity parameters in a nonlinear gray-box model of an industrial robot. An identification procedure has been proposed where the parameters are identified in the frequency domain, using estimates of the nonparametric FRFs for a number of robot configurations/positions. The nonlinear parametric gray-box model is linearized in the same positions and the optimal parameters are obtained by minimizing the discrepancy between the nonparametric and the parametric FRFs. Two different parameter estimators (NLS and LLS) have been analyzed. The estimators, as well as the selection of weights in the estimators, have been evaluated in an

experimental study. The conclusions of the experimental study are that:

- the LLS estimator is superior to the NLS estimator for this type of system,
- more than one position is needed in order to get a reasonable estimate,
- using phase information improves the estimate,
- rough user-defined weights work much better than the theoretically optimal weights,
- gearbox parameters are easier to estimate than arm structure parameters,
- spring parameters are easier to identify than damping parameters, and
- the theoretical uncertainties for the estimated parameters in Table 1 give a good indication of the quality of the estimated parameters.

The uncertainties in the dampers and in some of the springs in the best identified model are quite large, but the resulting global model should anyway be useful for many purposes.

An explanation to the fourth point is that the assumptions in Theorem A.1 are violated since the nonparametric FRF estimate has bias errors due to nonlinearities and closed-loop data, and the model is unable to describe every detail of the true system. The weights should, for such a case, primarily be selected to distribute the bias, not to get minimum variance. The theoretically optimal weights are low at the resonance and anti-resonance frequencies (due to uncertainties in the FRF), which in turn give large model errors there.

The fifth point comes as no surprise. The model structure, where all elastic effects in the arm structure are lumped into a few spring-damper pairs, can of course be modified and refined. Both regarding the location of these spring-damper pairs, as well as how many that are needed in order to properly model the system. Identifiability of these added parameters can also be discussed. Maybe additional sensors are needed, e.g., accelerometers attached to the structure, as is the case in experimental modal analysis (Behi and Tesar, 1991; Verboven, 2002).

The main reasons for the large uncertainties in the damping parameters probably are that the system is poorly damped and that the nonparametric FRFs contain errors at the resonances and anti-resonances such that unique damping parameters are hard to find.

A number of areas are still subject to future work. The selection of weights can certainly be improved by combining the user choices and the estimated FRF uncertainty. The selection of frequencies, as well as the amplitude spectrum for the nonparametric FRF estimation can be further improved. An experimental verification of the optimal positions for identification is still interesting to perform. The parameter accuracy and problems with local minima, versus measurement time and excitation energy are also interesting problems to study. Using a frequency-domain method for identification of a nonlinear system has some problems, as was pointed out in Section 2. Therefore, it would be interesting to apply time-domain prediction error methods as a comparison, even though that involves a number of hard problems to tackle, also mentioned in Section 2. A simulation-based study, using a realistic nonlinear model, could also be enlightening.

Finally, to conclude this paper: Identification of industrial robots is a challenging task. Using a general purpose method by pressing a button will almost surely fail. The problem instead requires a combination of tailored identification methods, experiment design, and a skilled user, using all available knowledge about the system.

References

Albu-Schäffer, A. and Hirzinger, G. (2001). Parameter identification and passivity based joint control for a 7DOF torque controlled light weight robot. In *Proc. 2001 IEEE International Conference on Robotics and Automation*, pages 2852–2858, Seoul, Korea.

Albu-Schäffer, A. and Hirzinger, G. (2000). State feedback controller for flexible joint robots: A globally stable approach implemented on DLR's light-weight robots. In *Proc. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1087–1093, Takamatsu, Japan.

An, C. H., Atkeson, C. G., and Hollerbach, J. M. (1988). *Model-Based Control of a Robot Manipulator*. MIT Press, Cambridge, Massachusetts.

Behi, F. and Tesar, D. (1991). Parametric identification for industrial manipulators using experimental modal analysis. *IEEE Transactions on Robotics and Automation*, 7(5):642–652.

Dobrowiecki, T. and Schoukens, J. (2007). Measuring a linear approximation to weakly nonlinear MIMO systems. *Automatica*, 43(10):1737–1751.

Johansson, R., Robertsson, A., Nilsson, K., and Verhaegen, M. (2000). State-space system identification of robot manipulator dynamics. *Mechatronics*, 10(3):403–418.

Khalil, W. and Gautier, M. (2000). Modeling of mechanical systems with lumped elasticity. In *Proc. 2000 IEEE International Conference on Robotics and Automation*, pages 3964–3969, San Francisco, California.

Kozlowski, K. (1998). *Modelling and identification in robotics*. Advances in Industrial Control. Springer, London.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Moberg, S. and Hanssen, S. (2007). A DAE approach to feedforward control of flexible manipulators. In *Proc. 2007 IEEE International Conference on Robotics and Automation*, pages 3439–3444, Roma, Italy.

Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach*. IEEE Press, New York.

Schoukens, J., Pintelon, R., Dobrowiecki, T., and Rolain, Y. (2005). Identification of linear systems with nonlinear distortions. *Automatica*, 41(3):491–504.

Spong, M. W. (1987). Modeling and control of elastic joint robots. *Journal of Dynamic Systems Measurement and Control*, 109(4):310–319.

Verboven, P. (2002). *Frequency-domain system identification for modal analysis*. PhD thesis, Vrije Universiteit Brussel, Belgium.

Wernholt, E. and Gunnarsson, S. (2006a). Detection and estimation of nonlinear distortions in industrial robots. In *Proc. 23rd IEEE Instumentation and Measurement Technology Conference*, pages 1913–1918, Sorrento, Italy.

Wernholt, E. and Gunnarsson, S. (2006b). Nonlinear identification of a physically parameterized robot model. In *Proc. 14th IFAC Symposium on System Identification*, pages 143–148, Newcastle, Australia.

Wernholt, E. and Gunnarsson, S. (2007). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Wernholt, E. and Löfberg, J. (2007). Experiment design for identification of nonlinear gray-box models with application to industrial robots. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Wernholt, E. and Moberg, S. (2007). Experimental comparison of methods for multivariable frequency response function estimation. Technical Report LiTH-ISY-R-2827, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Paper B

Analysis of Methods for Multivariable Frequency Response Function Estimation in Closed Loop

Edited version of the paper:

> Wernholt, E. and Gunnarsson, S. (2007a). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

An early version of the paper is published in:

> Wernholt, E. and Gunnarsson, S. (2004). On the use of a multivariable frequency response estimation method for closed loop identification. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 827–832, Nassau, Bahamas.

Preliminary version published as Technical Report LiTH-ISY-R-2775, Dept. of Electrical Engineering, Linköping University, SE–581 83 Linköping, Sweden.

Analysis of Methods for Multivariable Frequency Response Function Estimation in Closed Loop

Erik Wernholt and Svante Gunnarsson

Dept. of Electrical Engineering,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: {erikw,svante}@isy.liu.se.

Abstract

Nonparametric estimation methods for the multivariable frequency response function are analyzed, both in open and closed loop. Expressions for the bias and covariance are derived and the usefulness of these expressions is illustrated in simulations of an industrial robot where the different estimators are compared. The choice of estimator depends on the signal-to-noise ratio as well as the measurement setup and a bias-variance trade-off.

1 Introduction

This paper studies the properties of some nonparametric frequency response function (FRF) estimation methods for multivariable systems, both for open-loop and closed-loop data. Nonparametric estimates of the FRF give valuable information about the dynamics of a system and are often used as an intermediate step in a parametric identification process to assess the quality of the measurements and visualize the complexity of the modeling problem. In a second step, a parametric model can be estimated, either by,

1. treating the estimated FRF as a measurement and minimize the distance between the model and the estimated FRF as is done in, e.g., experimental modal analysis (Verboven, 2002), or,

2. directly from the measured input and output data (Ljung, 1999; Pintelon and Schoukens, 2001).

The paper mainly deals with the H_1 *estimator*, but three other estimators will be treated as well in various degrees: the *arithmetic mean (ARI) estimator*, the *joint input-output (JIO) estimator*, and the *errors-in-variables (EIV) estimator*. These estimators will

be described in more detail in Section 2, see also Wellstead (1981), Cobb and Mitchell (1990), Pintelon and Schoukens (2001), Verboven (2002), Guillaume et al. (1996, 1992) and Guillaume (1998).

Previous studies of the performance of the EIV and H_1 estimators for multivariable systems are presented in, e.g., Guillaume et al. (1996), Verboven (2002), and Dobrowiecki et al. (2005) for open-loop data and Wernholt and Gunnarsson (2004) contains some early work for closed-loop data. The open-loop covariance expressions from the two papers Guillaume et al. (1996) and Dobrowiecki et al. (2005) will here be combined and covariance expressions for the H_1, ARI, and EIV estimators are presented. For closed-loop data, an asymptotic expression for the bias of the H_1 estimator is derived together with asymptotic covariance expressions for the JIO and EIV estimators, as well as approximate covariance expressions for the H_1 and ARI estimators.

For a numerical illustration, a simulation model of an industrial robot will be used. The robot application is interesting since it gives many challenging problems for system identification methods, such as a multivariable nonlinear system, oscillatory behavior, and data collection under feedback control. An overview of identification in robotics can be found in Kozlowski (1998). See also Öhr et al. (2006), Berglund and Hovland (2000), and Khorrami et al. (1995) for examples where nonparametric FRF estimates are used for the identification of parametric robot models. The simulation results are evaluated using the derived bias and variance expressions. Experimental results from real robots are presented in, e.g., Wernholt (2004) and Wernholt and Gunnarsson (2006).

The paper is organized as follows. Section 2 presents the nonparametric estimation methods that will be used in the paper. The estimation methods are analyzed in Section 3 for open-loop data and Section 4 for closed-loop data. The estimators and the derived error expressions are illustrated in a numerical example in Section 5. Finally, Section 6 contains some conclusions.

2 FRF Estimation Methods

The setup considered in this paper is given by

$$y(t) = G(q)u(t) + v(t), \tag{1}$$

where $G(q)$ is the $n_y \times n_u$ multivariable discrete-time transfer operator, with q being the shift operator, and $v(t)$ is the measurement noise. The input and output signals, $u(t) \in \mathbb{R}^{n_u}$ and $y(t) \in \mathbb{R}^{n_y}$, are measured at time instants $t_n = nT_s,\ n = 1, 2, \ldots, N$, with sampling period T_s.

To avoid leakage effects in the discrete Fourier transform (DFT), which is used by the identification methods, the input signal, $u(t)$, is assumed to be N_P-periodic ($u(t + N_P T_s) = u(t)$) and an integer number of periods, P, of the steady-state response is collected, giving $N = PN_P$ samples for each experiment. Consider now the DFTs of the

input and output signals

$$U(\omega_k) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} u(nT_s) e^{j\omega_k n T_s},$$
$$Y(\omega_k) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} y(nT_s) e^{j\omega_k n T_s},$$

where only the N_P frequencies $\omega_k = k\frac{2\pi}{N_P T_s},\ k = 1, 2, \ldots, N_P$ are considered. Given periodic data, the following linear mapping holds exactly in the noise-free case

$$Y(\omega_k) = G(e^{j\omega_k T_s}) U(\omega_k),$$

where $G(e^{j\omega_k T_s}) \in \mathbb{C}^{n_y \times n_u}$ is the FRF. To be able to extract $G(e^{j\omega_k T_s})$ from data, at least n_u different experiments are needed. The data vectors from $n_e \geq n_u$ different experiments will be collected into matrices (bold-face in the sequel) where each column corresponds to one experiment. The input-output relation can then be written as

$$\mathbf{Y}(\omega_k) = G(e^{j\omega_k T_s}) \mathbf{U}(\omega_k), \tag{2}$$

where $\mathbf{U}(\omega_k) \in \mathbb{C}^{n_u \times n_e}$ and $\mathbf{Y}(\omega_k) \in \mathbb{C}^{n_y \times n_e}$. If $\mathbf{U}(\omega_k)$ has rank n_u, a nonparametric estimate of $G(e^{j\omega_k T_s})$ can be formed by, e.g., using the H_1 *estimator* (Guillaume et al., 1996; Pintelon and Schoukens, 2001; Verboven, 2002)

$$\widehat{G}^{\mathrm{H}_1}(e^{j\omega_k T_s}) = \mathbf{Y}(\omega_k)\mathbf{U}^H(\omega_k)[\mathbf{U}(\omega_k)\mathbf{U}^H(\omega_k)]^{-1}, \tag{3}$$

where $(\cdot)^H$ denotes complex conjugate transpose. When the measurements are corrupted by noise, (2) changes to

$$\mathbf{Y}(\omega_k) = G(e^{j\omega_k T_s}) \mathbf{U}(\omega_k) + \mathbf{V}(\omega_k). \tag{4}$$

The H_1 estimator (3) can still be used, but the estimate will contain errors due to the noise.

In general the choice of excitation signal offers a large freedom in terms of frequency contents, magnitude, and so on, as long as the matrix $\mathbf{U}(\omega_k)$ has rank n_u. In this paper, the *orthogonal random phase multisine* signal will be used. This signal has been suggested in Dobrowiecki et al. (2005) and Dobrowiecki and Schoukens (2005) to minimize the variance ($\det \sigma^2_{\widehat{G}}$, cf. (10)) given certain amplitude constraints for the input signal. Assuming $n_e = M \cdot n_u$, $\mathbf{U}(\omega_k)$ is partitioned in M blocks of size $n_u \times n_u$ as

$$\mathbf{U}(\omega_k) = \begin{bmatrix} \mathbf{U}^{[1]}(\omega_k) & \ldots & \mathbf{U}^{[M]}(\omega_k) \end{bmatrix}, \tag{5a}$$
$$\mathbf{U}^{[m]}(\omega_k) = \mathbf{U}^{[m]}_{\mathrm{diag}}(\omega_k)\mathbf{T}, \quad m = 1, \ldots, M, \tag{5b}$$

with $\mathbf{U}^{[m]}_{\mathrm{diag}}(\omega_k) = \mathrm{diag}\left\{U_l^{[m]}(\omega_k)\right\}_{l=1}^{n_u}$ a diagonal matrix where each $U_l^{[m]}(\omega_k)$ is a random phase multisine signal, and $\mathbf{T}$ is an orthogonal matrix. A scalar random phase multisine signal $u(t)$ can be written as

$$u(t) = \sum_{k=1}^{N_f} A_k \cos(\omega_k t + \phi_k),$$

with amplitudes A_k, frequencies ω_k chosen from the grid $\{\frac{2\pi l}{N_P T_s},\ l = 1, \ldots, N_P/2 - 1\}$ (N_P even), and random phases ϕ_k uniformly distributed on the interval $[0, 2\pi)$. The optimal matrix $\mathbf{T}$, with constraints $|\mathbf{T}_{il}| \leq 1$, is given by (Dobrowiecki and Schoukens, 2005)

$$\mathbf{T}_{il} = e^{\frac{2\pi j}{n_u}(i-1)(l-1)}. \tag{6}$$

For the single input, single output (SISO) case, a number of different FRF estimators have been suggested in the literature (see, e.g., Pintelon and Schoukens, 2001; Wellstead, 1981), which all have different properties regarding bias (for the output-noise case only in closed loop) and variance. These estimators can often be generalized to the multiple input, multiple output (MIMO) case, as for the H_1 estimator (3). By using (5a), (3) can be rewritten as

$$\widehat{G}^{\mathrm{H}_1} = \left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{Y}^{[m]}\mathbf{U}^{[m]H}\right]\left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{U}^{[m]}\mathbf{U}^{[m]H}\right]^{-1}.$$

(For notational simplicity, the frequency argument will be omitted when not explicitly needed.) Another useful estimator is the *arithmetic mean estimator* (Guillaume, 1998; Pintelon and Schoukens, 2001)

$$\widehat{G}^{\mathrm{ARI}} = \frac{1}{M}\sum_{m=1}^{M}\widehat{G}^{[m]} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{Y}^{[m]}[\mathbf{U}^{[m]}]^{-1}. \tag{7}$$

If the reference signal is measured, an asymptotically ($n_e \to \infty$) unbiased estimator has been proposed in Wellstead (1981), which can be generalized to the MIMO case as (Cobb and Mitchell, 1990; Verboven, 2002)

$$\widehat{G}^{\mathrm{JIO}} = \mathbf{Y}\mathbf{R}^H\left[\mathbf{U}\mathbf{R}^H\right]^{-1}, \tag{8}$$

where $\mathbf{R}$ is the reference DFT matrix, see (16), and JIO stands for *joint input-output estimator* (cf. Ljung, 1999, p. 438). This estimator is based on the instrumental variables principle and is also called H_{IV}, H_s, H_c, or 3-channel FRF estimator in the literature (see Cobb and Mitchell, 1990; Verboven, 2002). By using (5a), (8) is rewritten as

$$\widehat{G}^{\mathrm{JIO}} = \left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{Y}^{[m]}\mathbf{R}^{[m]H}\right]\left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{U}^{[m]}\mathbf{R}^{[m]H}\right]^{-1}.$$

If the same excitation is used in all blocks with synchronized measurements[1], then the JIO estimator reduces to the *errors-in-variables*[2] *estimator* (Guillaume et al., 1996; Pintelon and Schoukens, 2001)

$$\widehat{G}^{\mathrm{EIV}} = \left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{Y}^{[m]}\right]\left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{U}^{[m]}\right]^{-1}. \tag{9}$$

The EIV estimator does not require the reference signal to be known and this is also the maximum likelihood estimator for normally distributed noise (Guillaume et al., 1996).

[1]Synchronized measurements here means that all $\mathbf{U}^{[m]}$ are equal, except for noise. This can easily be achieved by using one orthogonal multisine signal with multiple periods, where each period then is treated as a "block".

[2]Measurements $y(t) = G(q)u(t) + v_y(t)$ and $u(t) + v_u(t)$ with measurement noise $v_u(t)$ and $v_y(t)$.

3 Open-Loop Error Analysis

In this section, the estimators will be analyzed for open-loop data. The analysis of the H_1 estimator is closely related to the work in Guillaume et al. (1996), Verboven (2002), and Dobrowiecki et al. (2005). In Dobrowiecki et al. (2005), covariance expressions are derived when the orthogonal multisine signal (5) is used as input. The MIMO system is there viewed as n_y separate MISO systems with output noise. Implicitly, the output noise is assumed to be independent over the different outputs so expressions for the covariance between different rows in $\widehat{G}$ are not presented. In this paper, the MIMO system will be considered without such noise assumption, which can be seen as applying (5) to Theorem B.1. In addition, the ARI and EIV estimators are treated. These results will be extended to the closed-loop case in Section 4.

3.1 Noise Assumptions

Assumption B.1. For each of the n_e experiments, the DFT of the noise, $V(\omega_k)$ satisfies

$$\mathrm{E}(V(\omega_k)) = 0, \quad \mathrm{E}(V(\omega_k)V^T(\omega_k)) = 0,$$
$$\mathrm{E}(V(\omega_k)V^H(\omega_k)) = \sigma_V^2(\omega_k),$$

for $k = 1, \ldots, N_f$, with $\mathrm{E}(X)$ the expected value of X. $V(\omega_k)$ is therefore circular complex. In addition, it is assumed that the noise is independent and identically distributed over the n_e different experiments.

This assumption is common and justifiable in most practical circumstances[3] and enables us to obtain the following result for the covariance of $\mathbf{V}(\omega_k)$.

Lemma B.1
Consider the DFT matrix $\mathbf{V}(\omega_k) \in \mathbb{C}^{n_y \times n_e}$ *in* (4)*. Under Assumption B.1, the covariance* $\sigma^2_{\mathbf{V}}(\omega_k)$ *is given by*

$$\sigma^2_{\mathbf{V}}(\omega_k) = \mathrm{E}(\mathrm{vec}(\mathbf{V}(\omega_k))\,\mathrm{vec}(\mathbf{V}(\omega_k))^H) = I \otimes \sigma_V^2(\omega_k),$$

where $\otimes$ *is the Kronecker product.*

Proof: Let $\mathbf{V}_i$ denote column i in $\mathbf{V}$. $\sigma^2_{\mathbf{V}}$ will then be a block matrix with blocks $\mathrm{E}(\mathbf{V}_i\mathbf{V}_j^H) \in \mathbb{C}^{n_y \times n_y}$. Expanding this expression and using Assumption B.1 gives that $\mathrm{E}(\mathbf{V}_i\mathbf{V}_j^H) = \sigma_V^2$ if $i = j$ and zero otherwise. $\sigma^2_{\mathbf{V}}$ will therefore be a block-diagonal matrix with σ_V^2 in all the diagonal blocks, which can be written as $I \otimes \sigma_V^2$. □

3.2 Bias and Covariance

The bias and covariance of the H_1 estimator (3) will now be derived in the following theorem. The same result can be found in Verboven (2002, p. 24) and Brillinger (1981, Theorem 8.2.5), see also Guillaume et al. (1996).

[3] For example, the DFT of filtered white noise is asymptotically circular complex when the number of samples, $N \to \infty$ (see Pintelon and Schoukens, 2001, Theorem 14.25).

Theorem B.1
Consider the H_1 estimator (3) *and assume an open-loop setup* (4) *where* $\mathbf{U}(\omega_k)$ *and* $\mathbf{V}(\omega_k)$ *are independent and the noise fulfills Assumption B.1. The estimate* $\widehat{G}^{H_1}(e^{j\omega_k T_s})$ *will then be unbiased and the covariance* $\sigma^2_{\widehat{G}^{H_1}}(\omega_k)$ *is given by*

$$\sigma^2_{\widehat{G}^{H_1}}(\omega_k) = \mathrm{E}(\mathrm{vec}(\tilde{G}^{H_1}(e^{j\omega_k T_s}))\,\mathrm{vec}(\tilde{G}^{H_1}(e^{j\omega_k T_s}))^H) = \\ = \mathrm{E}(\left[\mathbf{U}(\omega_k)\mathbf{U}^H(\omega_k)\right]^{-T}) \otimes \sigma^2_V(\omega_k). \quad (10)$$

Proof: The estimation error $\tilde{G}^{\mathrm{H}_1}$ can be rewritten as

$$\tilde{G}^{\mathrm{H}_1} = \widehat{G}^{\mathrm{H}_1} - G = (G\mathbf{U} + \mathbf{V})\mathbf{U}^\dagger - G = \mathbf{V}\mathbf{U}^\dagger, \quad (11)$$

with $\mathbf{U}^\dagger = \mathbf{U}^H(\mathbf{U}\mathbf{U}^H)^{-1}$. $\mathrm{E}(\tilde{G}^{\mathrm{H}_1}) = 0$ follows since $\mathrm{E}(\mathbf{V}) = 0$ and $\mathbf{U}$ and $\mathbf{V}$ are independent. For the covariance, $\mathrm{vec}(\tilde{G}^{\mathrm{H}_1})$ can be rewritten as $\mathrm{vec}(\tilde{G}^{\mathrm{H}_1}) = (\mathbf{U}^{\dagger T} \otimes I)\,\mathrm{vec}(\mathbf{V})$ by using $\mathrm{vec}(ADB) = (B^T \otimes A)\,\mathrm{vec}(D)$ (see, e.g., Pintelon and Schoukens, 2001, p. 420). Equation (10) then follows from

$$\begin{aligned}
\sigma^2_{\widehat{G}^{\mathrm{H}_1}} &= \mathrm{E}(\mathrm{vec}(\tilde{G}^{\mathrm{H}_1})\,\mathrm{vec}(\tilde{G}^{\mathrm{H}_1})^H) \\
&= \mathrm{E}([\mathbf{U}^{\dagger T} \otimes I]\,\mathrm{vec}(\mathbf{V})\,\mathrm{vec}(\mathbf{V})^H[\mathbf{U}^{\dagger T} \otimes I]^H) \\
&= \mathrm{E}_u([\mathbf{U}^{\dagger T} \otimes I]\mathrm{E}_v(\mathrm{vec}(\mathbf{V})\,\mathrm{vec}(\mathbf{V})^H)[\mathbf{U}^{\dagger T} \otimes I]^H) \\
&= \mathrm{E}_u([\mathbf{U}^{\dagger T} \otimes I][I \otimes \sigma^2_V][\overline{\mathbf{U}^\dagger} \otimes I]) \\
&= \mathrm{E}(\mathbf{U}^{\dagger T}\overline{\mathbf{U}^\dagger}) \otimes \sigma^2_V = \mathrm{E}([\mathbf{U}^{\dagger H}\mathbf{U}^\dagger]^T) \otimes \sigma^2_V \\
&= \mathrm{E}([(\mathbf{U}\mathbf{U}^H)^{-1}\mathbf{U}\mathbf{U}^H(\mathbf{U}\mathbf{U}^H)^{-1}]^T) \otimes \sigma^2_V,
\end{aligned}$$

using $\mathrm{E}(\cdot) = \mathrm{E}_u(\mathrm{E}_v(\cdot))$ in the second step since $\mathbf{U}$ and $\mathbf{V}$ are independent, Lemma B.1 in the third step, and $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ twice in the fourth step. □

Continuing the error analysis with the orthogonal multisine signal (5) gives the following results (cf. Dobrowiecki et al., 2005).

Corollary B.1
Using an orthogonal multisine signal (5) *with* $\mathbf{T}$ *from* (6) *simplifies* $\sigma^2_{\widehat{G}^{H_1}}$ *in Theorem B.1 to*

$$\sigma^2_{\widehat{G}^{H_1}} = \mathrm{diag}\left\{\frac{1}{n_u \sum_{m=1}^{M} |U_l^{[m]}|^2}\right\}_{l=1}^{n_u} \otimes \sigma^2_V,$$

and the variance of each element in $\widehat{G}^{H_1}$ *to*

$$\sigma^2_{\widehat{G}^{H_1}_{ij}} = \frac{\sigma^2_{V,ii}}{n_u \sum_{m=1}^{M} |U_j^{[m]}|^2}. \quad (12)$$

Proof: Follows by noting that $\mathbf{T}\mathbf{T}^H = n_u I$ for (6) as

$$\mathbf{U}\mathbf{U}^H = \sum\nolimits_m \mathbf{U}^{[m]}\mathbf{U}^{[m]H} = \sum\nolimits_m \mathbf{U}^{[m]}_{\mathrm{diag}}\mathbf{T}\mathbf{T}^H\mathbf{U}^{[m]H}_{\mathrm{diag}} \\ = n_u \sum\nolimits_m \mathbf{U}^{[m]}_{\mathrm{diag}}\mathbf{U}^{[m]H}_{\mathrm{diag}} = n_u \sum\nolimits_m \mathrm{diag}\left\{|U_l^{[m]}|^2\right\}_{l=1}^{n_u}.$$

□

This means that different columns in $\widehat{G}^{\mathrm{H}_1}$ are uncorrelated and that the covariance for a certain column is inversely proportional to the total input power in that particular input channel. Note that if another type of excitation signal $\mathbf{U}$ is used, without the property that $\mathbf{U}\mathbf{U}^H$ is diagonal, then the covariance is increased and the columns in $\widehat{G}^{\mathrm{H}_1}$ will be correlated. The variance (12) is given by a noise-to-signal ratio, where $\sigma^2_{V,ii}$ is the noise variance in output i and the denominator is the total power for input j during the n_e experiments.

The covariance for the ARI and EIV estimators, (7) and (9), can easily be derived, similarly to the proof of Theorem B.1,

$$\sigma^2_{\widehat{G}^{\mathrm{ARI}}} = \frac{1}{M^2}\mathrm{E}\Big(\sum_{m=1}^{M}\left[\mathbf{U}^{[m]}\mathbf{U}^{[m]H}\right]^{-T}\Big) \otimes \sigma^2_V, \tag{13}$$

$$\sigma^2_{\widehat{G}^{\mathrm{EIV}}} = M\mathrm{E}\Big(\left[\sum_{m=1}^{M}\mathbf{U}^{[m]}\sum_{n=1}^{M}\mathbf{U}^{[n]H}\right]^{-T}\Big) \otimes \sigma^2_V. \tag{14}$$

The H_1, ARI, and EIV estimators coincide in case all $\mathbf{U}^{[m]}$ are equal. The ARI and H_1 estimators coincide also when using the orthogonal multisine signal with the same magnitude in all blocks, $|U_l^{[m]}| = |U_l|,\ m = 1, \ldots, M$. In other cases the covariance is typically larger for the ARI estimator than the H_1 estimator, and the EIV estimator can have infinite covariance.

All estimators are unbiased in the case of open-loop data and output noise. The EIV (and JIO) estimators are in addition unbiased in the error-in-variables case and for closed-loop data, which will be treated next.

4 Closed-Loop Error Analysis

In this paper, the aim is to analyze also the properties of the estimation methods in a closed-loop setup. Therefore, let the input in (1) be given by

$$u(t) = F(q)(r(t) - y(t)), \tag{15}$$

where $F(q)$ is the controller and $r(t) \in \mathbb{R}^{n_y}$ is the reference signal. For simplicity, $n_y = n_u$ will be assumed in this section ($n_y \geq n_u$ is enough for most expressions). The input DFT matrix $\mathbf{U}(\omega_k)$ is then given by

$$\mathbf{U}(\omega_k) = G_u(e^{j\omega_k T_s})(\mathbf{R}(\omega_k) - \mathbf{V}(\omega_k)), \tag{16}$$

where

$$G_u(q) = (I + F(q)G(q))^{-1}F(q).$$

4.1 Bias

Since $\mathbf{U}$ and $\mathbf{V}$ now are correlated, the H_1 estimator (3) will be biased. To analyze the bias, we will assume a closed-loop setup according to (1) and (15), where the noise fulfills

Assumption B.1 and $\mathbf{R}$ and $\mathbf{V}$ are independent and uniformly bounded. The estimation error is given by

$$\begin{aligned}\tilde{G}^{\mathrm{H}_1} =& \widehat{G}^{\mathrm{H}_1} - G = \mathbf{V}\mathbf{U}^{\dagger} = \mathbf{V}(\mathbf{R}-\mathbf{V})^{\dagger}G_u^{-1} \\ =& [\mathbf{V}\mathbf{R}^H - \mathbf{V}\mathbf{V}^H][\mathbf{R}\mathbf{R}^H + \mathbf{V}\mathbf{V}^H - \mathbf{R}\mathbf{V}^H - \mathbf{V}\mathbf{R}^H]^{-1}G_u^{-1}.\end{aligned} \tag{17}$$

Calculating the bias $\mathrm{E}(\tilde{G}^{\mathrm{H}_1})$ is hard due to the matrix inverse, even though Taylor expansions can be used, see Guillaume et al. (1992) and Guillaume (1998). Instead, consistency will first be considered, where $\widehat{G}^{\mathrm{H}_1}$ is said to be (strongly) consistent if $\underset{n_e\to\infty}{\text{a.s.lim}}\, \widehat{G}^{\mathrm{H}_1} = G$. This analysis will use some fundamental results for stochastic limits, see, e.g., Pintelon and Schoukens (2001, Chap. 14) for a summary.

With suitable assumptions on $\mathbf{V}$ to avoid inversion of a singular matrix[4], $\tilde{G}^{\mathrm{H}_1}$ is uniformly bounded and continuous, viewed as a function of $\mathbf{V}$ and $\mathbf{R}$. The almost sure limit (a.s.lim) and a continuous function may be interchanged, such that the limit can be moved into the expression as

$$\begin{aligned}\underset{n_e\to\infty}{\text{a.s.lim}}\, \tilde{G}^{\mathrm{H}_1} =& \left[\underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{V}_i\mathbf{R}_i^H - \underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{V}_i\mathbf{V}_i^H\right]\times \\ & \left[\underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{R}_i\mathbf{R}_i^H + \underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{V}_i\mathbf{V}_i^H \right. \\ & \left. - \underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{R}_i\mathbf{V}_i^H - \underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{V}_i\mathbf{R}_i^H\right]^{-1} G_u^{-1}.\end{aligned}$$

The sums $\underset{n_e\to\infty}{\text{a.s.lim}}\, \frac{1}{n_e}\sum_{i=1}^{n_e}\mathbf{V}_i\mathbf{R}_i^H$, and so on, of independent and uniformly bounded variables can, by the strong law of large numbers, be replaced by expectation, which gives

$$\underset{n_e\to\infty}{\text{a.s.lim}}\, \tilde{G}^{\mathrm{H}_1} = -\sigma_V^2[\sigma_R^2 + \sigma_V^2]^{-1}G_u^{-1}, \tag{18}$$

where $\sigma_R^2 = \lim_{n_e\to\infty}\frac{1}{n_e}\sum_{i=1}^{n_e}\mathrm{E}(\mathbf{R}_i\mathbf{R}_i^H)$. The estimator is therefore inconsistent. For $\tilde{G}^{\mathrm{H}_1}$ uniformly bounded, $\text{a.s.lim}_{n_e\to\infty}\tilde{G}^{\mathrm{H}_1} = \lim_{n_e\to\infty}\mathrm{E}(\tilde{G}^{\mathrm{H}_1})$, which means that (18) also is an expression for the asymptotic bias. Using $G_u^{-1} = G + F^{-1}$ finally gives

$$\underset{n_e\to\infty}{\text{a.s.lim}}\, \widehat{G}^{\mathrm{H}_1} = \sigma_R^2[\sigma_R^2 + \sigma_V^2]^{-1}G - \sigma_V^2[\sigma_R^2 + \sigma_V^2]^{-1}F^{-1}. \tag{19}$$

For frequencies where the signal-to-noise ratio (SNR) is poor, the estimate will tend to the inverse controller. If σ_V^2 and σ_R^2 are diagonal, (18) simplifies to

$$\underset{n_e\to\infty}{\text{a.s.lim}}\, \tilde{G}_{ij}^{\mathrm{H}_1} = -\frac{\sigma_{V,ii}^2}{\sigma_{V,ii}^2 + \sigma_{R,ii}^2}(G_{ij} + F_{ij}^{-1}), \tag{20}$$

such that the asymptotic relative bias $(\lim_{n_e\to\infty}\mathrm{E}(\tilde{G}_{ij}^{\mathrm{H}_1})/G_{ij})$ will be (approximately) proportional to the noise-to-signal ratio.

[4] If the probability density function of $\mathbf{V}$ is short-tailed (e.g., uniform), that is handled if the SNR is large enough. See Guillaume et al. (1992) for details.

These bias expressions could be compared with the expressions for the ARI estimator (7) in the SISO case (Pintelon and Schoukens, 2001)

$$\mathrm{E}(\tilde{G}^{\mathrm{ARI}}) = -e^{-\sigma_R^2/\sigma_V^2} G_u^{-1},$$
$$\mathrm{E}(\widehat{G}^{\mathrm{ARI}}) = (1 - e^{-\sigma_R^2/\sigma_V^2})G - e^{-\sigma_R^2/\sigma_V^2} F^{-1},$$

which indicate a huge difference for large SNRs. A similar difference can be seen also in the MIMO case. When the SNR is small, on the other hand, the ARI estimator will deteriorate since $\mathbf{U}^{[m]}$ could loose rank for some blocks, giving infinite covariance. The H_1 estimator is, in that respect, a more robust estimator.

For closed-loop data, the EIV and JIO estimators should also be considered since they are consistent and asymptotically unbiased, which can be seen by studying their estimation errors

$$\tilde{G}^{\mathrm{EIV}} = \left[\frac{1}{M}\sum_{m=1}^{M} \mathbf{V}^{[m]}\right]\left[\frac{1}{M}\sum_{m=1}^{M} \mathbf{U}^{[m]}\right]^{-1}, \tag{21}$$

$$\tilde{G}^{\mathrm{JIO}} = \left[\frac{1}{n_e}\sum_{i=1}^{n_e} \mathbf{V}_i\mathbf{R}_i^H\right]\left[\frac{1}{n_e}\sum_{i=1}^{n_e} \mathbf{R}_i\mathbf{R}_i^H - \frac{1}{n_e}\sum_{i=1}^{n_e} \mathbf{V}_i\mathbf{R}_i^H\right]^{-1} G_u^{-1}, \tag{22}$$

and using similar arguments as for $\tilde{G}^{\mathrm{H}_1}$, see also Verboven (2002).

A problem with the asymptotic bias expressions (18)–(20) is that for large SNRs, n_e must be fairly large until these expressions are useful, mainly due to the variance of $X = \frac{1}{n_e}\sum_{i=1}^{n_e} \mathbf{V}_i(\mathbf{R}_i - \mathbf{V}_i)^H$. To see this, assume for simplicity normally distributed noise and $\sigma_R^2 = \alpha \cdot I$ and $\sigma_V^2 = \beta \cdot I$, with α and β scalars. The matrix X, with mean $-\sigma_V^2$, then has variance $\mathrm{E}(|(X+\sigma_V^2)_{ij}|^2) = (\beta\alpha+\beta^2)/n_e$. For the matrix to be diagonal dominant, we require $(\beta\alpha + \beta^2)/n_e \ll \beta^2$, or $n_e \gg 1 + \alpha/\beta$, where α/β is the SNR.

4.2 Covariance

It is hard to obtain exact covariance expressions for the considered estimators for the closed-loop case since the noise appears in the matrix inverse. However, for the JIO estimator an asymptotic expression is given by the following theorem.

Theorem B.2
Consider the JIO estimator (8) *and assume a closed-loop setup according to* (1) *and* (15)*, where the noise fulfills Assumption B.1 and* $\mathbf{R}$ *and* $\mathbf{V}$ *are independent and uniformly bounded. The covariance is then asymptotically* ($n_e \to \infty$) *given by*

$$\sigma^2_{\widehat{G}^{JIO}} = \left[G_u^{-H}\mathrm{E}((\mathbf{R}\mathbf{R}^H)^{-1})G_u^{-1}\right]^T \otimes \sigma_V^2, \tag{23}$$

with $G_u^{-1} = G + F^{-1}$.

Proof: (sketch) Consider the estimation error (22). The matrix inverse will be dominated by $[\frac{1}{n_e}\sum_{i=1}^{n_e} \mathbf{R}_i\mathbf{R}_i^H]^{-1}$ since $\frac{1}{n_e}\sum_{i=1}^{n_e} \mathbf{V}_i\mathbf{R}_i^H \to 0$ as $n_e \to \infty$. $\tilde{G}^{\mathrm{JIO}} = \mathbf{V}\mathbf{R}^\dagger G_u^{-1}$ is then a valid approximation for large n_e. This corresponds to using $\mathbf{U} = G_u\mathbf{R}$ in (11) and the covariance (23) is therefore obtained by inserting $\mathbf{U} = G_u\mathbf{R}$ into (10). □

Even though this is an asymptotic covariance expression, it still gives valuable insight to how the covariance is affected by the closed-loop setup, even for a finite number of experiments n_e. The validity for finite n_e will depend both on the SNR and n_e.

To obtain covariance expressions for the remaining estimators, the following approximation will be used

$$\mathbf{U} \approx G_u \mathbf{R}, \tag{24}$$

which requires $\mathbf{V} \ll \mathbf{R}$ in order to be valid for the H_1 and ARI estimators. Using (24) for the EIV estimator will give the asymptotic covariance expression, similarly as for the JIO estimator. The approximation (24) can be viewed as neglecting the noise in the feedback loop. Inserting (24) into the open-loop expressions from the previous section then yield:

$$\sigma^2_{\widehat{G}^{\mathrm{H}_1}} = \left[G_u^{-H} \mathrm{E}((\mathbf{R}\mathbf{R}^H)^{-1}) G_u^{-1}\right]^T \otimes \sigma_V^2, \tag{25}$$

$$\sigma^2_{\widehat{G}^{\mathrm{ARI}}} = \left[G_u^{-H} \frac{1}{M^2} \mathrm{E}(\sum_{m=1}^{M} (\mathbf{R}^{[m]}\mathbf{R}^{[m]H})^{-1}) G_u^{-1}\right]^T \otimes \sigma_V^2, \tag{26}$$

$$\sigma^2_{\widehat{G}^{\mathrm{EIV}}} = \left[G_u^{-H} M \mathrm{E}(\left[\sum_{m=1}^{M} \mathbf{R}^{[m]} \sum_{n=1}^{M} \mathbf{R}^{[n]H}\right]^{-1}) G_u^{-1}\right]^T \otimes \sigma_V^2. \tag{27}$$

According to the stated expressions (23), (25)–(27), the four estimators will have the same covariance in case all $\mathbf{R}^{[m]}$ are equal. Except for the EIV estimator, the expressions coincide also when using the orthogonal multisine signal (28) with the same magnitude in all blocks, $|R_l^{[m]}| = |R_l|,\ m = 1, \ldots, M$. However, when the SNR is reduced, the covariance of the H_1 and ARI estimators will deviate from the approximate expressions, typically giving a larger covariance for the ARI estimator and a lower covariance for the H_1 estimator. This can also be seen in the numerical illustration in the next section.

Now, similar expressions to Corollary B.1 will be derived using the orthogonal multisine signal

$$\mathbf{R}(\omega_k) = \begin{bmatrix}\mathbf{R}^{[1]}(\omega_k) & \ldots & \mathbf{R}^{[M]}(\omega_k)\end{bmatrix}, \tag{28a}$$

$$\mathbf{R}^{[m]}(\omega_k) = \mathbf{R}^{[m]}_{\mathrm{diag}}(\omega_k)\mathbf{T}, \quad m = 1, \ldots, M, \tag{28b}$$

as in (5) with $\mathbf{T}$ from (6). This parameterization corresponds to an optimal experiment design given output amplitude constraints, compared to the open-loop case (5) when having input amplitude constraints. If we actually have input constraints in the closed-loop case, $\mathbf{R} = G_u^{-1}\mathbf{U}$ can be used, yielding $\mathbf{R}_{\mathrm{diag}} = G_u^{-1}\mathbf{U}_{\mathrm{diag}}$. Note that $\mathbf{R}_{\mathrm{diag}}$ then no longer is diagonal.

Corollary B.2
Using an orthogonal multisine signal (28) *with* $\mathbf{T}$ *from* (6) *simplifies the covariance of the JIO and* H_1 *estimators to*

$$\sigma^2_{\widehat{G}} = \left[G_u^{-H}\, \mathrm{diag}\left\{\frac{1}{n_u \sum_{m=1}^{M} |R_l^{[m]}|^2}\right\}_{l=1}^{n_u} G_u^{-1}\right]^T \otimes \sigma_V^2,$$

with $G_u^{-1} = G + F^{-1}$. The variance of each element in $\widehat{G}$ can be written as

$$\sigma^2_{\widehat{G}_{ij}} = \sigma^2_{V,ii} \sum_{l=1}^{n_u} \frac{|(G + F^{-1})_{lj}|^2}{n_u \sum_{m=1}^{M} |R_l^{[m]}|^2}. \tag{29}$$

One immediately notes that to reduce the variance, the reference signal should be as large as possible. In addition, a large gain controller will also reduce the variance. Actually, $F = -G^{-1}$, would give zero error but is unrealistic since it corresponds to a marginally stable system and infinite input power. If the noise variance is the same for the different outputs, then all elements in a column will have the same variance. The variance will typically vary between different columns so for a symmetric system ($|G_{ij}| = |G_{ji}|$), the variance will usually be non-symmetric ($\sigma^2_{\widehat{G}_{ij}} \neq \sigma^2_{\widehat{G}_{ji}}$). Small elements in a column will also, in general, have a larger relative error ($\sigma_{\widehat{G}_{ij}}/|G_{ij}|$).

5 Numerical Illustration

As a numerical illustration, a simulation model of an industrial robot will be used. The model $G(q)$ is a linearized version of a nonlinear state-space model from Öhr et al. (2006) with 24 states, describing the dynamics of the first three axes ($n_u = n_y = 3$) from applied motor torque to achieved motor velocity (the FRF of G can be seen in all the figures). The controller $F(q)$ is a diagonal PI controller with a gain of 8 dB in the excited frequency interval. A sampling period of $T_s = 0.5$ ms is used with $N = N_P = 10^4$ samples from the steady-state response.

As excitation, the orthogonal multisine signal (6) and (28) is used with a flat amplitude spectrum from 2 to 60 Hz and $\mathbf{R}^{[m]}_{\text{diag}}$ the same in all $M = 100$ blocks. The EIV and JIO estimators will therefore coincide in this illustration. The noise v in (1) is normally distributed white noise, giving a diagonal covariance matrix σ^2_V. The resulting signal-to-noise ratio, $\text{SNR}_i = 20\log_{10}(\sigma_{R,ii}/\sigma_{V,ii})$, is constant for the excited frequencies. Simulations with two different SNRs, 3 dB and 22 dB will be presented. The statistical properties of the estimators are obtained by repeating the simulations 100 times with different noise realizations.

First, consider the 22 dB case and the three estimators: H_1 (3), ARI (7) and JIO (8). The bias $\text{E}(\tilde{G})$ and variance $\sigma^2_{\widehat{G}}$ (actually the standard deviation $\sigma_{\widehat{G}}$) of the estimators are plotted in Figures 1 and 2, respectively. In addition are the bias expression for the H_1 estimator, (18), and the variance expression (29) plotted. Note that all four covariance expressions, (23), (25)–(27), coincide for the excitation used in this illustration. That can also be seen in Figure 2, where the variance is approximately the same for the three estimators and the variance expression is fairly accurate.

As was mentioned in the end of Section 4, the variance is non-symmetric, which can be seen by comparing elements (1,3) and (3,1) in Figure 2. For small elements, the relative error is also larger, which can be seen by, e.g., comparing elements (1,1) and (3,1) in Figure 2. This is inherent in all the studied estimation methods.

The bias of the ARI and JIO estimators, see Figure 1, are approximately the same and always equal to or lower than for the H_1 estimator. The asymptotic bias expression (18) is a lower limit for the bias of the H_1 estimator and there is a good match when G

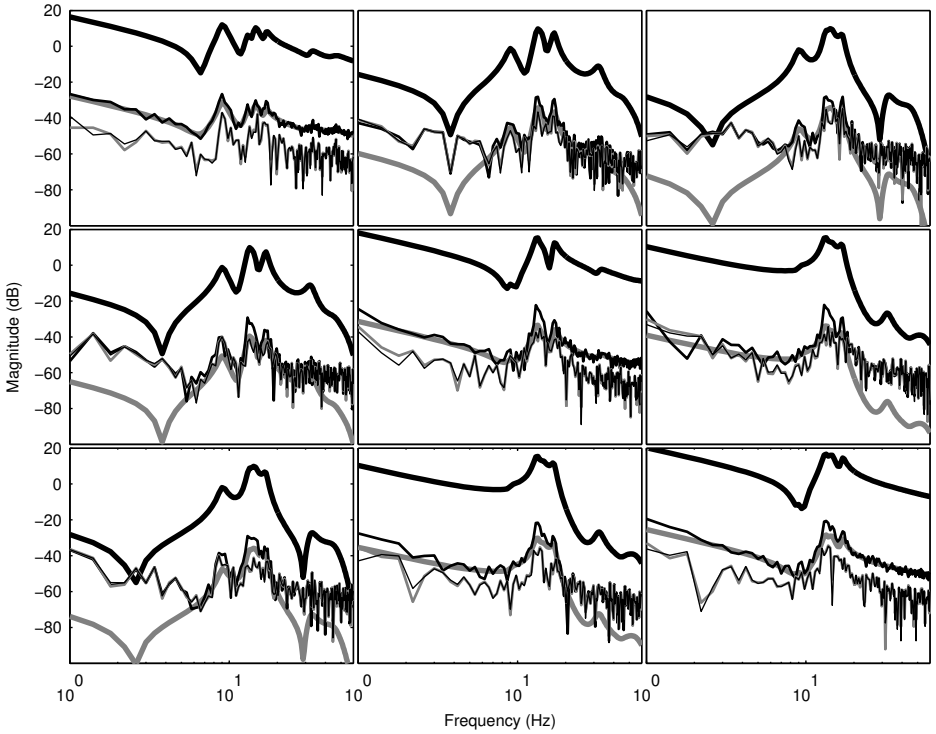


Figure 1: *Bias for the estimators with* 22 *dB SNR. Thick black line:* G*, thick gray line: bias expression* (18)*, black line:* H_1*, gray line: ARI, and thin black line: JIO.*

is large, especially for the diagonal elements. In this simulation, σ_V^2 and σ_R^2 are diagonal so the asymptotic expression (18) simplifies to (20). A problem is that with $M = 100$, the off-diagonal elements of $\mathbf{V}(\mathbf{R} - \mathbf{V})^\dagger$ in (17) are still quite large. If all elements of $\mathbf{V}(\mathbf{R} - \mathbf{V})^\dagger$ would be equal to α, then the bias is $\tilde{G}_{ij} = \alpha \sum_l (G_{lj} + F_{lj}^{-1})$, i.e., the same bias for all elements in a given column. That is almost the case here since $M = 100$ is not large enough, which explains the behavior for the off-diagonal elements in Figure 1.

To see the different properties of the three estimators, the SNR is reduced to 3 dB and the bias and variance are plotted in Figures 3 and 4, respectively. As can be seen, there is now a big difference both in bias and variance. The H_1 estimator gives largest bias and smallest variance and the JIO estimator gives smallest bias and a smaller variance than the ARI estimator. The JIO estimator should therefore be used if the reference signal is measured, or the EIV estimator in case the measurements are synchronized and the same excitation is used in all blocks. The variance expression (29) predicts the variance of the JIO estimator fairly accurately, but the bias expression (18) for the H_1 estimator differs quite much here. The validity of the bias expression with respect to the SNR therefore needs further studies.

Since the properties of the EIV and JIO estimators coincide for certain excitations, one might wonder which one is the preferred choice. Assuming a linear system, there is

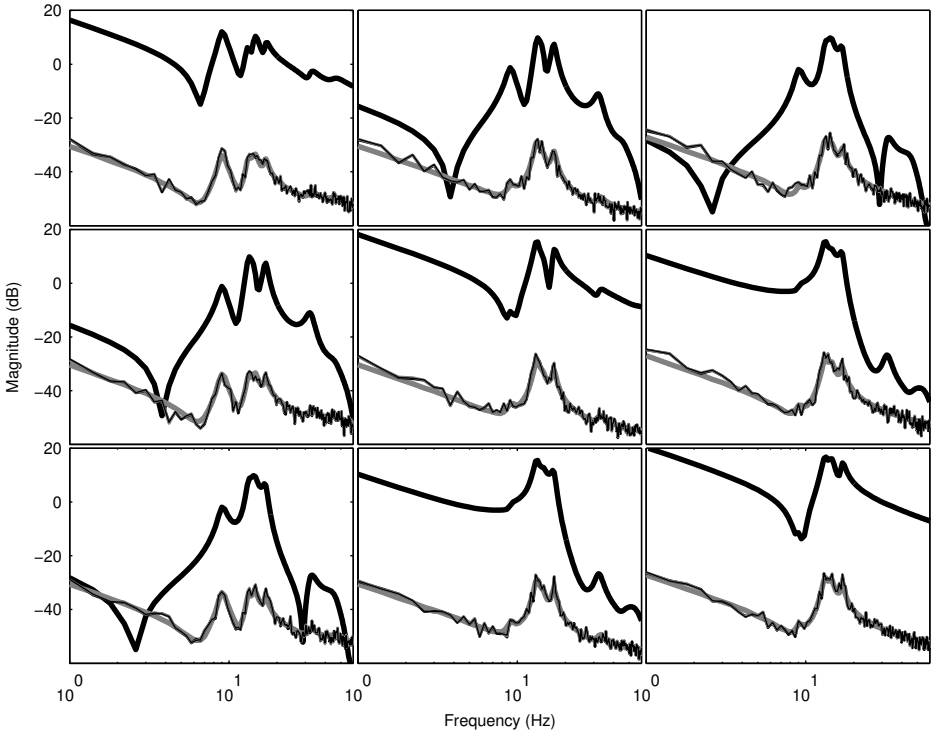


Figure 2: *Standard deviation for the estimators with* 22 *dB SNR. Thick black line:* G*, thick gray line: the variance expression* (29)*, black line:* H_1*, gray line: ARI, and thin black line: JIO.*

no point in using different realizations of the orthogonal multisine signal in the different blocks. Instead can the same excitation be applied for numerous periods, treating each period as a "block" (which is the case in these simulations). As long as the measurements can be synchronized, the EIV estimator is then the preferred choice. Both since the measurements of the reference signal are not needed and in case there are any nonlinearities in the feedback loop. On the other hand, consider the case with a nonlinear system and the aim to find a linear approximation of the system. Then it is crucial to use different realizations of the orthogonal multisine signal to find the best linear approximation (Dobrowiecki and Schoukens, 2005; Wernholt and Gunnarsson, 2006) and the EIV estimator cannot be used. The choice therefore depends on the measurement setup as well as assumptions about if the system is linear or not.

6 Conclusion

In this paper, four nonparametric FRF estimation methods (H_1, ARI, JIO, EIV) for multivariable systems have been studied. Covariance expressions for open-loop data are pre-

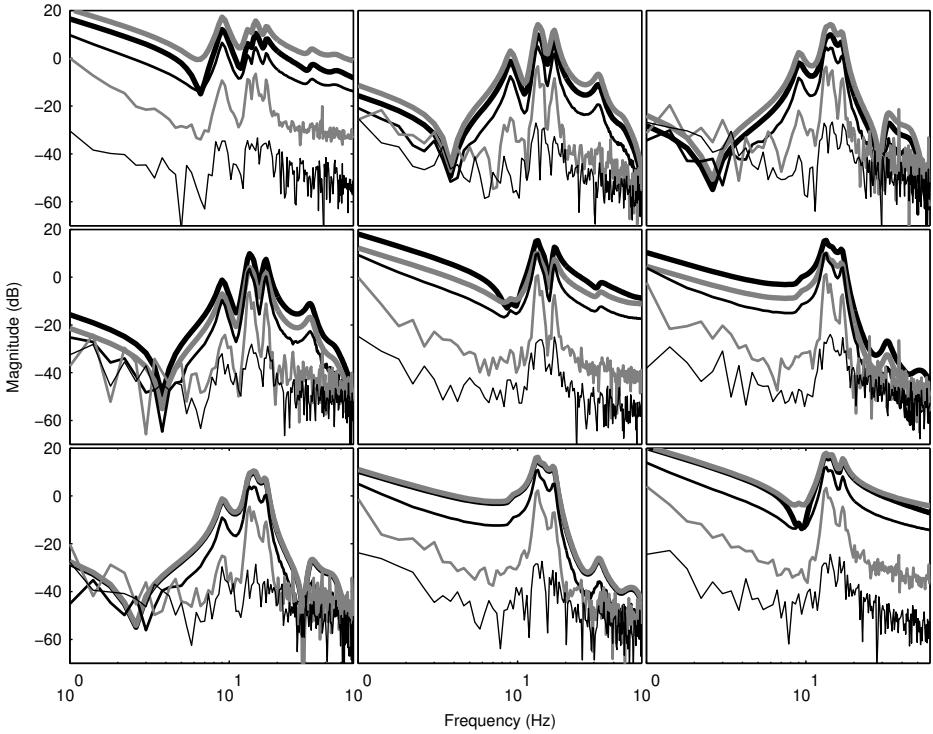


Figure 3: *Bias for the estimators with* 3 *dB SNR. Thick black line:* G*, thick gray line: bias expression* (18)*, black line:* H_1*, gray line: ARI, and thin black line: JIO.*

sented for the H_1, ARI, and EIV estimators. For closed-loop data, an asymptotic expression for the bias of the H_1 estimator is derived together with asymptotic covariance expressions for the JIO and EIV estimators, as well as approximate covariance expressions for the H_1 and ARI estimators.

In the simulations, all estimators give approximately the same variance for large SNRs and the bias of the H_1 estimator is larger or equal to the bias of the ARI and JIO estimators (JIO and EIV coincide in the simulations). For small SNRs, the H_1 estimator gives largest bias and smallest variance and the JIO estimator gives smallest bias and a smaller variance than the ARI estimator. Among these estimators, either the JIO or EIV estimator is the preferred choice, depending on if the reference signal is measured or if the same excitation is used in all blocks with synchronized measurements. Otherwise, the ARI estimator should be used if the SNR is guaranteed to be large enough in all blocks (since each $\mathbf{U}^{[m]}$ must be invertible, see also (13)). If that cannot be guaranteed, the H_1 estimator should be used. The H_1 estimator is also a good choice if the bias is negligible.

The proposed covariance expressions predict the covariance fairly accurately, especially for the JIO estimator, but the bias expression (18) for the H_1 estimator differ quite much for small SNRs.

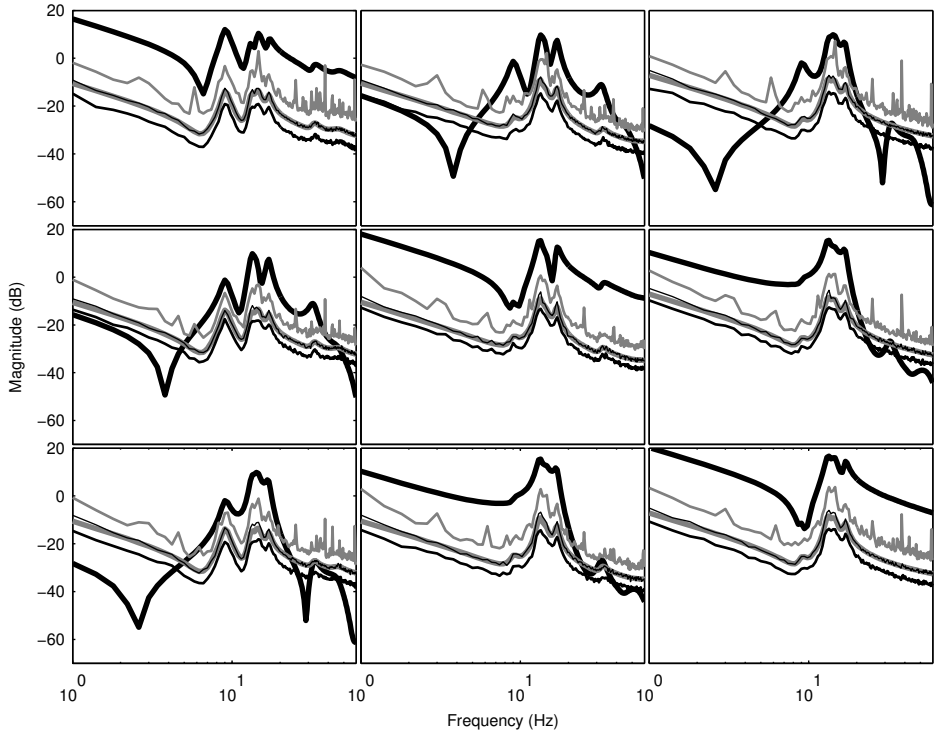


Figure 4: *Standard deviation for the estimators with* 3 *dB SNR. Thick black line: G, thick gray line: the variance expression* (29)*, black line:* H_1*, gray line: ARI, and thin black line: JIO.*

References

Berglund, E. and Hovland, G. E. (2000). Automatic elasticity tuning of industrial robot manipulators. In *Proc. 39th IEEE Conference on Decision and Control*, pages 5091–5096, Sydney, Australia.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. McGraw-Hill, New York, expanded edition.

Cobb, R. and Mitchell, L. (1990). Estimates of variance and confidence bands for the three-channel frequency response function estimator. *International Journal of Analytical and Experimental Modal Analysis*, 5(3):185–194.

Dobrowiecki, T. and Schoukens, J. (2005). Measuring linear approximation to weakly nonlinear MIMO systems. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Dobrowiecki, T., Schoukens, J., and Guillaume, P. (2005). Optimized excitation signals

for MIMO frequency response function measurments systems. In *Proc. 22nd IEEE Instrumentation and Measurement Technology Conference*, pages 1872–1877.

Guillaume, P. (1998). Frequency response measurements of multivariable systems using nonlinear averaging techniques. *IEEE Transactions on Instrumentation and Measurement*, 47(3):796–800.

Guillaume, P., Pintelon, R., and Schoukens, J. (1996). Accurate estimation of multivariable frequency response functions. In *Proc. 13th IFAC World Congress*, pages 423–428, San Francisco, California.

Guillaume, P., Pintelon, R., and Schoukens, R. (1992). Nonparametric frequency response function estimators based on nonlinear averaging techniques. *IEEE Transactions on Instrumentation and Measurement*, 41(6):739–746.

Khorrami, F., Jain, S., and Tzes, A. (1995). Experimental results on adaptive nonlinear control and input preshaping for multi-link flexible manipulators. *Automatica*, 31(1):83–97.

Kozlowski, K. (1998). *Modelling and identification in robotics.* Advances in Industrial Control. Springer, London.

Ljung, L. (1999). *System Identification: Theory for the User.* Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach.* IEEE Press, New York.

Pintelon, R. and Schoukens, J. (2001). Measurement of frequency response functions using periodic excitations, corrupted by correlated input/output errors. *IEEE Transactions on Instrumentation and Measurement*, 50(6):1753–1760.

Verboven, P. (2002). *Frequency-domain system identification for modal analysis.* PhD thesis, Vrije Universiteit Brussel, Belgium.

Wellstead, P. E. (1981). Nonparametric methods of system identification. *Automatica*, 17(1):55–69.

Wernholt, E. (2004). *On Multivariable and Nonlinear Identification of Industrial Robots.* Licentiate thesis no. 1131, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden.

Wernholt, E. and Gunnarsson, S. (2004). On the use of a multivariable frequency response estimation method for closed loop identification. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 827–832, Nassau, Bahamas.

Wernholt, E. and Gunnarsson, S. (2006). Detection and estimation of nonlinear distortions in industrial robots. In *Proc. 23rd IEEE Instumentation and Measurement Technology Conference*, pages 1913–1918, Sorrento, Italy.

Paper C

Experimental Comparison of Methods for Multivariable Frequency Response Function Estimation

Edited version of the paper:

Wernholt, E. and Moberg, S. (2007a). Experimental comparison of methods for multivariable frequency response function estimation. Technical Report LiTH-ISY-R-2827, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Experimental Comparison of Methods for Multivariable Frequency Response Function Estimation

Erik Wernholt[1] and Stig Moberg[1,2]

[1]Dept. of Electrical Engineering,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: {erikw,stig}@isy.liu.se.

[2]ABB AB – Robotics,
SE–721 68 Västerås, Sweden.

Abstract

Nonparametric estimation methods for the multivariable frequency response function are experimentally evaluated using closed-loop data from an industrial robot. Three classical estimators (H_1, joint input-output, arithmetic mean) and two estimators based on nonlinear averaging techniques (harmonic mean, geometric/logarithmic mean) are considered. The estimators based on nonlinear averaging give the best results, followed by the arithmetic mean estimator, which gives a slightly larger bias. The joint input-output estimator, which is asymptotically unbiased in theory, turns out to give large bias errors for low frequencies. Finally, the H_1 estimator gives the largest bias for all frequencies.

Keywords: System identification, frequency response methods, multivariable systems, non-parametric identification, closed-loop identification, industrial robots

1 Introduction

Nonparametric estimates of the frequency response function (FRF) give valuable information about the dynamics of a system and are often used as an intermediate step in a parametric identification process to assess the quality of the measurements and visualize the complexity of the modeling problem. In a second step, a parametric model can be estimated, either by, 1) treating the estimated nonparametric FRF as a measurement and

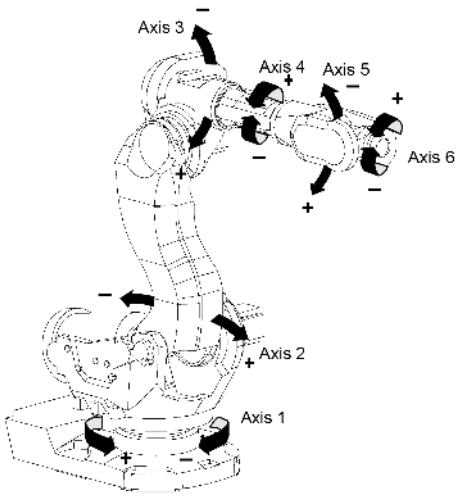


Figure 1: *The ABB manipulator IRB6600.*

minimize the discrepancy between the model and the estimated FRF as is done in, e.g., experimental modal analysis (Verboven, 2002), or, 2) directly from the measured input and output data, (Ljung, 1999; Pintelon and Schoukens, 2001).

This paper studies five different estimators: H_1, *joint input-output (JIO)*, *arithmetic mean (ARI)*, *logarithmic mean (LOG)*, and *harmonic mean (HAR)*. These estimators will be described in more detail in the next section. The experimental comparison of the estimators will be carried out using measurements from the first three axes of an industrial robot in the ABB IRB6600 series, see Figure 1. The robot application is interesting since it gives many challenging problems for system identification methods. The robot dynamics is multivariable, nonlinear, oscillatory, and, moreover, data must be collected in closed loop. Examples when nonparametric FRF estimates are used for the identification of parametric robot models can be found in, e.g., Wernholt and Moberg (2007), Öhr et al. (2006), Berglund and Hovland (2000), and Khorrami et al. (1995).

The five estimators differ in requirements on the measurement setup and signal-to-noise ratio (SNR), as well as in bias and variance properties. Using nonparametric frequency-domain methods for closed-loop data might give biased estimates. The JIO estimator is asymptotically unbiased, and the bias of the other estimators varies depending on the SNR. A recent study on the first three estimators can be found in Wernholt and Gunnarsson (2007). See also Pintelon and Schoukens (2001), Guillaume (1998), and Verboven (2002).

2 Measurement Setup

The open-loop system to be identified is unstable, which makes it necessary to collect data while the robot controller is running in closed loop. Consider therefore the setup in Figure 2, where the controller takes as input the difference between the reference signal

r and the measured and sampled output y, and u is the input. The disturbance v contains various sources of noise and disturbances.

Remark C.1. Figure 2 is a simplified block diagram since some disturbances enter early in the robot system, such as motor torque ripple (Gutt et al., 1996).

An experimental control system is used, which enables the use of off-line computed reference signals for each motor controller. The experiments are performed by first moving the robot arm into a position and then applying a speed reference signal to the robot controller. The resulting motor torques (actually the torque reference to the torque controller) and angular positions are sampled and stored. The measured angular positions are then filtered and differentiated to obtain estimates of the motor angular speeds, which are here considered as the output signals.

To avoid leakage effects in the discrete Fourier transform (DFT), which is used by the estimation methods, the excitation signal, r, is assumed to be N_P-periodic, $r(t+N_P T_s) = r(t)$ with T_s the sampling period, and an integer number of periods of the steady-state response are collected.

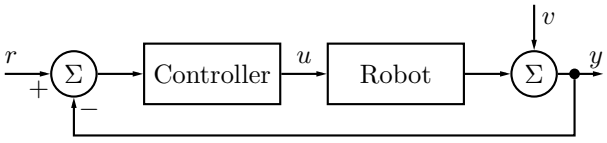


Figure 2: *Closed loop measurement setup.*

3 FRF Estimation Methods

3.1 Basic Idea

Consider the DFTs of the input, $U(\omega_k)$, output, $Y(\omega_k)$, and reference, $R(\omega_k)$, for the N_P frequencies $\omega_k = k\frac{2\pi}{N_P T_s}$, $k = 1, 2, \ldots, N_P$. Assuming a linear system and periodic data, the following mapping holds exactly in the noise-free case

$$Y(\omega_k) = G(\omega_k)U(\omega_k), \tag{1}$$

where $G(\omega_k) \in \mathbb{C}^{n_y \times n_u}$ is the FRF of the linear system. To be able to estimate $G(\omega_k)$ from data, at least n_u different experiments are needed. The data vectors from $n_e \geq n_u$ different experiments are collected into matrices (bold-face in the sequel) where each column corresponds to one experiment. The input-output relation can then be written as

$$\mathbf{Y}(\omega_k) = G(\omega_k)\mathbf{U}(\omega_k), \tag{2}$$

where $\mathbf{U}(\omega_k) \in \mathbb{C}^{n_u \times n_e}$ and $\mathbf{Y}(\omega_k) \in \mathbb{C}^{n_y \times n_e}$. If $\mathbf{U}(\omega_k)$ has rank n_u, an estimate of $G(\omega_k)$ can be formed by, e.g., using the H_1 *estimator* (Guillaume et al., 1996; Pintelon and Schoukens, 2001; Verboven, 2002)

$$\widehat{G}^{\mathrm{H}_1}(\omega_k) = \mathbf{Y}(\omega_k)\mathbf{U}^H(\omega_k)[\mathbf{U}(\omega_k)\mathbf{U}^H(\omega_k)]^{-1}, \tag{3}$$

where $(\cdot)^H$ denotes complex conjugate transpose.

Disturbances as well as nonlinearities will modify the mapping (1), which in turn will introduce errors in the estimated FRF. For a nonlinear system, we are looking for a linear approximation of the input-output behavior. The properties of such approximations have been extensively studied in, e.g., Schoukens et al. (2005) and Dobrowiecki and Schoukens (2007). See also Wernholt and Gunnarsson (2006) for some results on how nonlinearities affect the FRF estimation of industrial robots.

3.2 Some Classical Estimators

For the single input, single output (SISO) case, a number of different nonparametric FRF estimators have been suggested in the literature (see, e.g., Guillaume et al., 1992; Pintelon and Schoukens, 2001; Wellstead, 1981), which all have different properties regarding bias and variance. These estimators can often be generalized to the multiple input, multiple output (MIMO) case, as for the H_1 estimator (3).

If the reference signal is measured, an asymptotically ($n_e \to \infty$) unbiased estimator has been proposed in Wellstead (1981), which can be generalized to the MIMO case as (Cobb and Mitchell, 1990; Verboven, 2002)

$$\widehat{G}^{\mathrm{JIO}} = \mathbf{Y}\mathbf{R}^H \left[\mathbf{U}\mathbf{R}^H\right]^{-1}, \tag{4}$$

where JIO stands for *joint input-output estimator* (cf. Ljung, 1999, p. 438). (For notational simplicity, the frequency argument will be omitted when not explicitly needed.) This estimator is based on the instrumental variables principle and is also called H_{IV}, H_s, H_c, or 3-channel FRF estimator in the literature (Cobb and Mitchell, 1990; Verboven, 2002).

In this paper, we will assume that n_e is a multiple of n_u such that the DFT matrices can be partitioned into M blocks of size $n_u \times n_u$ as

$$\mathbf{R}(\omega_k) = \begin{bmatrix}\mathbf{R}^{[1]}(\omega_k) & \dots & \mathbf{R}^{[M]}(\omega_k)\end{bmatrix}, \tag{5}$$

and similarly for $\mathbf{U}(\omega_k)$ and $\mathbf{Y}(\omega_k)$. If the same excitation is used in all blocks with synchronized measurements (e.g., by using multiple periods and treating each period as a "block"), then the JIO estimator reduces to the *errors-in-variables*[1] *(EIV) estimator* (Guillaume et al., 1996; Pintelon and Schoukens, 2001)

$$\widehat{G}^{\mathrm{EIV}} = \left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{Y}^{[m]}\right]\left[\frac{1}{M}\sum_{m=1}^{M}\mathbf{U}^{[m]}\right]^{-1}. \tag{6}$$

The EIV estimator does not require the reference signal to be known and this is also the maximum likelihood estimator for Normal distributed noise (Guillaume et al., 1996). The estimator is therefore often regarded as the optimal one to use, given that the measurements are synchronized and the same excitation is used in all blocks.

Another useful estimator is the *arithmetic mean (ARI) estimator* (Guillaume, 1998; Pintelon and Schoukens, 2001)

$$\widehat{G}^{\mathrm{ARI}} = \frac{1}{M}\sum_{m=1}^{M}\widehat{G}^{[m]}, \tag{7}$$

[1] Measurements $Y = GU + V_y$ and $U + V_u$ with measurement noise V_u and V_y.

where

$$\widehat{G}^{[m]} = \mathbf{Y}^{[m]}[\mathbf{U}^{[m]}]^{-1}. \tag{8}$$

In case of measurement noise on the input, or closed-loop data, this estimator gives less bias than the H_1 estimator (Pintelon and Schoukens, 2001).

3.3 Estimators Based on Nonlinear Averaging Techniques

The averaging in (7) can be generalized to include nonlinear averaging techniques. This has been studied in Guillaume et al. (1992) for SISO systems and Guillaume (1998) extends some of these results to MIMO systems. The FRF is then estimated as

$$\widehat{G} = g^{-1}\left(\frac{1}{M}\sum_{m=1}^{M} g\left(\widehat{G}^{[m]}\right)\right),$$

where $z = g(x)$ is some nonlinear function with inverse $x = g^{-1}(z)$. Here, we will consider the following functions

$$g(x) = \begin{cases} x & \text{arithmetic mean,} \\ \log x & \text{logarithmic mean,} \\ x^{-1} & \text{harmonic mean.} \end{cases}$$

In Guillaume et al. (1992), the properties of FRF estimators based on nonlinear averaging are studied and compared with the classical estimators. They include analytical bias expressions for a number of different estimators when the measurement noise in the input and output signals are uncorrelated (i.e., not a closed-loop setup). Their conclusion is that the use of many classical estimators, such as H_1, is not advisable in most practical situations due to larger bias errors. For non-synchronous measurements (where EIV is not applicable), they propose to use estimators based on nonlinear averaging, where the logarithmic mean estimator (also called geometric mean) performs well in all their analyzed situations. This estimator is also particularly robust to outliers.

Using the complex logarithm, $\log x = \log|x| + j\arg x$, will introduce some problems when averaging the phase, due to the phase wrapping at $-\pi$ ($-\pi < \arg x \leq \pi$). A solution to this problem is given in Guillaume et al. (1992) for the scalar case by multiplying each $\widehat{G}^{[m]}(\omega_k)$ by $e^{-j\varphi(\omega_k)}$ before the averaging, and afterward multiply the result by $e^{j\varphi(\omega_k)}$, where the phases $\varphi(\omega_k)$ are obtained using some other estimator. The observations are then close to the positive real axis when performing the averaging.

For multivariable systems, the nonlinear averaging functions can either be applied element-wise or by treating them as matrix functions. The element-wise approach gives worse estimates, which has been showed in Guillaume (1998) and also seen in our experimental results. The matrix function approach $g(A)$, where $A \in \mathbb{C}^{n\times n}$, can be described by using an eigenvalue decomposition

$$A = V\Lambda V^{-1},$$

where V contains the n orthogonal (if $A = A^T$) eigenvectors and Λ is a diagonal matrix, $\Lambda = \operatorname{diag}\{\lambda_l\}_{l=1}^{n}$, with the n eigenvalues. A matrix function $g(A)$ is then given by

$$g(A) = Vg(\Lambda)V^{-1} = V\operatorname{diag}\{g(\lambda_l)\}_{l=1}^{n}V^{-1}. \tag{9}$$

Using this, the *logarithmic mean (LOG) estimator* is finally given by

$$\widehat{G}^{\text{LOG}} = P_1^{-1} \exp\left(\frac{1}{M}\sum_{m=1}^{M} \log\left(P_1 \widehat{G}^{[m]}\right)\right), \tag{10}$$

where $\exp$ and $\log$ are treated as matrix functions (9). The matrix P_1 is used to avoid the phase wrapping problems. Here, the matrix is selected as

$$P_1 = V^{[1]} \operatorname{diag}\left\{e^{-j \arg \lambda_l^{[1]}}\right\}_{l=1}^{n} [V^{[1]}]^{-1}, \tag{11}$$

where $\widehat{G}^{[1]} = V^{[1]}\Lambda^{[1]}[V^{[1]}]^{-1}$ (see Guillaume, 1998).

The last estimator to consider, the *harmonic mean (HAR) estimator*, is given by

$$\widehat{G}^{\text{HAR}} = \left[\frac{1}{M}\sum_{m=1}^{M}\left[\widehat{G}^{[m]}\right]^{-1}\right]^{-1}. \tag{12}$$

Since $[\widehat{G}^{[m]}]^{-1} = \mathbf{U}^{[m]}[\mathbf{Y}^{[m]}]^{-1}$, this actually corresponds to calculating the arithmetic mean of the output-input behavior, and then taking the inverse ($n_y = n_u$ is required). The output DFT matrices, $\mathbf{Y}^{[m]}$, $m = 1, \ldots, M$, are then inverted, compared to inverting the input DFT matrices when using the ARI estimator (7). Note that the HAR estimator also uses the matrix function approach.

In this paper, these estimators will be evaluated experimentally. For a theoretical discussion about the estimators and their properties, see the included references in this paper as well as the discussion in Section 5.3.

4 Excitation Signals

In general the choice of excitation signal offers a large freedom in terms of frequency contents, magnitude, and so on, as long as the matrix $\mathbf{U}(\omega_k)$ has rank n_u. In this paper, the *orthogonal random phase multisine* signal will be used. This signal has been suggested in Dobrowiecki et al. (2005) and Dobrowiecki and Schoukens (2005, 2007) as input in the open-loop case to minimize the FRF uncertainty, given input amplitude constraints. Here, the orthogonal multisine signal will be used as reference signal, which corresponds to an optimal experiment design given output amplitude constraints. Each block in (5) is then given by

$$\mathbf{R}^{[m]}(\omega_k) = \mathbf{R}^{[m]}_{\text{diag}}(\omega_k)\mathbf{T}, \quad m = 1, \ldots, M, \tag{13}$$

with

$$\mathbf{R}^{[m]}_{\text{diag}}(\omega_k) = \operatorname{diag}\left\{R_l^{[m]}(\omega_k)\right\}_{l=1}^{n_y},$$

a diagonal matrix where each $R_l^{[m]}(\omega_k)$ is a random phase multisine signal, and $\mathbf{T}$ is an orthogonal matrix. A scalar random phase multisine signal $r(t)$ can be written as

$$r(t) = \sum_{k=1}^{N_f} A_k \cos(\omega_k t + \phi_k), \tag{14}$$

with amplitudes A_k, frequencies ω_k chosen from the grid $\{\frac{2\pi l}{N_P T_s},\ l = 1, \ldots, N_P/2 - 1\}$ (N_P even), and random phases ϕ_k uniformly distributed on the interval $[0, 2\pi)$.

The optimal matrix $\mathbf{T}$, with constraints $|\mathbf{T}_{il}| \leq 1$, is given by (Dobrowiecki and Schoukens, 2007)

$$\mathbf{T}_{il} = e^{\frac{2\pi j}{n_u}(i-1)(l-1)}. \tag{15}$$

The number of frequencies as well as the amplitude spectrum will affect the quality of the FRF estimate. Using too many frequencies will give a low SNR, which increases both the bias and the variance in the estimate. The FRF estimate can be improved by averaging over multiple blocks and/or periods. The covariance matrix can then also be estimated.

For a linear system, averaging over different periods is sufficient, whereas for a nonlinear system, it is essential to average over blocks where $\mathbf{R}_{\text{diag}}$ in each block should have different realizations of the random phases. The reason is that nonlinearities otherwise will distort the estimate and give a too low uncertainty estimate. To reduce these nonlinear distortions, one should also excite only odd frequencies (odd l in the frequency grid for (14)). For details, see Dobrowiecki and Schoukens (2007) and Pintelon and Schoukens (2001, Chap. 3). For the industrial robot, the nonlinearities cause large distortions (Wernholt and Gunnarsson, 2006). Averaging over multiple blocks is therefore important, so the EIV estimator is not applicable here.

5 Experimental Results

5.1 Measurement Data

As excitation, the orthogonal random phase multisine signal is used as motor speed reference r with period time $T_0 = N_P T_s = 10\,\text{s}$. A flat amplitude spectrum is used for 100 log-spaced odd frequencies between 1 and 60 Hz (i.e., the grid $1.1, 1.3, \ldots, 59.9\,\text{Hz}$). To reduce nonlinear effects, $M = 100$ blocks with different realizations of the random phases are used. The magnitude of $\mathbf{R}_{\text{diag}}^{[m]}$ is the same in all blocks. The commanded input u and measured output y are both sampled at $f_s = 1/T_s = 2\,\text{kHz}$.

To reduce the effect of static friction, it is common to use an excitation signal that avoids zero velocity as much as possible. Therefore the reference is given by $r(t) = r_{ms}(t) + r_s(t)$, with $r_{ms}(t)$ the multisine signal and $r_s(t)$ a single sinusoid with frequency $0.7\,\text{Hz}$. The amplitude of $r_s(t)$ is selected sufficiently larger than $\max_t |r_{ms}(t)|$ to avoid passing through zero velocity due to the multisine. Reducing the effect of static friction is really important for this type of system, as has been pointed out in Wernholt and Gunnarsson (2006). A drawback is that the sinusoid introduces large errors at a few of its overtones due to nonlinearities in the system. The first five overtones, $2.1, 3.5, 4.9, 6.3, 7.7\,\text{Hz}$, are therefore removed in the resulting FRF estimate.

The FRF estimates turn out to have fairly large errors at low frequencies. Therefore, 10 additional blocks of data are collected using only four frequencies in the multisine ($1.1, 2.7, 4.3, 5.9\,\text{Hz}$). This improves the SNR, which in turn will reduce the error at these frequencies.

In the figures, the FRF estimates from motor torque to motor acceleration are plotted. That avoids the integrator ($-20\,\text{dB/decade}$) such that resonances appear more clearly.

5.2 Evaluation of FRF Estimators

The five estimators from Section 3 (H_1, JIO, ARI, HAR, LOG) will now be used for FRF estimation. Since the true system is unknown, we cannot calculate the bias of the estimators. The assessment of the FRF estimates will therefore be partly based on experience from robot experts, as well as knowledge about the system and theoretical results.

The magnitude and phase of the resulting estimates, using 100 blocks of measurement data, are shown in Figure 3. The magnitude estimates agree quite well for frequencies above 5 Hz, except for the H_1 estimator which tends to give much lower values. For low frequencies (1–5 Hz), the estimates fluctuate quite much. By comparing the estimates with the ones using 10 blocks with four frequencies (circles in Figure 3), it is evident that the JIO and H_1 estimators give large errors there. The estimators based on nonlinear averaging, HAR and LOG, give almost the same estimate and agree very well with the estimates using four frequencies. The ARI estimator works much better than JIO and H_1, but does not match the performance of HAR and LOG (compare, e.g., element (2,3)). Considering the phase estimates, all five estimators agree remarkably well for frequencies above 5 Hz. For low frequencies, the HAR and LOG estimators give much better estimates than the others.

In the simulation study in Wernholt and Gunnarsson (2007) it was noted that the uncertainties in nonparametric FRF estimates usually are non-symmetric, even if the true system is symmetric. This can be seen in Figure 3 as well, where one standard deviation of the ARI estimator is included. Compare, e.g., elements (1,2) and (2,1), or elements (2,3) and (3,2). For small elements, the relative error is also larger, which can be seen by comparing elements (1,1) and (2,1). These effects are inherent in all the studied estimation methods. From the standard deviation, it is also evident that the classical estimators (H_1, JIO, and ARI) cannot handle a low SNR (at low frequencies and resonances) as good as the estimators based on nonlinear averaging techniques (HAR and LOG).

The estimates in Figure 3 are very clear and accurate, but measuring 100 blocks takes almost 3 hours in our experimental setup (approximately 1 minute per block of effective measurements, plus additional time for saving data, etc.). It is therefore interesting to study the quality of the estimates using less data.

Figure 4 shows the estimates using only 5 blocks, where LOG using 100 blocks is included for comparison. For low frequencies, all estimators give biased estimates in element (2,1). Otherwise, the same comments as for Figure 3 are valid here as well. To be noted is that ARI fluctuates quite much at the resonances. Since the reference signal has a flat amplitude spectrum, the input is small at the resonances. The fluctuations are therefore probably caused by inverting a matrix that is quite sensitive to noise.

Next, we will compare the different estimators as a function of the number of blocks. Let $\widehat{G}_M$ denote the FRF estimate calculated from M blocks of data. Consider now the difference between $\widehat{G}^{\text{LOG}}_{100}$ (the LOG estimate in Figure 3) and the estimates $\widehat{G}^{\text{H}_1}_M$, $\widehat{G}^{\text{JIO}}_M$, $\widehat{G}^{\text{ARI}}_M$, $\widehat{G}^{\text{HAR}}_M$, and $\widehat{G}^{\text{LOG}}_M$, where the number of blocks, M, will vary between 1 and 100. To measure the difference between two estimates, the following cost is used

$$c_{log}(G^1, G^2) = \left(\sum_{i,j,k} \left| \log G^1_{ij}(\omega_k) - \log G^2_{ij}(\omega_k) \right|^2 \right)^{1/2}. \tag{16}$$

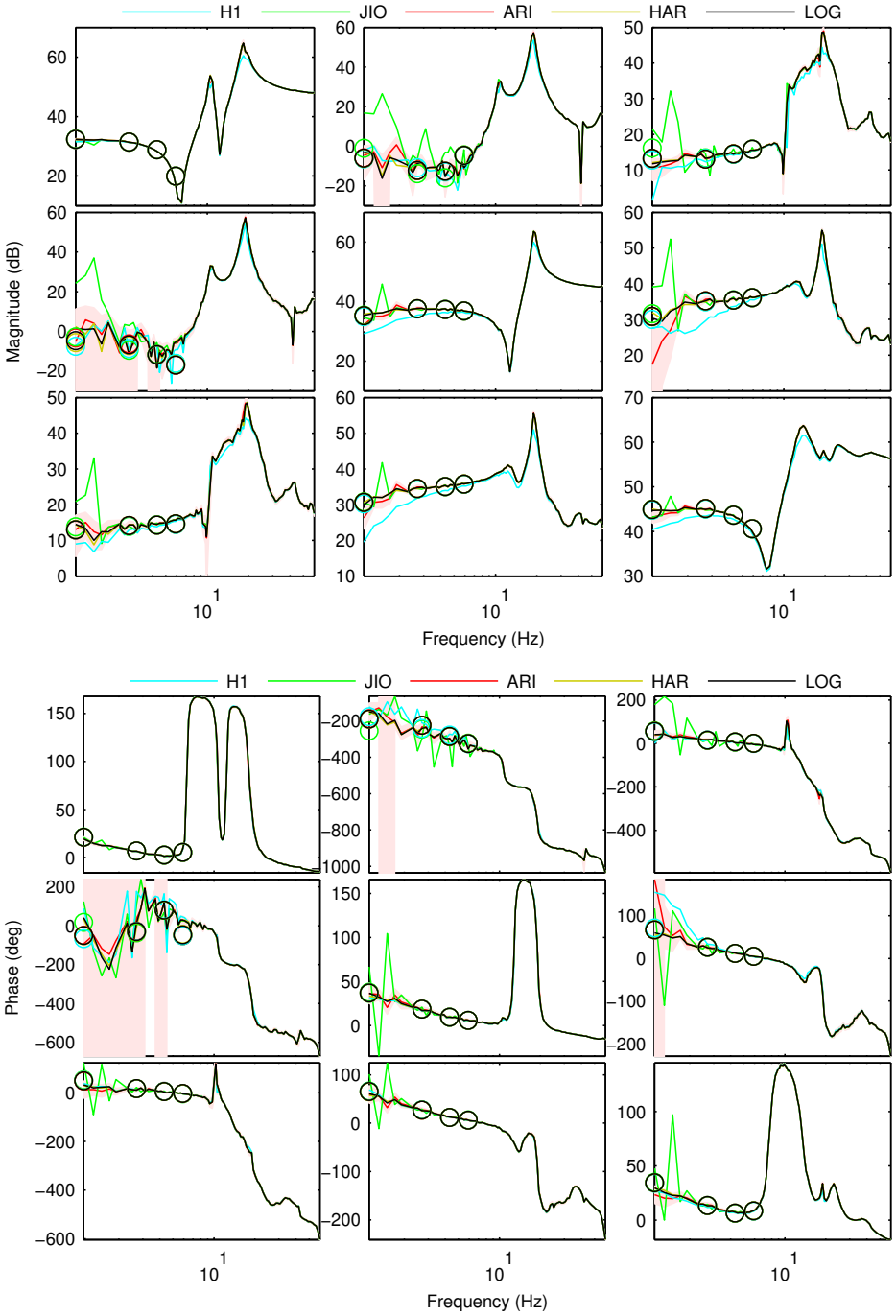


Figure 3: *Estimated FRFs using 100 blocks of data. ARI is plotted with one standard deviation (shaded). Circles: estimates using* 10 *blocks with four frequencies.*

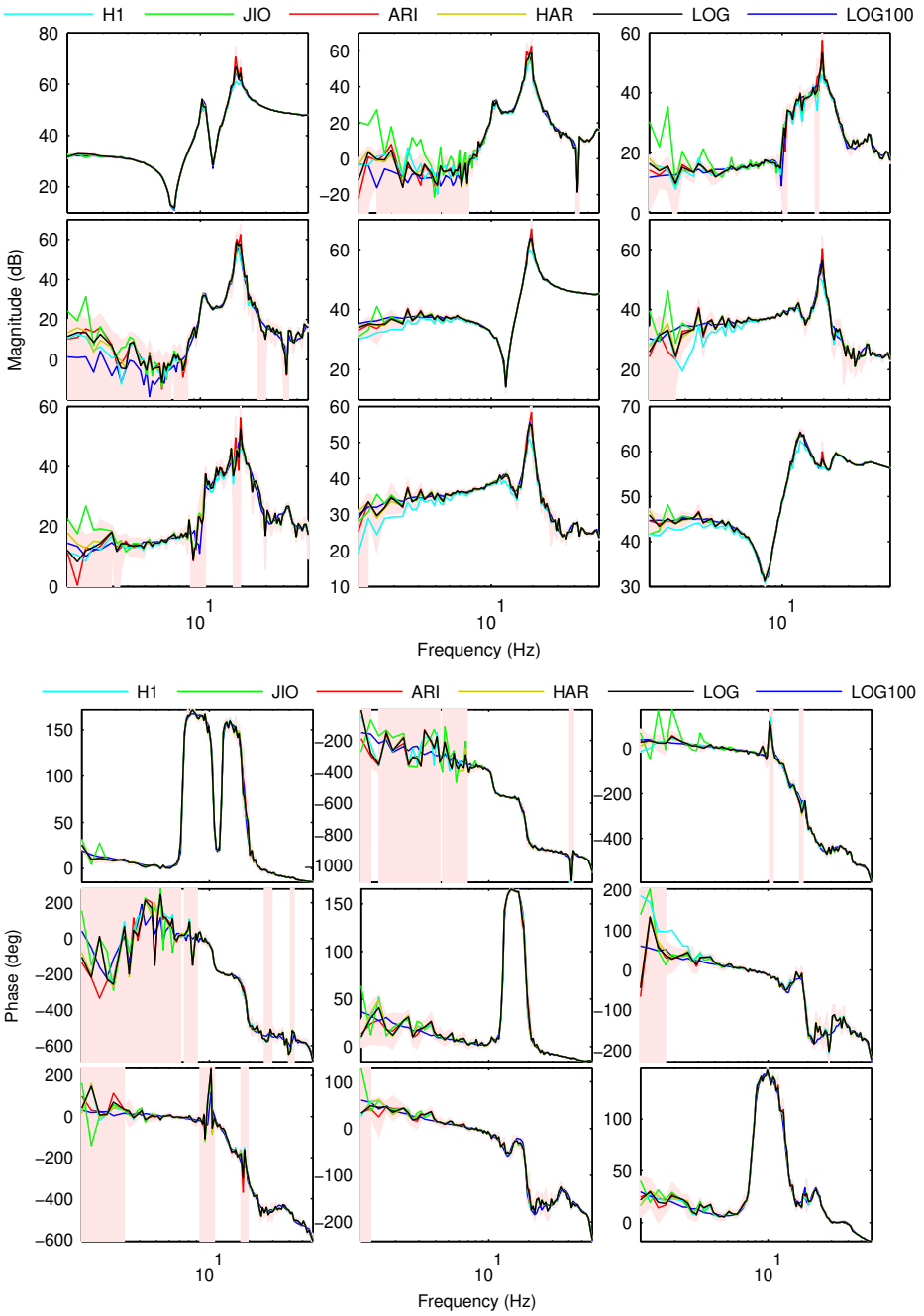


Figure 4: *Estimated FRFs using* 5 *blocks of data. For comparison is LOG with* 100 *blocks included as well. ARI is plotted with one standard deviation (shaded).*

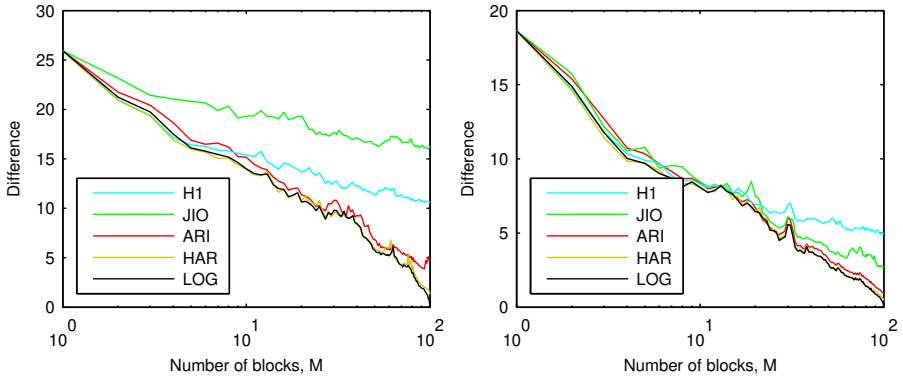


Figure 5: *Difference* $c_{log}(\widehat{G}_{100}^{LOG}, \widehat{G}_M)$ *for* 1–60 *Hz (left) and* 5–60 *Hz (right) as a function of the number of blocks,* M*.* $\widehat{G}_M$ *is replaced by each of the five estimators.*

In addition, phase wrapping problems in the complex logarithm are handled to ensure that $|\arg G_{ij}^1(\omega_k) - \arg G_{ij}^2(\omega_k)| \leq \pi$. The difference between $\widehat{G}_{100}^{\text{LOG}}$ and the other estimates can be seen in Figure 5.

First, note that all estimators coincide for $M = 1$. When M is increased, so is the difference between the estimators. The most interesting thing to note from Figure 5 might be that HAR and LOG give better results than the other estimators for all $M > 1$. HAR might be slightly better than LOG for small M, but they otherwise give approximately the same performance. What might be surprising, given the previous results in Figures 3 and 4, is that H_1 gives almost the same cost as LOG and HAR for low values of M. The explanation can be found in Figure 4 where all estimates differ from $\widehat{G}_{100}^{\text{LOG}}$ for low frequencies. Another interesting thing to note is that ARI gives a larger cost than H_1 when $M \leq 10$. The main reason is the large errors at the resonances for ARI. JIO gives the worst overall performance, mainly due to the errors at low frequencies. Considering only the interval 5–60 Hz in Figure 5, JIO and ARI give approximately the same cost for $M \leq 30$, except that the cost for JIO fluctuates more.

Since we are comparing the estimates with $\widehat{G}_{100}^{\text{LOG}}$, the fact that LOG and HAR are the best comes as no surprise when M is close to 100. However, this is also the case if we compare with $\widehat{G}_{100}^{\text{ARI}}$ (for $M \leq 50$) or some of the other estimators. The previous conclusion that LOG and HAR are the best estimators therefore holds also in this comparison. Finally, this section is ended by Figure 6, where estimates using LOG for some values of M, including $M = 1$, are given. As can be seen, the fluctuations are fairly large when using only one block of data.

5.3 Discussion

From theoretical studies, as well as experimental experience, it is known that the bias depends on the SNR, such that the bias is reduced when the SNR is increased. The data set with 10 blocks and four frequencies can therefore be considered to give a smaller bias. This can also be seen in Figure 3 where all five estimators give approximately the

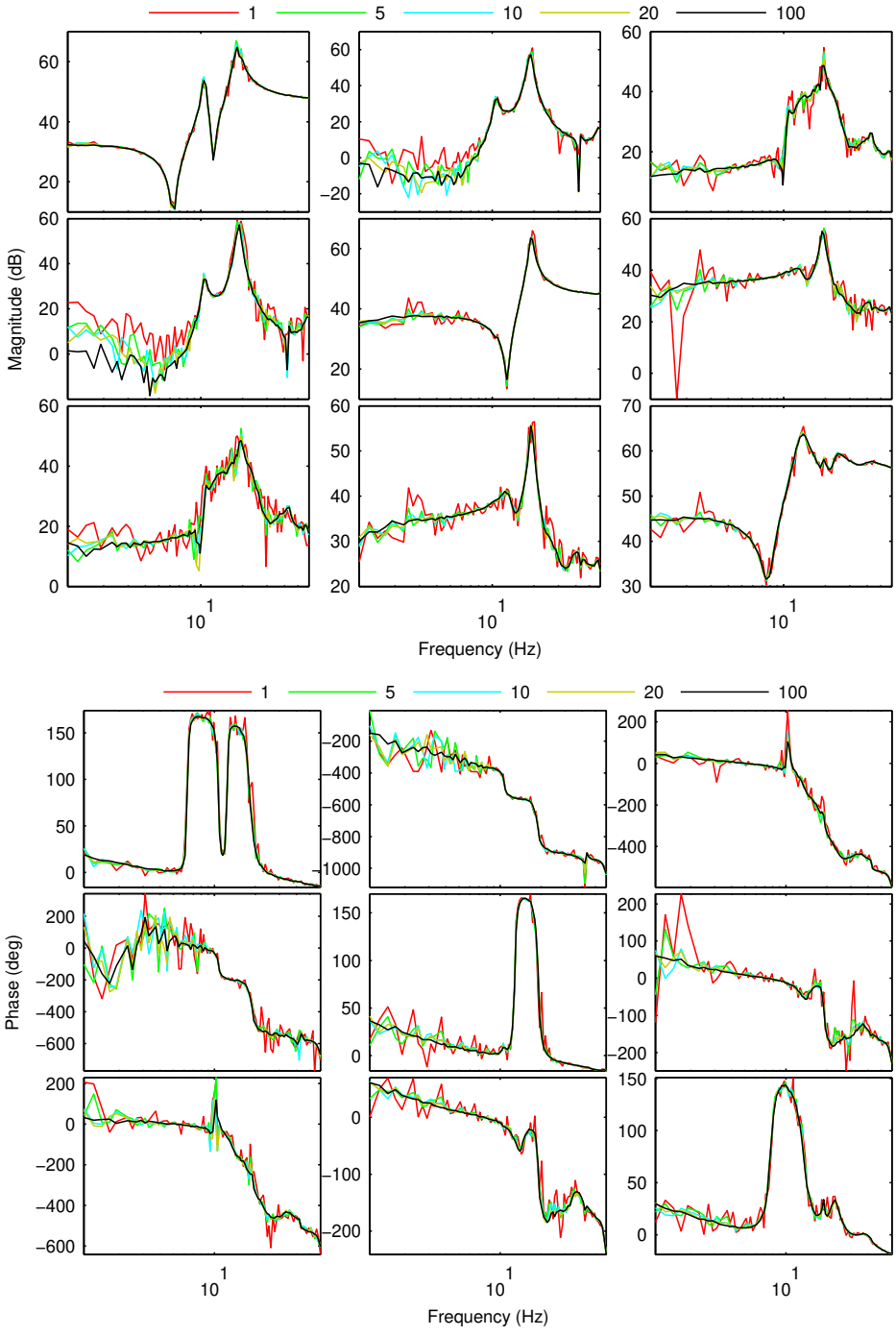


Figure 6: *Estimated FRFs using the LOG estimator for different number of blocks.*

same estimate. Since we are dealing with linear approximations of a nonlinear system, a comparison of estimates using different types of excitation signals should however be done with some care. The best linear approximation might be different when changing the amplitude spectrum and/or the signal shape. This can, e.g., be seen in Wernholt and Gunnarsson (2006) where estimates are compared when using different excitation amplitudes as well as with and without a filtered square wave (to reduce the effect of static friction).

The H_1, ARI, and JIO estimators have been analyzed in, e.g., Wernholt and Gunnarsson (2007), Pintelon and Schoukens (2001) and Verboven (2002). The H_1 and ARI estimators are known to give biased estimates for closed-loop data. The bias of the H_1 estimator is much larger, except for small SNRs where the ARI estimator deteriorates. The JIO estimator is asymptotically unbiased ($M \to \infty$) and is therefore expected to give the best performance when the number of blocks increases. This is also the conclusion in Wernholt and Gunnarsson (2007), where the JIO, ARI, and H_1 estimators are compared in a simulation study with a similar setup as in this paper, but with a linear robot model. In our experimental results, the JIO estimator does not seem to perform that well, at least not if the low-frequency behavior is considered.

A limitation with the JIO estimator, besides that an additional signal is needed, is that the whole measurement setup must be linear, including the controller. This is due to the fact that $\widehat{G}^{\text{JIO}}$ actually is based on H_1 estimates of the FRFs from r to y, $\widehat{G}_{ry}$, and r to u, $\widehat{G}_{ru}$, as $\widehat{G}^{\text{JIO}} = \widehat{G}_{ry}\widehat{G}_{ru}^{-1}$. Both these FRFs incorporate the whole measurement setup, including the controller. In our application, the controller, e.g., has a saturation to ensure that the motors are not overloaded. The other methods considered in this paper are only affected by nonlinearities in the input-output system.

The properties of the LOG and HAR estimators have been analyzed in e.g., Guillaume et al. (1992), Guillaume (1998), and Pintelon and Schoukens (2001). The HAR estimator is particularly useful when the input SNR is low since the input DFT matrix $\mathbf{U}^{[m]}$ is not inverted. In our setup we are using the orthogonal multisine signal as reference $\mathbf{R}^{[m]}$, so the output $\mathbf{Y}^{[m]}$ will be approximately orthogonal as well, or at least well-conditioned. The input $\mathbf{U}^{[m]}$, on the other hand, can be quite ill-conditioned for certain frequencies (e.g., close to resonances, or for low frequencies since we are using a flat speed reference signal). The HAR estimator is therefore well suited for our setup. The LOG estimator is inherently robust to outliers since the logarithmic function attenuates the effect of occasional large errors. For the classical estimators, a large number of blocks are needed to average out such effects. This can, e.g., be noted with the ARI estimator at the resonances in Figure 4.

As was mentioned in Section 5.1, it is important to compensate for the static friction. Using the single sinusoid caused large errors at some of its overtones due to nonlinearities in the system. This might depend on the fact that its phase is the same in all experiments. Different ways of reducing the effects of static friction would therefore be interesting to study.

As future work remains a thorough simulation study using a realistic nonlinear robot model where friction as well as other nonlinear effects in actuators and sensors are added, such as torque ripple (Gutt et al., 1996) and resolver ripple (Hanselman, 1990). The bias can then be studied for the different estimators as a function of the excitation signal and the nonlinearities.

6 Conclusion

This paper has dealt with nonparametric multivariable FRF estimation using closed-loop data from an industrial robot. Three classical estimators (H_1, JIO, ARI) and two estimators based on nonlinear averaging techniques (HAR, LOG) have been compared. The orthogonal random phase multisine signal has been used, where the number of experiments have been varied to see how the estimators depend on the amount of data. The estimators based on nonlinear averaging give the best results, followed by the arithmetic mean estimator, which gives a slightly larger bias. The joint input-output estimator, which is asymptotically unbiased in theory, turns out to give large bias errors for low frequencies. Finally, the H_1 estimator gives the largest bias for all frequencies.

References

Berglund, E. and Hovland, G. E. (2000). Automatic elasticity tuning of industrial robot manipulators. In *Proc. 39th IEEE Conference on Decision and Control*, pages 5091–5096, Sydney, Australia.

Cobb, R. and Mitchell, L. (1990). Estimates of variance and confidence bands for the three-channel frequency response function estimator. *International Journal of Analytical and Experimental Modal Analysis*, 5(3):185–194.

Dobrowiecki, T. and Schoukens, J. (2005). Measuring linear approximation to weakly nonlinear MIMO systems. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Dobrowiecki, T. and Schoukens, J. (2007). Measuring a linear approximation to weakly nonlinear MIMO systems. *Automatica*, 43(10):1737–1751.

Dobrowiecki, T., Schoukens, J., and Guillaume, P. (2005). Optimized excitation signals for MIMO frequency response function measurments systems. In *Proc. 22nd IEEE Instrumentation and Measurement Technology Conference*, pages 1872–1877.

Guillaume, P. (1998). Frequency response measurements of multivariable systems using nonlinear averaging techniques. *IEEE Transactions on Instrumentation and Measurement*, 47(3):796–800.

Guillaume, P., Pintelon, R., and Schoukens, J. (1996). Accurate estimation of multivariable frequency response functions. In *Proc. 13th IFAC World Congress*, pages 423–428, San Francisco, California.

Guillaume, P., Pintelon, R., and Schoukens, R. (1992). Nonparametric frequency response function estimators based on nonlinear averaging techniques. *IEEE Transactions on Instrumentation and Measurement*, 41(6):739–746.

Gutt, H.-J., Scholl, F. D., and Blattner, J. (1996). High precision servo drives with DSP-based torque ripple reduction. In *Proc. IEEE AFRICON, 1996*, volume 2, pages 632–637.

Hanselman, D. (1990). Resolver signal requirements for high accuracy resolver-to-digital conversion. *IEEE Transactions on Industrial Electronics*, 37(6):556–561.

Khorrami, F., Jain, S., and Tzes, A. (1995). Experimental results on adaptive nonlinear control and input preshaping for multi-link flexible manipulators. *Automatica*, 31(1):83–97.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach*. IEEE Press, New York.

Pintelon, R. and Schoukens, J. (2001). Measurement of frequency response functions using periodic excitations, corrupted by correlated input/output errors. *IEEE Transactions on Instrumentation and Measurement*, 50(6):1753–1760.

Schoukens, J., Pintelon, R., Dobrowiecki, T., and Rolain, Y. (2005). Identification of linear systems with nonlinear distortions. *Automatica*, 41(3):491–504.

Verboven, P. (2002). *Frequency-domain system identification for modal analysis*. PhD thesis, Vrije Universiteit Brussel, Belgium.

Wellstead, P. E. (1981). Nonparametric methods of system identification. *Automatica*, 17(1):55–69.

Wernholt, E. and Gunnarsson, S. (2006). Detection and estimation of nonlinear distortions in industrial robots. In *Proc. 23rd IEEE Instumentation and Measurement Technology Conference*, pages 1913–1918, Sorrento, Italy.

Wernholt, E. and Gunnarsson, S. (2007). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Wernholt, E. and Moberg, S. (2007). Frequency-domain gray-box identification of industrial robots. Technical Report LiTH-ISY-R-2826, Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. *Submitted to the 17th IFAC World Congress*, Seoul, Korea.

Paper D

Estimation of Nonlinear Effects in Frequency-Domain Identification of Industrial Robots

Edited version of the paper:

> Wernholt, E. and Gunnarsson, S. (2007b). Estimation of nonlinear effects in frequency-domain identification of industrial robots. *Accepted for publication in IEEE Transactions on Instrumentation and Measurement.*

Parts of the paper in:

> Wernholt, E. and Gunnarsson, S. (2006a). Detection and estimation of nonlinear distortions in industrial robots. In *Proc. 23rd IEEE Instumentation and Measurement Technology Conference*, pages 1913–1918, Sorrento, Italy.

Preliminary version published as Technical Report LiTH-ISY-R-2740, Dept. of Electrical Engineering, Linköping University, SE–581 83 Linköping, Sweden.

Estimation of Nonlinear Effects in Frequency-Domain Identification of Industrial Robots

Erik Wernholt and Svante Gunnarsson

Dept. of Electrical Engineering,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: {erikw,svante}@isy.liu.se.

Abstract

A method for the detection and estimation of nonlinear distortions when identifying nonparametric multivariable frequency response functions (FRF) is considered. The method is successfully applied to experimental data from an industrial robot, collected in closed loop. The results show that nonlinear distortions are indeed present and cause larger variability in the FRF than the measurement noise contributions.

Keywords: Frequency response functions, multivariable systems, nonlinear distortions, non-parametric identification, industrial robots

1 Introduction

Industrial robots pose a challenging modeling problem. Usually a robot has six joints, with coupled dynamics, giving a truly multivariable system. The dynamics is nonlinear, both with respect to the rigid body dynamics and other things such as friction, torque and resolver ripple, and a transmission with backlash, hysteresis, and nonlinear stiffness. The robot arm and transmission are more or less elastic and, in addition, the data collection must usually be carried out in closed loop. The identification of such a complex system is a huge task, both in finding suitable model structures and efficient identification methods. Under certain simplifying assumptions, a subset of the parameters can be identified. Neglecting elastic effects, a nonlinear model of the rigid body dynamics can be estimated using least squares techniques. This is a much studied problem in the literature (see, for example, Kozlowski (1998) for an overview). Taking elastic effects into account makes the identification problem more involved, since now typically only a subset of the state variables are measured and one can therefore not use linear regression. A common

remedy is to study the dynamic behavior around certain positions and there estimate parametric or nonparametric (linear) models (see, e.g., Albu-Schäffer and Hirzinger, 2001; Behi and Tesar, 1991; Johansson et al., 2000; Wernholt and Gunnarsson, 2006b). One application area for linear models is control design, where a global controller is achieved through gain scheduling. The linear models could also be used for the tuning of elastic parameters (typically springs and dampers) in a global nonlinear elastic model, as in Öhr et al. (2006). For both of these application areas, it is important that the estimated linear models are accurate, or at least are delivered with some estimated uncertainty regions. In the presence of nonlinearities, these uncertainty regions tend to be underestimated. It is therefore of great importance to use an identification method that can detect the presence of nonlinearities and in addition quantify how the estimated models are affected. In this paper a method presented in, e.g., Pintelon et al. (2004a,b) and Schoukens et al. (2005) will be applied and evaluated. The aim of the method is to make it possible to detect the presence of nonlinearities and quantify how much they affect the nonparametric estimate of the frequency response function (FRF). In Pintelon et al. (2004a,b) the method was applied to electronic systems, and it will here be used to analyze closed-loop experimental data (both SISO and MIMO measurements) from an industrial robot. See also Verboven et al. (2006) for experimental results of a 2x2 mechanical system (a car body, using shakers as excitation and measuring applied forces and accelerations). An early version of the results presented here was given in Wernholt and Gunnarsson (2006a).

2 Measurement Setup

The data used in this paper are collected at ABB Robotics in Västerås, Sweden, using an experimental industrial robot system, similar to the ABB IRB6600 robot in Figure 1. Each axis of the robot arm is actuated by an electric motor via a transmission and the movements are controlled by a computer system. The first three axes will be considered in this work, giving a multivariable system with three inputs (commanded motor torques) and three outputs (motor velocities).

For this particular industrial robot, the motor angles are measured using resolvers (Hanselman, 1990). The measured angles are then filtered and differentiated to obtain estimates of the motor velocities, which are here considered as the output signals. As actuators, AC permanent magnet motors are used.

For this kind of application it is necessary to use feedback control while data are collected, both for safety reasons and in order to keep the robot around its position. Consider therefore the setup in Figure 2, where the controller takes as input the difference between the reference signal r and the measured and sampled output y, and u is the input. The measurement noise, v, contains various sources of noise and disturbances (e.g., torque and resolver ripple). An experimental control system is used, which makes it possible to use off-line computed reference signals for each motor controller. The excitation signals are applied as reference signals for the motor velocities. The experimental controller can approximately be seen as a diagonal PI-controller.

Both the resolvers and the actuators, as well as the transmission, are sources of nonlinearities. Due to non-ideal resolver characteristics, a nonlinear measurement error appears (often called resolver ripple) which is periodic for full turns of the motor (Hanselman,

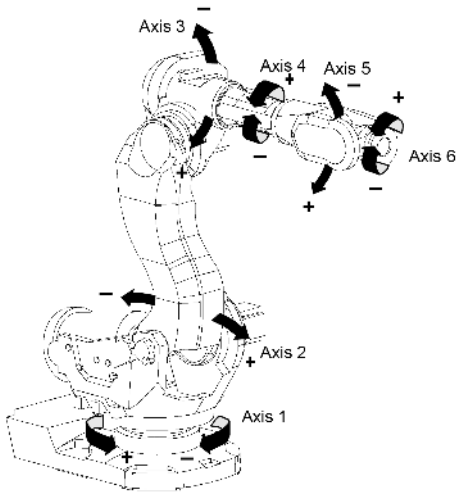


Figure 1: *The ABB IRB6600 robot.*

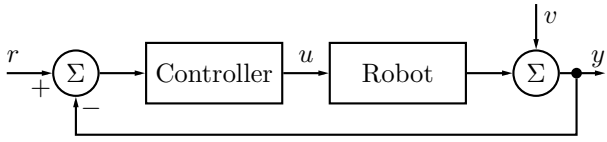


Figure 2: *Closed loop measurement setup.*

1990). AC motors give rise to a similar ripple in the motor torque, also periodic for full turns of the motor but in addition proportional to the motor current. For details, see Gutt et al. (1996). As transmission, compact gearboxes (RV gears) are used due to their low backlash, compact size, and high torque transmission. This will however introduce nonlinear stiffness, friction, and kinematic errors (see Dhaouadi et al. (2003) for an example with harmonic drives).

3 Estimation Method

In this section, the estimation method from Pintelon et al. (2004a,b) and Schoukens et al. (2005) will be described. First, in Section 3.1, some properties of nonparametric FRF estimates are introduced. The estimation method is then described in Section 3.2 for the SISO case, and in Section 3.3 for the MIMO case. Since the theoretical background to the estimation method is based on an open-loop setup, that is mainly how the material will be presented in this section as well, with some comments on the differences for the closed-loop case.

3.1 Introduction

The basic quantities for estimation of both the nonparametric FRF of the input-output relation and the nonlinear distortions are the DFTs of the input and output signals, i.e.,

$$U(\omega_k) = \frac{1}{\sqrt{N_P}} \sum_{n=1}^{N_P} u(nT_s) e^{j\omega_k n T_s}, \quad Y(\omega_k) = \frac{1}{\sqrt{N_P}} \sum_{n=1}^{N_P} y(nT_s) e^{j\omega_k n T_s}, \tag{1}$$

where T_s and N_P denote the sampling interval and the number of data, respectively. The FRF estimate is obtained by forming

$$\widehat{G}(\omega_k) = \frac{Y(\omega_k)}{U(\omega_k)}. \tag{2}$$

To avoid leakage effects in the DFTs, the input signal, $u(t)$, is assumed to be N_P-periodic and an integer number of periods, P, of the steady-state response are collected, giving $N = PN_P$ samples for each experiment. The properties of the nonparametric FRF estimate have been investigated in, e.g., Ljung (1999) and Pintelon and Schoukens (2001). Given that the true input-output relationship is linear and the output is subject to an additive zero mean disturbance,

$$y(t) = G_0 u(t) + v(t), \tag{3}$$

the analysis shows that the estimate is unbiased ($\mathrm{E}(\widehat{G}) = G_0$) and its variance decays as $1/P$ if the input signal is periodic, as stated above, and uncorrelated with the disturbance.

In this paper, the input-output relationship is nonlinear, which requires additional results. A special class of periodic input signals will then be used, namely random phase multisines, defined as

$$u(t) = \sum_{k=1}^{N_f} A_k \sin(\omega_k t + \phi_k), \tag{4}$$

with amplitudes A_k, frequencies ω_k chosen from the grid $\{\frac{2\pi l}{N_P T_s},\ l = 1, \ldots, N_P/2-1\}$, and random phases ϕ_k uniformly distributed on the interval $[0, 2\pi)$. One of the benefits from using this kind of input signal is that it enables quantification of nonlinear effects in the system. This can be done, either by: 1) looking at non-excited harmonics in the output spectrum, or 2) by performing multiple experiments with different realizations of the random phase multisine signal and then comparing FRF estimates from different periods and experiments (Pintelon et al., 2004a,b). The fundamental observation, from Pintelon and Schoukens (2001, 2002), is that using random phase multisines as input signal in an open-loop setting, with N_f sufficiently large, the FRF of a wide class of nonlinear systems[1] can be written as

$$G(\omega_k) = G_R(\omega_k) + G_S(\omega_k) + N_G(\omega_k), \tag{5}$$

with G_R the best linear approximation to the nonlinear system[2], G_S a zero mean stochastic nonlinear contribution, and N_G the measurement noise contribution. For different realizations of the random phase multisine, the stochastic nonlinear contribution, G_S, acts

[1] Systems that can be represented by a Volterra description (Pintelon and Schoukens, 2001, p. 74).

[2] For the class of Normal input signals (e.g., random phase multisine) with a given power spectrum.

as circular complex noise. Therefore G_S cannot be distinguished from the measurement noise. Still, for a given realization, G_S is a deterministic component that will look the same for different periods. The best linear approximation, G_R, depends only on the input spectrum and odd nonlinearities.

It is usually recommended to use so-called odd random phase multisines, i.e., to only excite the odd harmonics (l odd in the frequency grid). The motivation is that in that case, only odd nonlinearities contribute to the variance of G_S (at the odd frequencies). Exciting all harmonics increases the variance of G_S also at the odd frequencies due to even nonlinearities and a combination of odd and even frequencies.

For closed-loop data, the properties of the FRF are much harder to analyze. The input signal will then be correlated with the measurement noise which, even in a linear setting, will give biased FRF estimates (see, e.g., Pintelon and Schoukens, 2001; Wernholt and Gunnarsson, 2007). What is worse, even if a multisine excitation is used as reference signal, the input signal now also depends on the nonlinearities due to the feedback. This changes the amplitude distribution of the input signal and a theoretical analysis becomes hard. The estimation method, presented next, can still be applied to closed-loop data for detection of nonlinearities. The resulting estimates should, however, be treated with some care since the assumptions for deriving (5) are not fulfilled.

3.2 SISO Measurements

To be able to distinguish between the measurement noise and the stochastic nonlinear contributions, two different measurement strategies were proposed in Pintelon et al. (2004a,b) for the characterization of operational amplifiers. The first strategy uses a single experiment with $P \geq 6$ periods of the steady-state response. The excitation signal is an odd random phase multisine signal, where certain odd frequencies are eliminated as well. The stochastic nonlinear contributions are then found by analyzing the non-excited frequencies, where even nonlinearities show up at even frequencies and odd nonlinearities at the non-excited odd frequencies. The strategy also includes a procedure to handle closed-loop data (the excitation is then applied as reference signal to the controller), where the problem is that the input is nonzero also at non-excited frequencies due to feedback of the nonlinear distortions. Linear contributions must then be compensated for, which is handled by interpolating the FRF, see Pintelon et al. (2004a,b) for details.

Here, their second strategy will be used, which requires $M \geq 6$ different experiments with $P \geq 2$ periods in each experiment[3]. This strategy is more time consuming and will not give a classification in odd and even degree nonlinear distortions for closed-loop data. Still, no approximation (interpolation) is used and the strategy can easily be extended to the multivariable case. The two strategies could also be combined in order to obtain the classification in odd and even degree nonlinearities.

To simplify the notation, the estimation method will now be described for the SISO case. In the MIMO case, the method is applied element-wise, see Section 3.3.

From the measurements, $M \times P$ FRFs, $\widehat{G}^{[m,p]}$, can be calculated as (see (2))

$$\widehat{G}^{[m,p]}(\omega_k) = \frac{Y^{[m,p]}(\omega_k)}{U^{[m,p]}(\omega_k)}, \tag{6}$$

[3]Needed if the nonparametric FRF estimate is used for parametric modeling to preserve the properties of the maximum-likelihood estimator (Pintelon and Schoukens, 2001).

where $U^{[m,p]}$ and $Y^{[m,p]}$ are the DFTs, see (1), of the input and output signals from period p in experiment m. In the following variance expressions, (5) is used together with the assumption that $N_G^{[m,p]}$ is independent over m and p and that $G_S^{[m,p]}$ is equal over p but independent over m.

For each experiment, one can calculate the average FRF $\widehat{G}^{[m]}$ and its sample variance $\hat{\sigma}^2_{\widehat{G}^{[m]}}$ according to

$$\widehat{G}^{[m]} = \frac{1}{P}\sum_{p=1}^{P} \widehat{G}^{[m,p]}, \tag{7}$$

$$\hat{\sigma}^2_{\widehat{G}^{[m]}} = \sum_{p=1}^{P} \frac{|\widehat{G}^{[m,p]} - \widehat{G}^{[m]}|^2}{P(P-1)}. \tag{8}$$

The FRF $\widehat{G}$ and its sample variance are obtained by averaging over the M experiments as

$$\widehat{G} = \frac{1}{M}\sum_{m=1}^{M} \widehat{G}^{[m]}, \tag{9}$$

$$\hat{\sigma}^2_{\widehat{G}} = \sum_{m=1}^{M} \frac{|\widehat{G}^{[m]} - \widehat{G}|^2}{M(M-1)}. \tag{10}$$

The mean values of the sample variances (8) and (10) are given by

$$\mathrm{E}(\hat{\sigma}^2_{\widehat{G}^{[m]}}) = \frac{\sigma^2_{N_G}}{P}, \tag{11}$$

$$\mathrm{E}(\hat{\sigma}^2_{\widehat{G}}) = \frac{\sigma^2_{G_S} + \frac{\sigma^2_{N_G}}{P}}{M}. \tag{12}$$

For a linear system, $\hat{\sigma}^2_{\widehat{G}}$ originates only from the measurement noise source N_G and is therefore, using (11) and (12) with $\sigma^2_{G_S} = 0$, approximately equal to the mean value of $\hat{\sigma}^2_{\widehat{G}^{[m]}}$ divided by M

$$\hat{\sigma}^2_{\widehat{G}n} = \frac{1}{M^2}\sum_{m=1}^{M} \hat{\sigma}^2_{\widehat{G}^{[m]}}. \tag{13}$$

If $\hat{\sigma}^2_{\widehat{G}}$ is significantly larger than $\hat{\sigma}^2_{\widehat{G}n}$, this indicates a nonlinear behavior and

$$\hat{\sigma}^2_{G_S} = \max\left(M\left(\hat{\sigma}^2_{\widehat{G}} - \hat{\sigma}^2_{\widehat{G}n}\right), \quad 0\right), \tag{14}$$

is a variance estimate for the stochastic nonlinear contribution.

3.3 MIMO Measurements

For the MIMO case (n_y outputs, n_u inputs), the input vector is not invertible and it is therefore impossible to calculate an estimate like in (6) directly from data. To handle this, $M \cdot n_u$ experiments are carried out and grouped into M blocks of n_u experiments. For each block m and period p the sampled data are collected into matrices,

$\mathbf{U}^{[m,p]}(\omega_k) \in \mathbb{C}^{n_u \times n_u}$ and $\mathbf{Y}^{[m,p]}(\omega_k) \in \mathbb{C}^{n_y \times n_u}$ (bold-face in the sequel) where each column corresponds to one experiment. If $\mathbf{U}^{[m,p]}(\omega_k)$ has full rank, a nonparametric estimate of the multivariable FRF can be formed as

$$\widehat{G}^{[m,p]}(\omega_k) = \mathbf{Y}^{[m,p]}(\omega_k)\big(\mathbf{U}^{[m,p]}(\omega_k)\big)^{-1}. \tag{15}$$

As excitation signal, the so-called *orthogonal random phase multisine* (Dobrowiecki and Schoukens, 2005) will be used. This signal has been suggested in Dobrowiecki and Schoukens (2005) as input to maximize $|\det \mathbf{U}|$, given certain input amplitude constraints, which in turn will minimize the uncertainty in the FRF estimate (cf. (17)). Using this type of signal as reference signal instead corresponds to the case of output amplitude constraints. For each block of n_u experiments, the excitation signal is then generated as

$$\mathbf{R}(\omega_k) = \mathbf{R}_{\text{diag}}(\omega_k)\mathbf{T}, \tag{16}$$

where $\mathbf{R}_{\text{diag}}(\omega_k)$ is a diagonal matrix where each diagonal element contain a random phase multisine signal and $\mathbf{T}$ is an arbitrary, deterministic, orthogonal matrix. Here, the elements of $\mathbf{T}$ are chosen as

$$\mathbf{T}_{kn} = e^{\frac{2\pi j}{n_u}(k-1)(n-1)}.$$

See, e.g., Dobrowiecki and Schoukens (2005) for a background to this choice. For each of the M blocks, a new $\mathbf{R}_{\text{diag}}(\omega_k)$ is generated, in the sense that a new realization of the random phases is created.

To obtain variance expressions for the MIMO case, the same method as in Section 3.2 is used, but now taken element-wise.

4 Experimental Results

To illustrate the estimation method, both SISO and MIMO experiments have been carried out. For the SISO case, only axis 1 of the industrial robot is excited. The SISO measurements are easier to analyze and give insights that can be used when trying to analyze the MIMO measurements.

As excitation, an odd random phase multisine signal will be used as motor velocity reference $r(t)$ with period time $T_0 = N_P T_s = 10\,\text{s}$. A flat amplitude spectrum is used in the frequency interval 1–40 Hz, giving 195 excited frequencies. The commanded input u and measured output y are both sampled at $f_s = 1/T_s = 2\,\text{kHz}$. Since nonlinear effects are studied, the results will vary with the excitation signal, such as its shape and amplitude. This is, for example, the case when having a nonlinear spring stiffness in the transmission. Typically, the spring gets stiffer for larger amplitudes. Different signal amplitudes will be studied for the MIMO case. To reduce the effect of static friction, it is common to use an excitation signal which avoids zero velocity as much as possible. Therefore some experiments are carried out where $r(t) = r_{ms}(t) + r_{sq}(t)$, with $r_{ms}(t)$ the multisine signal and $r_{sq}(t)$ a square wave with frequency 0.1 Hz. To avoid too large input torques, the square wave is low-pass filtered with a cut-off frequency of 1 Hz.

In the figures, the FRF estimate from motor torque to motor acceleration is plotted. That avoids the integrator (-20 dB/decade) such that the resonances appear more clearly.

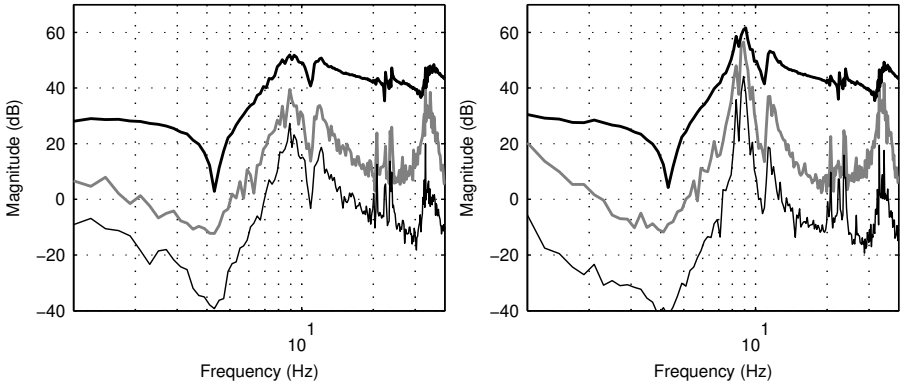


Figure 3: *Estimated FRF $\widehat{G}$ (black line) together with standard deviations $\hat{\sigma}_{\widehat{G}}$ (gray line) and $\hat{\sigma}_{\widehat{G}n}$ (thin black line) for data sets 1 (left) and 2 (right).*

4.1 SISO Measurements

Two data sets are collected from axis 1 of the industrial robot with amplitudes:

Data set 1: $\max_t |r_{ms}(t)| = 16\,\text{rad/s}$, $\max_t |r_{sq}(t)| = 0\,\text{rad/s}$.

Data set 2: $\max_t |r_{ms}(t)| = 16\,\text{rad/s}$, $\max_t |r_{sq}(t)| = 20\,\text{rad/s}$.

For each data set, $M = 6$ experiments are performed and $P = 2$ periods of the steady-state response are collected. The estimation method, according to (6)–(10) and (13), is then applied to the two different data sets. The estimated FRF, $\widehat{G}$, together with the total sample variance $\hat{\sigma}^2_{\widehat{G}}$ and the noise sample variance $\hat{\sigma}^2_{\widehat{G}n}$ (actually the standard deviations $\hat{\sigma}_{\widehat{G}}$ and $\hat{\sigma}_{\widehat{G}n}$) can be seen in Figure 3 for the two data sets. From the figure, it is evident that nonlinear distortions are indeed present and cause larger variability in the FRF than the measurement noise contributions ($\hat{\sigma}_{\widehat{G}} \gg \hat{\sigma}_{\widehat{G}n}$). Using the filtered square wave also gives slightly more pronounced resonances since the influence of static friction is reduced.

Since the nonlinear distortions give larger variability than the measurement noise, averaging over different experiments is important in order to get a more accurate FRF estimate. Averaging over several periods will not be as effective since that only reduce the measurement noise contributions. A somewhat puzzling result is that the variance in the estimate is increased around the resonance frequency at 9 Hz when using the filtered square wave. In Figure 4 the input and output spectra are plotted for the two different data sets. By comparing the input spectra for the two cases it can be seen that without the filtered square wave the input power around the resonance frequency is larger. This is reasonable since the static friction acts as an increased damping, which in turn requires a larger input signal in order to follow the reference signal. Assuming similar noise levels, the filtered square wave will therefore give a lower signal-to-noise ratio (SNR). That explains the increased variance, which, in a linear setup, usually is given by $\sigma^2_{\widehat{G}} = \sigma^2_V/|U|^2$, where $\sigma^2_V(\omega_k) = \mathrm{E}(|V(\omega_k)|^2)$ is the noise variance (assuming open loop, see (Ljung, 1999, p. 176)).

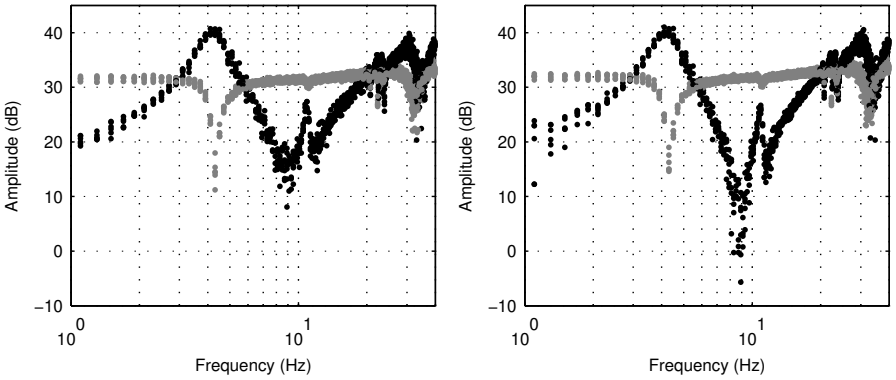


Figure 4: *Input spectrum $|U|^2$ (black) and output spectrum $|Y|^2$ (gray) for the six experiments of data sets 1 (left) and 2 (right).*

4.2 MIMO Measurements

For the MIMO case, three different data sets are collected with amplitudes:

Data set 3: $\max_t |r_{ms}(t)| = 5\,\text{rad/s}$, $\max_t |r_{sq}(t)| = 6\,\text{rad/s}$.

Data set 4: $\max_t |r_{ms}(t)| = 10\,\text{rad/s}$, $\max_t |r_{sq}(t)| = 12\,\text{rad/s}$.

Data set 5: $\max_t |r_{ms}(t)| = 10\,\text{rad/s}$, $\max_t |r_{sq}(t)| = 0\,\text{rad/s}$.

For each data set, $M = 9$ blocks of experiments are performed and $P = 2$ periods of the steady-state response are collected. The estimation method, according to (15), (7)–(10), and (13), is then applied (element-wise) to the three different data sets. The estimated FRF, $\widehat{G}$, together with the total sample variance $\hat{\sigma}^2_{\widehat{G}}$ and the noise sample variance $\hat{\sigma}^2_{\widehat{G}n}$ (actually the standard deviations $\hat{\sigma}_{\widehat{G}}$ and $\hat{\sigma}_{\widehat{G}n}$) can be seen in Figures 5–7 for the three data sets.

The influence of the excitation amplitude can be studied in Figures 5 and 6. As can be seen, a lower amplitude gives more fluctuations in the FRF estimate. For the (1,1) element one can also see a slightly increased anti-resonance frequency at $4\,\text{Hz}$, indicating a nonlinear stiffness (to see this effect more clearly, the difference in amplitudes should actually be larger). Comparing Figures 6 and 7, one clearly sees the influence of static friction and the importance of using the square wave. In Figure 7, almost all resonances have disappeared for axes 2 and 3. Still, similar to the SISO case, the variance of the estimate is increased around the resonance frequencies when using the square wave.

The FRF estimate in Figure 7 is, according to its sample variance, fairly accurate (at least for the diagonal elements), but still almost all resonances have disappeared. How can this be the case? To answer this question, one should study the properties of $G_R(\omega_k)$ in (5), the best linear approximation to the nonlinear system. A linear approximation of a nonlinear system will depend on both the amplitude distribution and the spectrum of the excitation signal. Adding a square wave changes the shape of the excitation signal and will therefore change the best linear approximation. The user must therefore decide

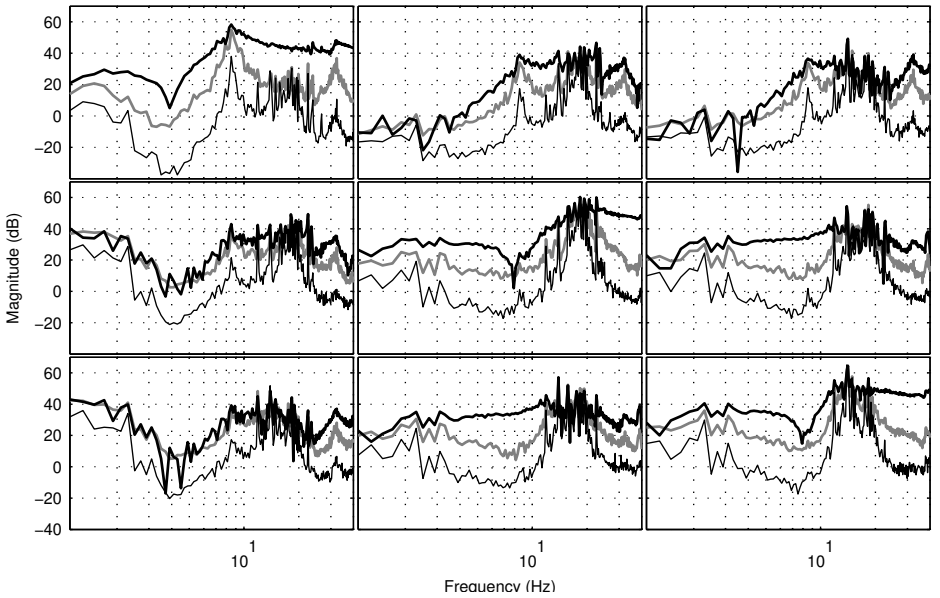


Figure 5: *Estimated FRF $\widehat{G}$ (black line) for data set 3 together with standard deviations $\hat{\sigma}_{\widehat{G}}$ (gray line) and $\hat{\sigma}_{\widehat{G}n}$ (thin black line).*

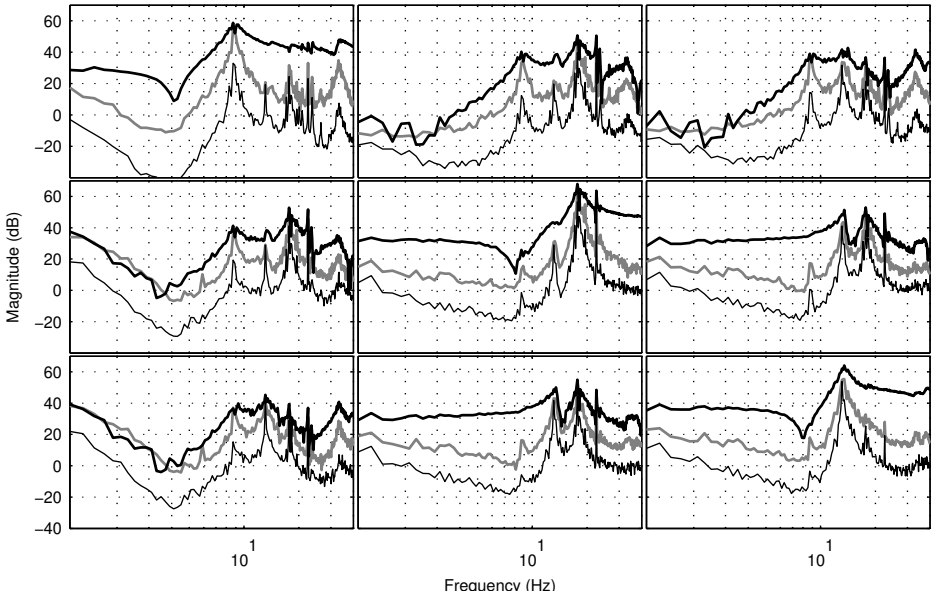


Figure 6: *Estimated FRF $\widehat{G}$ (black line) for data set 4 together with standard deviations $\hat{\sigma}_{\widehat{G}}$ (gray line) and $\hat{\sigma}_{\widehat{G}n}$ (thin black line).*

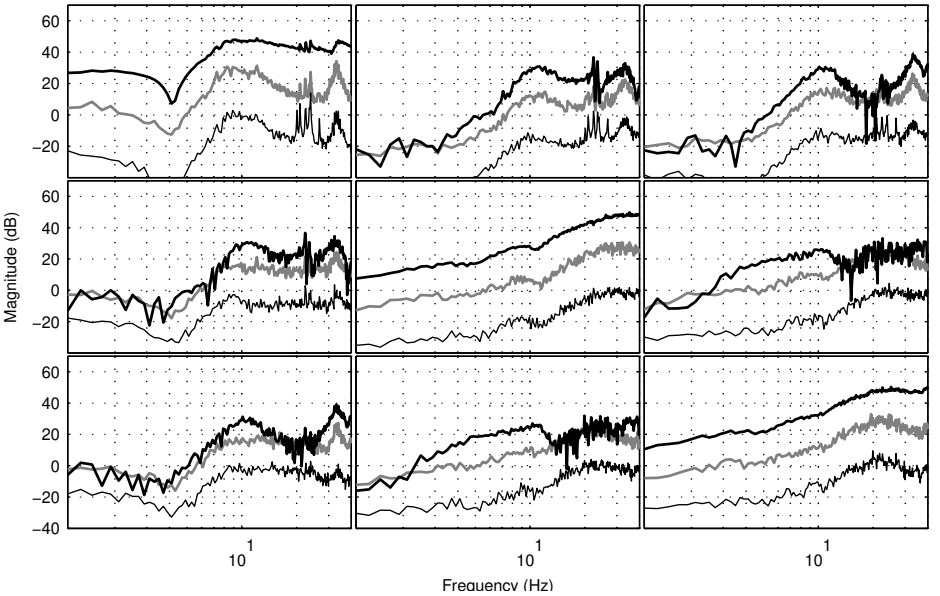


Figure 7: *Estimated FRF $\widehat{G}$ (black line) for data set 5 together with standard deviations $\hat{\sigma}_{\widehat{G}}$ (gray line) and $\hat{\sigma}_{\widehat{G}n}$ (thin black line).*

which "best" linear approximation that is most suitable for the modeling purpose and select a corresponding excitation signal for the experiments.

As was noted in the SISO case, the SNR was a possible explanation for the increased variability in the estimates when using a square wave. For the MIMO case, the average uncertainty (standard deviation) in the FRF estimate is inversely proportional to $|\det \mathbf{U}|^{1/n_u}$. To see this, we first need the covariance matrix[4]

$$\sigma^2_{\widehat{G}} = \mathrm{Cov}(\mathrm{vec}(\widehat{G})) = [\mathbf{U}\mathbf{U}^H]^{-T} \otimes \sigma^2_V, \tag{17}$$

where $\otimes$ is the Kronecker product and $\sigma^2_V = \mathrm{E}(VV^H)$ is the covariance matrix of the measurement noise. As a measure of the size of this covariance matrix, the geometric mean of its volume is used, i.e., $(\det \sigma^2_{\widehat{G}})^{\frac{1}{n_u n_y}} = (\det \sigma^2_V)^{1/n_y} / |\det \mathbf{U}|^{2/n_u}$. Taking the square root then gives the expression.

In Figure 8, the average of $|\det \mathbf{U}^{[m,p]}|^{1/n_u}$ (over the $M = 9$ different blocks) of the input matrix can be seen for the three different data sets. As can be seen, the average input power is reduced by adding the square wave (compare thin black line and gray line), except for frequencies above 25 Hz. The biggest differences are for frequencies 1–3 Hz and 9–20 Hz. This can also be noted by comparing Figure 8 with the total sample variance $\hat{\sigma}^2_{\widehat{G}}$ in Figures 6 and 7 (e.g., element (2,2)). To conclude: the square wave is useful to obtain nonparametric FRF estimates with more pronounced resonances by avoiding static friction, but the price to pay for this is increased variance.

[4] Assuming a linear system and neglecting correlation between the input signal and the measurement noise, see Wernholt and Gunnarsson (2007), Theorem 1.

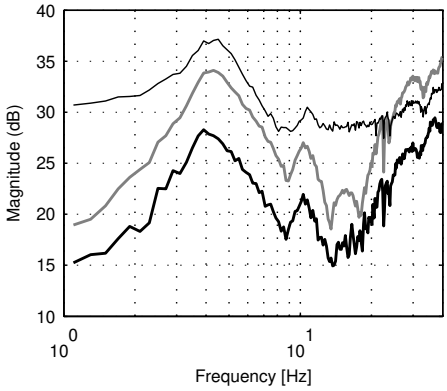


Figure 8: *Average of* $|\det \mathbf{U}^{[m,p]}|^{1/n_u}$ *(over the* $M = 9$ *different blocks) of the input matrix for data sets 3 (black line), 4 (gray line), and 5 (thin black line).*

Nonparametric MIMO FRF estimates also tend to be non-symmetric, even if the underlying linear system is symmetric ($G_{ij} = G_{ji}$) (Wernholt and Gunnarsson, 2004, 2007). This can be noted here as well in Figures 5–7 for frequencies below 4 Hz, especially for the elements (1,2), (1,3), (2,1) and (3,1). To understand this, consider once more the covariance matrix in (17). The variance of element $\widehat{G}_{ij}$ is given by the noise-to-signal ratio

$$\sigma^2_{\widehat{G}_{ij}} = \left[\left[\mathbf{U}\mathbf{U}^H\right]^{-1}\right]_{jj} \left[\sigma^2_V\right]_{ii}, \tag{18}$$

where $[\sigma^2_V]_{ii}$ is the noise variance in output i. One thing to note here is that all elements in a column will have the same influence from the input signal. If the noise variance would be approximately equal for the different outputs, then the uncertainty of all elements in a column would be approximately equal. By using (18), a non-symmetric FRF estimate comes as no surprise. There are, of course, no guarantees that the variance should be equal for elements $\widehat{G}_{ij}$ and $\widehat{G}_{ji}$.

5 Conclusion

A method for the detection and estimation of nonlinear distortions in nonparametric FRF estimates has successfully been applied to experimental data from an industrial robot. The results show that nonlinear distortions are indeed present and cause larger variability in the FRF than the measurement noise contributions. To get more accurate nonparametric FRF estimates one should therefore use available measurement time to average estimates from several experiments with random phase multisines. Averaging over several periods will not be as effective since that only reduces the measurement noise contributions. Using a filtered square wave to reduce the effect of friction will give more pronounced resonances, but will at the same time give larger variability in the FRF estimates due to a poorer SNR. This could be handled by modifying the amplitude spectrum of the excitation signal.

Acknowledgments

This work has been supported by VINNOVA's center of excellence ISIS at Linköping University and the Swedish Research Council (VR). The experiments have been carried out in the research lab of ABB AB – Robotics, Västerås, Sweden, which is gratefully acknowledged.

References

Albu-Schäffer, A. and Hirzinger, G. (2001). Parameter identification and passivity based joint control for a 7DOF torque controlled light weight robot. In *Proc. 2001 IEEE International Conference on Robotics and Automation*, pages 2852–2858, Seoul, Korea.

Behi, F. and Tesar, D. (1991). Parametric identification for industrial manipulators using experimental modal analysis. *IEEE Transactions on Robotics and Automation*, 7(5):642–652.

Dhaouadi, R., Ghorbel, F. H., and Gandhi, P. S. (2003). A new dynamic model of hysteresis in harmonic drives. *IEEE Transactions on Industrial Electronics*, 50(6):1165–1171.

Dobrowiecki, T. and Schoukens, J. (2005). Measuring linear approximation to weakly nonlinear MIMO systems. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Gutt, H.-J., Scholl, F. D., and Blattner, J. (1996). High precision servo drives with DSP-based torque ripple reduction. In *Proc. IEEE AFRICON, 1996*, volume 2, pages 632–637.

Hanselman, D. (1990). Resolver signal requirements for high accuracy resolver-to-digital conversion. *IEEE Transactions on Industrial Electronics*, 37(6):556–561.

Johansson, R., Robertsson, A., Nilsson, K., and Verhaegen, M. (2000). State-space system identification of robot manipulator dynamics. *Mechatronics*, 10(3):403–418.

Kozlowski, K. (1998). *Modelling and identification in robotics.* Advances in Industrial Control. Springer, London.

Ljung, L. (1999). *System Identification: Theory for the User.* Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.

Pintelon, R., Rolain, Y., Vandersteen, G., and Schoukens, J. (2004a). Experimental characterization of operational amplifiers: A system identification approach – part II: Calibrations and measurements. *IEEE Transactions on Instrumentation and Measurement*, 53(3):863–876.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach*. IEEE Press, New York.

Pintelon, R. and Schoukens, J. (2002). Measurement and modeling of linear systems in the presence of non-linear distortions. *Mechanical Systems and Signal Processing*, 16:785–801.

Pintelon, R. and Schoukens, J. (2001). Measurement of frequency response functions using periodic excitations, corrupted by correlated input/output errors. *IEEE Transactions on Instrumentation and Measurement*, 50(6):1753–1760.

Pintelon, R., Vandersteen, G., De Locht, L., Rolain, Y., and Schoukens, J. (2004b). Experimental characterization of operational amplifiers: A system identification approach – part I: Theory and simulations. *IEEE Transactions on Instrumentation and Measurement*, 53(3):854–862.

Schoukens, J., Pintelon, R., Dobrowiecki, T., and Rolain, Y. (2005). Identification of linear systems with nonlinear distortions. *Automatica*, 41(3):491–504.

Verboven, P., Guillaume, P., Vanlanduit, S., and Cauberghe, B. (2006). Assessment of nonlinear distortions in model testing and analysis of vibrating automotive structures. *Journal of Sound and Vibration*, 293(1-2):299–319.

Wernholt, E. and Gunnarsson, S. (2004). On the use of a multivariable frequency response estimation method for closed loop identification. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 827–832, Nassau, Bahamas.

Wernholt, E. and Gunnarsson, S. (2006a). Detection and estimation of nonlinear distortions in industrial robots. In *Proc. 23rd IEEE Instumentation and Measurement Technology Conference*, pages 1913–1918, Sorrento, Italy.

Wernholt, E. and Gunnarsson, S. (2006b). Nonlinear identification of a physically parameterized robot model. In *Proc. 14th IFAC Symposium on System Identification*, pages 143–148, Newcastle, Australia.

Wernholt, E. and Gunnarsson, S. (2007). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Paper E

Experiment Design for Identification of Nonlinear Gray-Box Models with Application to Industrial Robots

Edited version of the paper:

> Wernholt, E. and Löfberg, J. (2007). Experiment design for identification of nonlinear gray-box models with application to industrial robots. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Preliminary version published as Technical Report LiTH-ISY-R-2774, Dept. of Electrical Engineering, Linköping University, SE–581 83 Linköping, Sweden.

Experiment Design for Identification of Nonlinear Gray-Box Models with Application to Industrial Robots

Erik Wernholt and Johan Löfberg

Dept. of Electrical Engineering,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: {erikw,johanl}@isy.liu.se.

Abstract

Experiment design involving selection of optimal experiment positions for nonlinear gray-box models is studied. From the derived Fisher information matrix, a convex optimization problem is posed. By considering the dual problem, the experiment design is efficiently solved with linear complexity in the number of candidate positions, compared to cubic complexity for the primal problem. In the numerical illustration, using an industrial robot, the parameter covariance is reduced by a factor of six by using the 15 optimal positions compared to using the optimal single position in all experiments.

1 Introduction

Experiment design has been treated in the literature for almost half a century in both statistics and engineering areas, where some basic references are Fedorov (1972) and Goodwin and Payne (1977). A recent survey can also be found in Gevers (2005). The experiment design problem aims at finding the optimal experiment conditions, χ, contained in some set $\boldsymbol{\chi}$, that minimize a criterion related to the expected outcome of the experiment under certain constraints.

For parameter estimation, the parameter accuracy is a function of both the experimental conditions and the estimator. It is common, see Goodwin and Payne (1977), to assume that the estimator is efficient in the sense that the parameter covariance matrix achieves the Cramér-Rao lower bound (inverse of the Fisher information matrix), which is a lower limit on the covariance matrix that can be obtained with an unbiased estimator. The Fisher information matrix can be obtained without specifying an estimator. Usually, an assumption of normally distributed noise is made. According to Goodwin and Payne (1977, p. 127), this is not very restrictive since, asymptotically, the covariance matrix in

many cases will be the same even for other noise distributions. This is, for example, the case for the prediction error method (Ljung, 1999).

The classical approach has then been to minimize some scalar measure of the (asymptotic) covariance matrix $P_\theta(\chi, \theta_0)$ with constraints on the input and/or the output power. Some commonly used measures are A-optimality $\mathrm{Tr}\, P_\theta(\chi, \theta_0)$, D-optimality $\det P_\theta(\chi, \theta_0)$, E-optimality $\lambda_{max}\left(P_\theta(\chi, \theta_0)\right)$, and L-optimality $\mathrm{Tr}\left(W P_\theta(\chi, \theta_0)\right)$. An inherent problem in most experiment design problems is that the covariance matrix depends on the true system parameters θ_0. These parameters are, at least partly, unknown or uncertain, otherwise there would be no need for an experiment. This has been handled in different ways in the literature. One could either assume a good prior estimate of θ_0 and hope for the best, or use some robust experiment design methods. Various strategies have been suggested, including sequential design (iterate between parameter estimation and experiment design), Bayesian design (minimize the expected value over the prior parameter distribution) and min-max design (minimize the worst case when the parameters are contained in a given set). For references, see, e.g., Goodwin et al. (2006).

In addition to the problem of knowing θ_0, experiment design problems are, in their original form, often intractable due to non-convex and infinite-dimensional constraints, as well as the problem of finding a signal realization which has the desired spectral properties. However, due to great advances in the optimization community, there exist today many useful methods to reformulate the experiment design problems into tractable convex optimization problems. Here, Jansson and Hjalmarsson (2005) is a good example, where a framework for this reformulation is presented, based on linear parameterizations of the signal spectrum. Constraints on these spectra can be included, as well as quality constraints and robustness constraints as long as they can be rewritten as convex functions of the inverse (asymptotic) covariance matrix.

2 Problem Description

The work in this paper is inspired by the problem of identification of flexibility parameters in a nonlinear industrial robot model in Öhr et al. (2006). The industrial robot is challenging for system identification since it is a multivariable, nonlinear, unstable, and highly resonant system. Consider therefore the nonlinear gray-box model

$$\dot{x}(t) = f(x(t), u(t), \theta), \tag{1a}$$

$$y(t) = h(x(t), u(t), \theta), \tag{1b}$$

with state vector $x(t) \in \mathbb{R}^{n_x}$, input $u(t) \in \mathbb{R}^{n_u}$, output $y(t) \in \mathbb{R}^{n_y}$, and the nonlinear functions $f(\cdot)$ and $h(\cdot)$, parameterized by $\theta \in \mathbb{R}^d$. A discrete-time feedback controller[1] $u(t) = F(q)(r(t) - y(t))$ is used to stabilize the system, where $F(q)$ is the controller and q is the difference operator. The experiment design problem involves finding the excitation signal $r(t)$ (and possibly $F(q)$ if that is a design variable) that minimizes the parameter uncertainty, given certain constraints on $r(t), u(t), x(t)$. This is a nonlinear optimal control problem which is extremely hard to solve. One option would be to apply

[1] Physical modeling makes continuous-time gray-box models most natural. Still, a discrete-time controller will be used which gives that the excitation signal $r(t)$ in the experiment design is a discrete-time signal. For simplicity, blocks for sample and hold are left out in the sequel.

a prediction error method for the parameter estimation. A covariance matrix could then be calculated from the cost function. However, minimizing some measure of this covariance matrix would probably be a hard non-convex optimization problem. In addition comes the problem of obtaining a stable predictor for the nonlinear system.

In this paper, a different approach will be used, where we assume that $r(t)$ is a small perturbation around a certain operating point (x_0, u_0), called position in this paper. The perturbation is assumed to be small enough to justify linearization of the nonlinear system (1) and the use of linear theory. The identification will be carried out in the frequency domain by comparing the experimental FRF (Frequency Response Function) with the parametric FRF obtained by linearizing the nonlinear model.

The information about the unknown parameters will differ between different positions. Therefore, given a limited total measurement time, one should perform experiments in the position(s) that contribute the most to the information about the unknown parameters. Here, a set of Q_c candidate positions will be assumed and a total of M experiments should be performed. For each position, a separate experiment design problem can be solved, as in Jansson and Hjalmarsson (2005), where the excitation signal is optimized in order to meet constraints and minimize the parameter uncertainty, given a fixed measurement time.

This will be a suboptimal solution to the overall experiment design problem since the optimization is done in two steps. However, in principle it is possible to solve the overall problem by, for each position, parameterize the excitation signal (N_f parameters) and add constraints, and then solve the resulting problem with $Q_c \cdot N_f$ variables. The optimization problem is still convex if the constraints are convex functions of the variables that define the excitation signals (for example, a power constraint). This is an interesting extension of the work presented here and preliminary studies show promising results.

The paper is organized as follows: In Section 3, the Fisher information matrix is derived, assuming certain properties of the estimated FRFs. This enables the formulation of the experiment design problem in Section 4. The solution to this problem is discussed in Section 5. A numerical illustration, using the industrial robot, is given in Section 6. Finally, some conclusions are drawn in Section 7.

3 The Information Matrix

In this section, we assume that information about the unknown parameters can be obtained by collecting measurements from the nonlinear system in Q_c different positions. The complex-valued FRF measurements are assumed to have the following properties.

Assumption E.1. Let $\widehat{\mathcal{G}}^{(i)}(l) \in \mathbb{C}^n$ denote measurement l in position i. For each position $i = 1, \ldots, Q_c$, m_i experiments will be carried out, giving a total of $M = \sum_{i=1}^{Q_c} m_i$ experiments. The measurements $\widehat{\mathcal{G}}^{(i)}(l)$ can be written as

$$\widehat{\mathcal{G}}^{(i)}(l) = \mathcal{G}^{(i)}(\theta_0) + \eta^{(i)}(l), \tag{2}$$

where $\mathcal{G}^{(i)}(\theta)$ is the parametric FRF for position i and θ_0 is the same for all positions. The measurement noise $\eta^{(i)}(l)$ is a zero mean, circular complex normally distributed random variable, independent (over i and l) and identically distributed (over l) with covariance matrix $\Lambda_0^{(i)}$.

For FRF measurements with N_f frequencies, n would be $n = n_u \cdot n_y \cdot N_f$, see (13). Given the measurements and a model, one must come up with an estimator $\hat{\theta}_M$ for the unknown parameters. If the estimator is unbiased ($\mathrm{E}(\hat{\theta}_M) = \theta_0$), the covariance of any such estimator is bounded by the *Fisher information matrix* H,

$$\mathrm{E}((\hat{\theta}_M - \theta_0)(\hat{\theta}_M - \theta_0)^T) \succeq H^{-1}. \tag{3}$$

This bound is often referred to as the *Cramér-Rao lower bound*. Next, the Fisher information matrix, H, as well as the maximum likelihood estimator are derived, given Assumption E.1.

Theorem E.1
Given Assumption E.1, the Fisher information matrix is given by

$$H = \sum_{i=1}^{Q_c} m_i H_i, \tag{4}$$

where H_i is the information matrix in position i,

$$H_i = 2\Re\left\{\overline{\Psi^{(i)}(\theta_0)}[\Lambda_0^{(i)}]^{-1}[\Psi^{(i)}(\theta_0)]^T\right\}, \tag{5}$$

with the Jacobian matrix $[\Psi^{(i)}(\theta)]^T = \frac{\partial \mathcal{G}^{(i)}(\theta)}{\partial \theta} \in \mathbb{C}^{n\times d}$ and $\overline{(\cdot)}$ denoting complex conjugate.

Proof: The information matrix H is calculated as

$$H = \mathrm{E}\Big(\left[\frac{\partial \log f(\widehat{\mathcal{G}}|\theta)}{\partial \theta}\right] \left[\frac{\partial \log f(\widehat{\mathcal{G}}|\theta)}{\partial \theta}\right]^T \Big|_{\theta=\theta_0} \Big),$$

where $f(\widehat{\mathcal{G}}|\theta)$ is the conditional probability density function (PDF) for $\widehat{\mathcal{G}}$, given θ, where $\widehat{\mathcal{G}}$ denotes the set of measurements $\{\widehat{\mathcal{G}}^{(i)}(l),\ l = 1,\dots,m_i,\ i = 1,\dots,Q_c\}$. Independent measurements give that

$$f(\widehat{\mathcal{G}}|\theta) = \prod_{i=1}^{Q_c}\prod_{l=1}^{m_i} f(\widehat{\mathcal{G}}^{(i)}(l)|\theta) = \prod_{i=1}^{Q_c}\prod_{l=1}^{m_i} f_{\eta^{(i)}}(\mathcal{E}^{(i)}(l,\theta)),$$

where $\mathcal{E}^{(i)}(l,\theta) = \widehat{\mathcal{G}}^{(i)}(l) - \mathcal{G}^{(i)}(\theta)$ and $f_{\eta^{(i)}}(x)$ is the PDF for the noise $\eta^{(i)}(l) \in \mathbb{C}^n$. Zero mean circular complex normally distributed noise has the PDF (Pintelon and Schoukens, 2001, p. 437)

$$f_{\eta^{(i)}}(x) = \frac{1}{\pi^n \det(\Lambda_0^{(i)})} \exp(-x^H(\Lambda_0^{(i)})^{-1}x),$$

which gives

$$\log f(\widehat{\mathcal{G}}|\theta) = -\sum_{i=1}^{Q_c}\sum_{l=1}^{m_i}\Big(n\log\pi + \log\det\Lambda_o^{(i)} + [\mathcal{E}^{(i)}(l,\theta)]^H[\Lambda_o^{(i)}]^{-1}\mathcal{E}^{(i)}(l,\theta)\Big), \tag{6}$$

and

$$\left.\frac{\partial \log f(\widehat{\mathcal{G}}|\theta)}{\partial \theta}\right|_{\theta=\theta_0} = \sum_{i=1}^{Q_c}\sum_{l=1}^{m_i} 2\Re\left\{\overline{\Psi^{(i)}(\theta_0)}[\Lambda_0^{(i)}]^{-1}\eta^{(i)}(l)\right\}.$$

Since $\eta^{(i)}(l)$ is independent over i and l, the information matrix H is finally given by

$$\begin{aligned} H &= \sum_{i=1}^{Q_c}\sum_{l=1}^{m_i} 4\mathrm{Cov}(\Re\left\{\overline{\Psi^{(i)}(\theta_0)}[\Lambda_0^{(i)}]^{-1}\eta^{(i)}(l)\right\}) \\ &= \sum_{i=1}^{Q_c}\sum_{l=1}^{m_i} 2\Re\left\{\mathrm{Cov}(\overline{\Psi^{(i)}(\theta_0)}[\Lambda_0^{(i)}]^{-1}\eta^{(i)}(l))\right\} \\ &= \sum_{i=1}^{Q_c} m_i \cdot 2\Re\left\{\overline{\Psi^{(i)}(\theta_0)}[\Lambda_0^{(i)}]^{-1}[\Psi^{(i)}(\theta_0)]^T\right\}, \end{aligned}$$

where the fact that $\eta^{(i)}(l)$ is a circular complex random vector has been used. [If x is a zero mean circular complex random vector ($\mathrm{Cov}(x) = \mathrm{E}(xx^H)$, $\mathrm{E}(xx^T) = 0$), then $y = Ax$, with A a complex matrix, is circular complex as well. For a circular complex random vector, $\mathrm{Cov}(\Re\{x\}) = 0.5\Re\{\mathrm{Cov}(x)\}$ is also easy to verify.] □

Corollary E.1
Given Assumption E.1, the maximum likelihood estimator $\hat{\theta}_M^{ML} = \arg\max_\theta \log f(\widehat{\mathcal{G}}|\theta)$ *is given by*

$$\hat{\theta}_M^{ML} = \arg\min_\theta \sum_{i=1}^{Q_c}\sum_{l=1}^{m_i} [\mathcal{E}^{(i)}(l,\theta)]^H[\Lambda_o^{(i)}]^{-1}\mathcal{E}^{(i)}(l,\theta), \tag{7}$$

with $\mathcal{E}^{(i)}(l,\theta) = \widehat{\mathcal{G}}^{(i)}(l) - \mathcal{G}^{(i)}(\theta)$. *The estimator is also efficient, in the sense that the Cramér-Rao lower bound in Theorem E.1 is attained asymptotically* ($M \to \infty$).

Proof: The likelihood function is given by (6) which immediately gives (7). Efficiency follows from classical results for the maximum likelihood estimator in case of independent measurements (see, e.g., Ljung, 1999, p. 215). □

4 The Experiment Design Problem

Consider the information matrix H in Theorem E.1. The experiment design problem can be stated as follows: *Given Q_c different candidate positions and a total number of M experiments to be carried out, choose the numbers m_i, $i = 1, \ldots, Q_c$ to minimize some measure of the Cramér-Rao lower bound H^{-1}.*

Some commonly used measures were mentioned in Section 1. The solution will in general depend on what measure that is used. For simplicity, we will here consider the D-optimal design problem. One could also argue that some quality constraints should be added that reflect the intended model application, e.g., the weighted relative error (see Jansson and Hjalmarsson, 2005). That is, however, beyond the scope of this paper, whose

main purpose is to illustrate that the optimal solution involves experiments in more than one position. The D-optimal design problem can be posed as

$$\begin{aligned} &\text{minimize} \quad \log\det\left[\frac{1}{M}\sum_{i=1}^{Q_c} m_i H_i\right]^{-1}, \\ &\text{subject to} \quad m \succeq 0, \quad \mathbf{1}^T m = M, \quad m \in \mathbb{Z}^{Q_c}, \end{aligned} \tag{8}$$

where $H_i = H_i^T \in \mathbb{R}^{d\times d}$ is the information matrix from position i, see (5) and the factor $1/M$ is added to obtain the average information matrix. This is a *combinatorial experiment design problem* since the m_i are constrained to be integers. Given a large number of candidates Q_c, the problem will relatively quickly become intractable[2]. This problem fits into the framework of Boyd and Vandenberghe (2004, Chap. 7.5) where also E-optimal and A-optimal designs are treated. See also Vandenberghe et al. (1998, Section 2.4). When M is large, a reasonable approximate solution can be obtained by considering the *relaxed experiment design problem*

$$\begin{aligned} &\text{minimize} \quad \log\det\left[\sum_{i=1}^{Q_c} \lambda_i H_i\right]^{-1}, \\ &\text{subject to} \quad \lambda \succeq 0, \quad \mathbf{1}^T \lambda = 1, \end{aligned} \tag{9}$$

For the relaxed problem, M will no longer affect the solution λ^* and the Cramér-Rao lower bound H^{-1} is thus inversely proportional to the number of experiments, M, that will be carried out. For the combinatorial problem (8), this is only asymptotically true (cf. Figure 4).

From the solution λ^* of the relaxed problem (9) it is possible to obtain a suboptimal solution $\tilde{m}$ to the combinatorial problem (8) by using Algorithm E.1.

Algorithm E.1

$\tilde{m} := \text{round}(M\lambda^*)$
while $\sum_{i=1}^{M} \tilde{m}_i \neq M$ ***do***
  if $\sum_{i=1}^{M} \tilde{m}_i > M$ ***then***
    $k := \arg\max_i(\tilde{m}_i - M\lambda_i^*)$
    $\tilde{m}_k := \tilde{m}_k - 1$
  else
    $k := \arg\min_i(\tilde{m}_i - M\lambda_i^*)$
    $\tilde{m}_k := \tilde{m}_k + 1$
  end if
end while

Even though this is a suboptimal solution, it is possible to evaluate the solution by comparing the cost $V(M\lambda^*)$ and $V(\tilde{m})$, where

$$V(x) = \log\det\left[\sum_i x_i H_i\right]^{-1}.$$

[2]The mixed integer conic solver in YALMIP (Löfberg, 2004) is capable of solving (8) for the design problem presented in Section 6 for a few hundred positions.

$V(M\lambda^*)$ is a lower bound on $V(m^*)$, so we have $V(M\lambda^*) \leq V(m^*) \leq V(\tilde{m})$. If the difference $V(\tilde{m}) - V(M\lambda^*)$ is small, then $V(\tilde{m})$ must be close to $V(m^*)$. Note that when M is small, the difference $V(\tilde{m}) - V(M\lambda^*)$ will be large, but that is usually the case also for $V(m^*) - V(M\lambda^*)$. (This is, for example, evident for $M = 1$ in the numerical illustration, see Figure 4.)

5 Solving the Experiment Design Problem

Determinant maximization problems arise in many fields (see Vandenberghe et al. (1998) for an overview) and efficient solvers for this class of optimization problems are available. In fact, any semidefinite programming solver can be used to solve determinant maximization problems, although it requires cumbersome reformulations (Ben-Tal and Nemirovski, 2001). A recent example of a specialized solver is SDPT3 (Toh et al., 2006), which is designed to solve conic programming problems (which includes linear, quadratic, and semidefinite programs) with additional logarithmic barrier terms in the objective. This includes the determinant maximization problem (9) as a special case.

The purpose of this section is to motivate an alternative representation of (9), derived from duality. A common mis-conception is that when a so-called primal-dual solver such as SDPT3 is used, it does not make any difference how the model is implemented, in terms of primal and dual models. However, this is a dangerous fallacy.

Although the details are beyond the scope of this paper, we note that the complexity of solving the determinant optimization problem (9), with $d \ll Q_c$, can be shown to be at least cubic in Q_c (a linear system of size $Q_c \times Q_c$ has to be solved repeatedly during the solution process), when the model (9) is interpreted as the dual problem in the primal-dual pair in a standard conic programming model.[3]. Accordingly, the complexity of solving (9) scales badly with an increasing number of positions, Q_c.

Hence, blindly stating and solving the problem (9) can easily lead to unnecessarily complex models. Instead, we derive the dual of (9) as (cf. Vandenberghe et al., 1998, Eq. (3.4))

$$\begin{array}{ll} \text{minimize} & \log\det W^{-1}, \\ \text{subject to} & W = W^T \succ 0, \\ & \mathrm{Tr}(WH_i) \leq d, \quad i = 1, \ldots, Q_c. \end{array} \tag{10}$$

This problem, once again interpreted as the dual problem of the primal-dual pair, has only $d(d+1)/2$ variables (the variables parameterizing the symmetric matrix W). An increasing number of positions now instead leads to a growing number of linear inequality constraints. Although the complexity of solving a semidefinite program depends on the number of linear inequalities in the dual, it scales much better than increasing the number of variables. Roughly speaking, while complexity with respect to the number of variables in the dual for a standard SDP solver is at least cubic, the computational complexity with respect to the number of linear inequality constraints is essentially linear.

[3]Following the primal-dual definitions in Toh et al. (2006), which also is the way the modeling language YALMIP (Löfberg, 2004) and most people in the control community interprets problems.

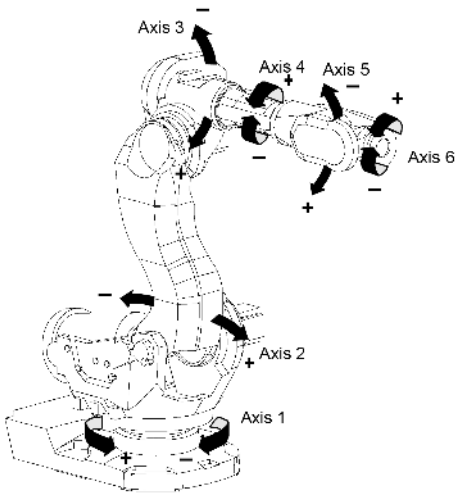


Figure 1: *The ABB manipulator IRB6600.*

Note that the original variables λ are dual variables related to the inequality constraints in the model (10). Hence, after solving the problem using the primal-dual solver SDPT3, the original variables are easily recovered.

6 Numerical Illustration

As a numerical illustration, we will use an industrial robot, see Figure 1. First, the robot model will be described in Section 6.1. The measurements $\widehat{\mathcal{G}}^{(i)}(l)$ will, for each experiment, be the nonparametric FRF estimate for a number of frequencies, which is explained in Section 6.2. In Section 6.3, the experiment design problem will be solved using YALMIP (Löfberg, 2004) and SDPT3 (Toh et al., 2006) and the results are presented in Section 6.4.

6.1 Robot Model

The model comes from Öhr et al. (2006) and consists of a kinematic chain of rigid bodies representing the links in the manipulator arm. The links are connected in joints by gearboxes and motors, and due to flexibilities in the gearboxes, spring-damper pairs are introduced between all motors and links. Two additional spring-damper pairs are introduced in some joints to model bearing flexibilities. The model dynamics can be described by the following set of equations.

$$\begin{bmatrix}\dot{v}_b\\ \dot{v}_a\end{bmatrix} = M_a^{-1}(q_b, q_a)\left[c(q_b, q_a, v_b, v_a) + g(q_b, q_a) - \begin{bmatrix}\tau_b\\ \tau_g\end{bmatrix}\right],$$
$$\dot{v}_m = M_m^{-1}\left(\tau_g + u\right), \quad \dot{q}_a = v_a, \quad \dot{q}_m = v_m, \quad \dot{q}_b = v_b,$$
$$\tau_g = K_g(q_a - q_m) + C_g(v_a - v_m), \quad \tau_b = K_b q_b + C_b v_b,$$

where q_a, q_m, q_b are angular positions for arms, motors, and bearings, respectively. The matrix $M_a(q_b, q_a)$ is the inertia matrix for the arms and M_m is the diagonal inertia matrix for the motors. The Coriolis and centrifugal torques are described by the function $c(q_b, q_a, v_b, v_a)$, and $g(q_b, q_a)$ represents gravity torque. The time t is omitted in the expressions. The matrices K_g, C_g and K_b, C_b describe the flexibilities and are here diagonal.

The model used in the numerical illustration has three-dimensional spring-damper pairs in the three main axes joints (1–3) and one-dimensional spring-damper pairs (gearbox) in the remaining joints (4–6), see Figure 1. This gives 12 spring-damper pairs and 18 degrees of freedom (6 arms + 6 motors + $2 \cdot 3$ joints having bearing flexibilities) and the nonlinear gray-box model (1) will have 36 states. As output, the motor angular velocity will be used, $y = v_m$. The rigid body parameters are known from a CAD (Computer Aided Design) model and the nominal spring-damper parameters are obtained from the identification described in Öhr et al. (2006). For the experiment design, θ will contain the 12 spring parameters ($d = 12$). In addition, these are scaled by the nominal values such that $\theta_0 = \mathbf{1}$.

The model equations are linearized in Q_c different positions $q_m = q_m^{(i)},\ i = 1, \ldots, Q_c$, with $v_m = v_a = v_b = 0$, $u = \tau_g$ and q_b, q_a the solution to $g(q_b, q_a) = \begin{bmatrix} \tau_b^T & \tau_g^T \end{bmatrix}^T$. (For simplicity, $q_a = q_m$, $q_b = 0$ can be used with minor errors if the springs are reasonably stiff.) From a resulting continuous-time transfer function $G_c^{(i)}(s, \theta)$, the discrete-time FRF $G^{(i)}(e^{j\omega T_s}, \theta)$ is given by (Åström and Wittenmark, 1984, Chap. 4.5)

$$G^{(i)}(e^{j\omega T_s}, \theta) = \frac{1 - e^{-j\omega T_s}}{j\omega T_s} G_c^{(i)}(j\omega, \theta), \tag{11}$$

where higher-order terms in the Poisson summation formula have been neglected. These terms are small since T_s is small and frequencies above π/T_s rad/s are well attenuated by the dynamics in the torque controllers (left out in the modeling).

6.2 Measurements

Assume a closed-loop setup

$$y(t) = G(q)u(t) + v(t), \quad u(t) = F(q)(r(t) - y(t)),$$

where $y, r \in \mathbb{R}^{n_y}$ and $u \in \mathbb{R}^{n_u}$. The nonparametric FRF estimate is then calculated from the DFT matrices $\mathbf{Y}(\omega_k) \in \mathbb{C}^{n_y \times n_u}$ and $\mathbf{U}(\omega_k) \in \mathbb{C}^{n_u \times n_u}$ like

$$\widehat{G}(e^{j\omega_k T_s}) = \mathbf{Y}(\omega_k)\mathbf{U}^{-1}(\omega_k),$$

where each column of $\mathbf{Y}(\omega_k)$ and $\mathbf{U}(\omega_k)$ contain the DFT of the sampled data from each sub-experiment (for a multivariable system, n_u sub-experiments are needed to form the FRF estimate). As excitation, an orthogonal random phase multisine signal (Dobrowiecki and Schoukens, 2005) will be used, which gives $\mathbf{R}(\omega_k) = \mathbf{R}_{\text{diag}}(\omega_k)\mathbf{T}$, where $\mathbf{R}_{\text{diag}}(\omega_k) = \text{diag}\left\{R_1(\omega_k), \ldots, R_{n_y}(\omega_k)\right\}$, each $R_l(\omega_k)$ is a random phase multisine signal, and $\mathbf{T}$ is an orthogonal matrix with $\mathbf{T}\mathbf{T}^H = n_u I$. Assuming $n_y = n_u$ and $v(t)$ being noise, independent and identically distributed for different sub-experiments, with

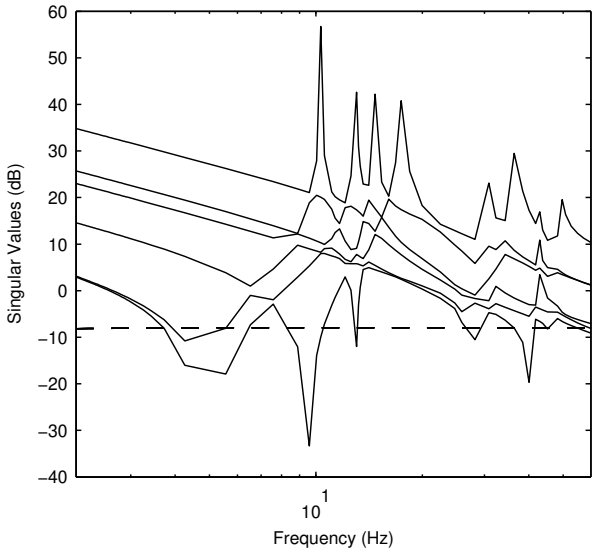


Figure 2: *Singular values for the model* (11) *in position* $q_m = 0$ *together with the inverse controller* F^{-1} *(dashed line).*

power spectrum $\Phi_v(\omega_k) \in \mathbb{C}^{n_y \times n_y}$ gives the uncertainty $\Lambda_{\widehat{G}}^{(i)}(\omega_k)$ in the nonparametric FRF estimate $\widehat{G}^{(i)}(e^{j\omega_k T_s})$ as follows (omitting the frequency argument).

$$\Lambda_{\widehat{G}}^{(i)} = \mathrm{Cov}(\mathrm{vec}(\widehat{G}^{(i)})) = \frac{1}{n_u \Phi_{r_0}} \left[(G^{(i)} + F^{-1})^H (G^{(i)} + F^{-1}) \right]^T \otimes \Phi_v, \qquad (12)$$

where $\Phi_{r_0}(\omega_k) = |R_0(\omega_k)|^2$ (assuming $|R_l(\omega_k)|^2 = |R_0(\omega_k)|^2$, $l = 1, \ldots, n_y$). It should be mentioned that (12) is an approximation since the FRF will be slightly biased due to data collection under feedback control. For details and proof, see Wernholt and Gunnarsson (2007).

In the numerical illustration, $\Phi_v(\omega) = C_v \omega^2 I$ and $\Phi_{r_0}(\omega_k) = C_r$ for $N_f = 50$ logarithmically spaced frequencies between 2 and 60 Hz and zero otherwise. The uncertainty will therefore be scaled by the factor C_v/C_r. A diagonal PI controller is used and the singular values for $F^{-1}(e^{j\omega_k T_s})$ as well as $G^{(i)}(e^{j\omega_k T_s}, \theta_0)$ for position $q_m = 0$ can be seen in Figure 2. To view all 36 elements, see (Öhr et al., 2006).

For the experiment design, $[\Psi^{(i)}(\theta)]^T = \frac{\partial \mathcal{G}^{(i)}(\theta)}{\partial \theta}$ and $\Lambda_0^{(i)}$ are needed to calculate the information matrix H_i in (5). Here, the model $\mathcal{G}^{(i)}(\theta)$ is given by

$$\mathcal{G}^{(i)}(\theta) = \begin{bmatrix} \mathrm{vec}(G^{(i)}(e^{j\omega_1 T_s}, \theta)) \\ \vdots \\ \mathrm{vec}(G^{(i)}(e^{j\omega_{N_f} T_s}, \theta)) \end{bmatrix}, \qquad (13)$$

where $G^{(i)}(e^{j\omega_k T_s}, \theta)$ is the parametric FRF in (11) and $\Psi^{(i)}(\theta)$ is calculated numerically by central differences. The covariance matrix $\Lambda_0^{(i)}$ is a block-diagonal matrix with

$\Lambda_{\widehat{G}}^{(i)}(\omega_1), \ldots, \Lambda_{\widehat{G}}^{(i)}(\omega_{N_f})$ on the diagonal. As candidate positions, all possible combinations of

$$\begin{aligned} q_{m1} &= \{0,\ 45,\ 90\}, \quad q_{m2} = \{-60,\ -45,\ \cdots,\ 90\}, \\ q_{m3} &= \{-165,\ -150,\ \cdots,\ 60\}, \quad q_{m4} = \{0,\ 45,\ 90\}, \\ q_{m5} &= \{-90,\ -45,\ 0,\ 45,\ 90\}, \quad q_{m6} = 0, \end{aligned}$$

will be used, giving a total of $Q_c = 7920$ candidate positions. $q_{m6} = 0$ is used since the payload is symmetric with respect to q_{m6}. The position in Figure 1 corresponds to $q_m = 0$.

6.3 Solution using YALMIP and SDPT3

The experiment design problem (9) (or actually the dual problem (10)) will now be solved using YALMIP (Löfberg, 2004) and SDPT3 (Toh et al., 2006). Although YALMIP always assumes that a model represents the dual in the primal-dual pair, it is possible to use the *automatic dualization* feature. If this is used, the dual problem (10) is automatically derived from (9) and solved, conveniently allowing us to solve the most efficient model without manually deriving any duals. This is particularly useful if we want to change the original problem (9) by adding additional constraints or terms in the objective, without deriving a new dual problem by hand. The original λ variables are automatically recovered by YALMIP from dual variables in SDPT3.

As we mentioned in Section 5, the computational complexity of solving the design problem is roughly linear when dualized, and cubic when interpreted in the naive original form. Although a dualized model gives us several orders of magnitudes faster computations, there is still room for improvement. The solver SDPT3 is developed with large sparse problems in mind, whereas our problem turns out to be dense. By changing one line of linear algebra code in SDPT3 from sparse to dense, the solver ran approximately four times faster on the dualized model. The optimization models were evaluated experimentally[4] and the computation times are illustrated in Figure 3. Note that the primal model only could be solved up to $Q_c \approx 4000$ due to memory limitations. The curves confirm the linear growth in Q_c for the dualized models, whereas the naive primal formulation grows slower than the expected cubic, but still requires far more computational effort. This simply indicates that there are other major contributors to the computations than the solution of the linear system, which was used for the crude asymptotic complexity estimate.

6.4 Results

The solution to the relaxed experiment design problem (9) with $Q_c = 7920$ candidate positions gives a λ^* with 15 nonzero elements. The corresponding positions and λ^*-values are presented in Table 1 together with the geometric mean $(\det H_i^{-1})^{1/d}$ as a measure of the covariance matrix. In addition are the positions presented with min, max, and median covariance. These values should be compared with a covariance value of 1 for

[4] Using a 1.7 GHz Pentium M processor with 1 GB RAM. A complete implementation, with illustrative data, can be found at http://control.ee.ethz.ch/~joloef/wiki/pmwiki.php?n=Examples.Experiment

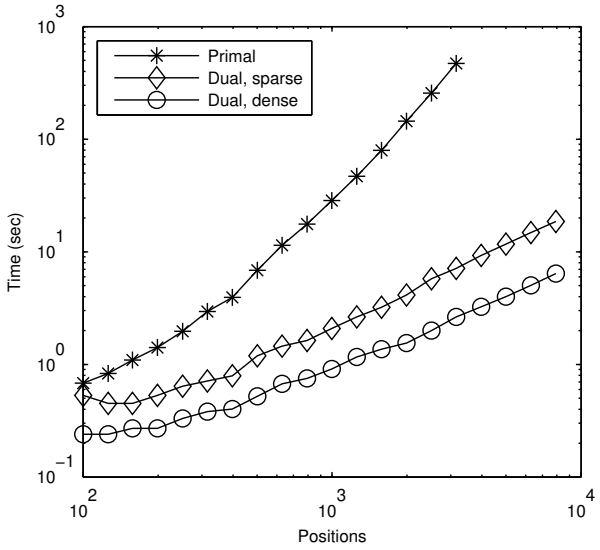


Figure 3: *Computation time when solving the primal and dual problems as a function of the number of candidate positions Q_c.*

the optimal solution[5] to the relaxed problem. The 15 optimal positions are also illustrated in Figure 5.

As was mentioned in Section 4, a sub-optimal solution to the original combinatorial experiment design problem can be obtained by using Algorithm E.1. The resulting covariance, $V(\tilde{m})$ with $V(x) = (\det \sum_i x_i H_i)^{-1/d}$, is plotted in Figure 4 together with the relaxed value $V(M\lambda^*)$ as a function of M. In addition are the corresponding values plotted for the optimal single position (min in Table 1) and the mean value using M random positions.

The difference between the combinatorial problem, $V(\tilde{m})$, and the relaxed problem, $V(M\lambda^*)$, gets small already for quite small M values ($M \approx 10$). The accuracy obtained by performing $M = 10$ experiments distributed over the optimal positions (then actually ≤ 10 positions) corresponds to almost 60 experiments in the optimal single position. Or the other way, 3 experiments in different positions give the same accuracy as 10 experiments in the optimal single position.

In addition, using multiple positions will probably make the experiment design more robust to uncertainties in θ_0, but that analysis is left for future work. To robustify the experiment design, one could also add the convex constraint that no more than a certain fraction of the M experiments is concentrated in less than a given fraction of the Q_c possible experiments, see Vandenberghe et al. (1998, Section 2.4) for details[6].

If the user has no knowledge about the system, multiple random positions might be a

[5] A covariance value of 1 corresponds to $\Phi_v(\omega_k)/\Phi_r(\omega_k) \approx 58 \cdot \omega_k^2$, which means a really poor SNR. In the real application, the SNR is large which will give a much better parameter accuracy.

[6] Unfortunately, the structure in the design problem that was exploited to derive an efficient dual formulation will typically be lost when adding additional complicated constraints.

Table 1: *The solution to the relaxed experiment design problem* (9)*. The covariance is the geometric mean* $(\det H_i^{-1})^{1/d}$*. In addition are the positions presented with min, max, and median covariance.*

Pos	q_{m1}	q_{m2}	q_{m3}	q_{m4}	q_{m5}	Cov	λ
1	0	−30	−75	0	0	15.4	0.1083
2	45	−60	−30	0	45	19.6	0.0256
3	45	−60	−30	45	45	21.2	0.0093
4	45	−30	−30	45	−45	6.3	0.1606
5	45	−30	60	45	90	7.4	0.0509
6	45	15	−150	90	90	20.1	0.0350
7	45	15	−15	90	−45	26.6	0.0881
8	45	45	−150	45	−90	15.4	0.0895
9	45	45	−105	0	−90	26.4	0.0927
10	90	−60	−60	90	90	23.9	0.0451
11	90	−30	−30	90	90	12.9	0.0238
12	90	−30	0	90	−45	40.6	0.0225
13	90	0	−105	45	−90	23.9	0.0192
14	90	30	30	0	90	11.7	0.1300
15	90	60	15	0	0	24.7	0.0993
Min	45	15	−135	45	45	5.9	
Med	45	30	−90	90	−90	122.6	
Max	0	60	−90	90	0	673.6	

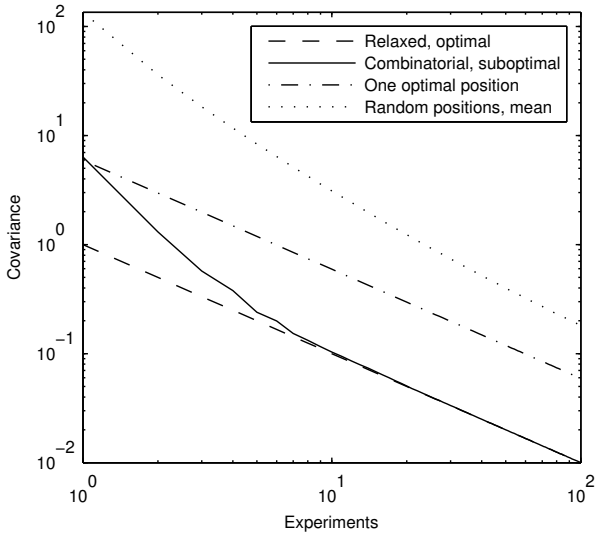


Figure 4: *Covariance,* $(\det H)^{-1/d}$*, as a function of the number of experiments* M*.*

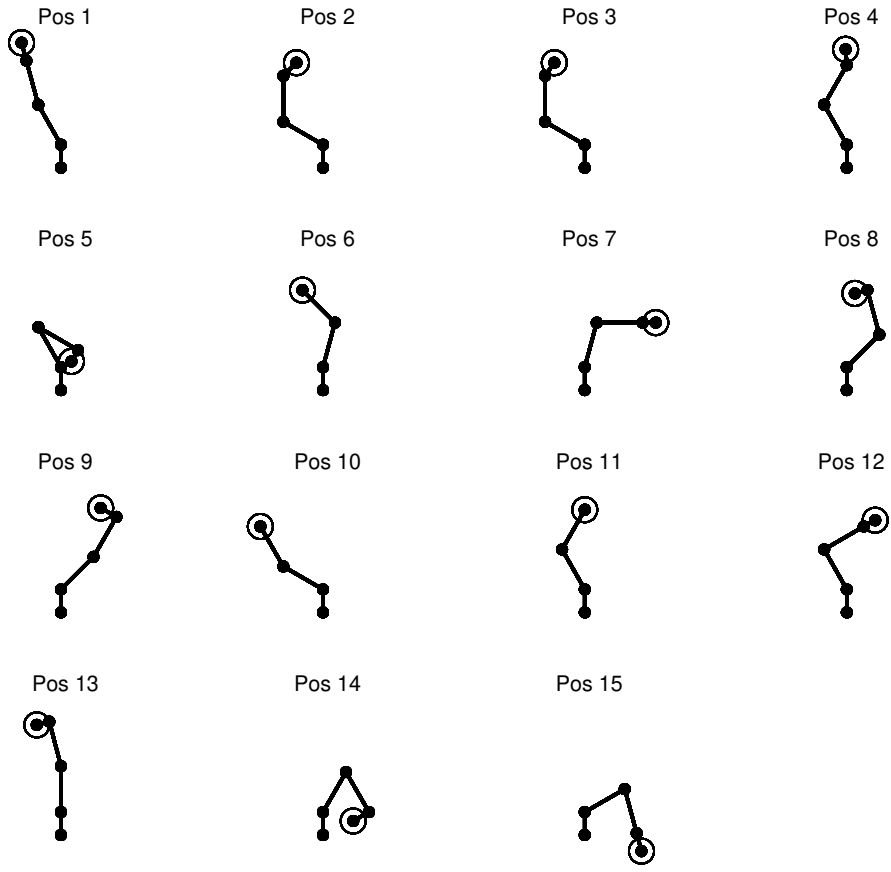


Figure 5: *Illustration of the optimal positions from Table 1, here viewed from the side. The payload is marked by a circle.*

good alternative, compared to using one random position, which is indicated by the dotted curve in Figure 4. Using 1 random position for $M = 100$ experiments will in average give 10 times worse accuracy compared to using 100 random positions. This might not be so surprising, but it is interesting to notice that with $M = 100$, the optimal single position is only 3 times better than using 100 random positions.

7 Conclusion

Experiment design involving selection of optimal experiment positions has been treated. From the derived Fisher information matrix, a convex optimization problem is posed. By considering the dual problem, the experiment design is efficiently solved with linear complexity in the number of candidate positions, compared to cubic complexity for the primal problem. In the numerical illustration, the 15 optimal positions reduce the pa-

rameter uncertainty by a factor of 6, compared to using the optimal single position in all experiments.

Acknowledgments

This work was supported by ABB AB – Robotics and VINNOVA's Center of Excellence ISIS at Linköping University. The authors also want to thank Martin Enqvist for inspiring discussions.

References

Åström, K. and Wittenmark, B. (1984). *Computer Controlled Systems – Theory and Design*. Prentice-Hall, Englewood Cliffs, New Jersey.

Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM series on Optimization. SIAM, Philadelphia, Pennsylvania.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Dobrowiecki, T. and Schoukens, J. (2005). Measuring linear approximation to weakly nonlinear MIMO systems. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Fedorov, V. V. (1972). *Theory of optimal experiments*. Academic Press, New York.

Gevers, M. (2005). Identification for control: From the early achievements to the revival of experiment design. *European Journal of Control*, 11(4-5):335–352.

Goodwin, G. C. and Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press, New York.

Goodwin, G. C., Rojas, J. S., and Welsh, J. S. (2006). Good, bad and optimal experiments for identification. In Glad, T. and Hendeby, G., editors, *Forever Ljung in System Identification*, pages 103–125. Studentlitteratur, Lund.

Jansson, H. and Hjalmarsson, H. (2005). Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, 50(10):1534–1549.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Löfberg, J. (2004). YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proc. CACSD Conference*, Taipei, Taiwan. Available from `http://control.ee.ethz.ch/~joloef/yalmip.php`.

Öhr, J., Moberg, S., Wernholt, E., Hanssen, S., Pettersson, J., Persson, S., and Sander-Tavallaey, S. (2006). Identification of flexibility parameters of 6-axis industrial manipulator models. In *Proc. ISMA2006 International Conference on Noise and Vibration Engineering*, pages 3305–3314, Leuven, Belgium.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach*. IEEE Press, New York.

Toh, K. C., Tütüncü, R. H., and Todd, M. J. (2006). On the implementation and usage of SDPT3 – a MATLAB software package for semidefinite-quadratic-linear programming, version 4.0. Available from `http://www.math.nus.edu.sg/~mattohkc/guide4-0-draft.pdf`.

Vandenberghe, L., Boyd, S., and Wu, S.-P. (1998). Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19(2):499–533.

Wernholt, E. and Gunnarsson, S. (2007). Analysis of methods for multivariable frequency response function estimation in closed loop. In *46th IEEE Conference on Decision and Control*, New Orleans, Louisiana. Accepted for publication.

Paper F

Nonlinear Gray-Box Identification of a Flexible Manipulator

Edited version of the paper:

> Wernholt, E. and Gunnarsson, S. (2007c). Nonlinear gray-box identification of a flexible manipulator. *Submitted to Control Engineering Practice*.

Parts of the paper in:

> Wernholt, E. and Gunnarsson, S. (2006b). Nonlinear identification of a physically parameterized robot model. In *Proc. 14th IFAC Symposium on System Identification*, pages 143–148, Newcastle, Australia.
>
> Wernholt, E. and Gunnarsson, S. (2005). Nonlinear grey-box identification of industrial robots containing flexibilities. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Preliminary version published as Technical Report LiTH-ISY-R-2739, Dept. of Electrical Engineering, Linköping University, SE–581 83 Linköping, Sweden.

Nonlinear Gray-Box Identification of a Flexible Manipulator

Erik Wernholt and Svante Gunnarsson

Dept. of Electrical Engineering,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: {erikw,svante}@isy.liu.se.

Abstract

A three-step procedure for nonlinear gray-box identification of an industrial manipulator containing flexibilities is studied. The aim of the first two steps is to obtain good initial values for the third prediction error minimization step. In the first step, rigid body dynamics and friction are identified using a separable least-squares method. In the second step, initial values for flexibilities are obtained using an inverse eigenvalue method. Finally, in the last step, the remaining parameters of a nonlinear gray-box model are identified directly in the time domain using prediction error minimization. The identification procedure is exemplified using real data from an experimental industrial robot. The estimated physical parameters have realistic numerical values and give a model with good correspondence to FRF measurements.

Keywords: Identification, Robotics, Flexible arms, Friction, Manipulators

1 Introduction

System identification in robotics is a vast research area and can be divided into, at least, three different levels or application areas. These levels involve the estimation of the kinematic description, the dynamic model (often divided into rigid body and flexible body dynamics), and the joint model (e.g., motor inertia, gearbox elasticity and backlash, motor characteristics, and friction parameters). An overview of identification in robotics can be found in Kozlowski (1998), and examples of previous work are also given in Section 4 in connection with the presentation of the steps in the identification procedure.

The work reported here considers identification of rigid body dynamics, friction, and flexibilities using a three-step procedure. The procedure was first introduced in Wernholt and Gunnarsson (2005), and it is here extended in two ways. First, a nonlinear friction

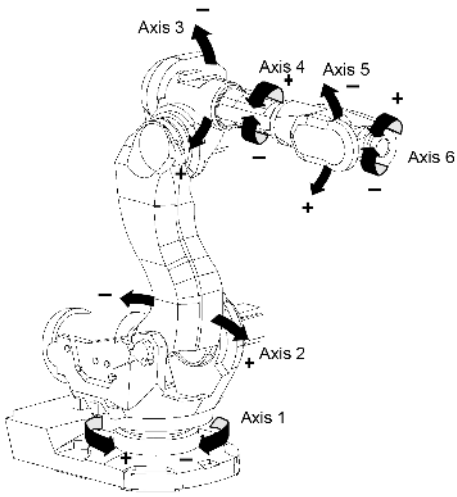


Figure 1: *The ABB manipulator IRB6600.*

model is included, leading to a pseudo-linear regression, and, second, the example is based on a three-mass model. An early version of the results presented here was also given in Wernholt and Gunnarsson (2006).

The problem treated here is closely related to the problems considered in, for example, Östring et al. (2003) and Isaksson et al. (2003). In Östring et al. (2003), a method is applied where inertial parameters as well as parameters describing the flexibility of a three-mass model are identified directly in the time domain. However, only linear models were considered in their work, and the initial values in the identification were selected in an *ad hoc* manner. Isaksson et al. (2003) consider gray-box identification of a two-mass model with backlash, where black-box modeling is used to find initial parameter values. The three-step procedure is also utilized in Gunnar et al. (2006) where an actuator model, including hysteresis, is identified for a particular type of parallel robot structure.

2 Problem Description

The starting point for the paper is an industrial robot of the type shown in Figure 1. For simplicity the presentation will consider a situation where the robot movements are not affected by gravity. This is the case when the movement only involves the first joint, denoted Axis 1 in Figure 1, and this is the situation that will be considered here. The case when also the gravity is involved will be dealt with in future work. Previous work, see, e.g., Östring et al. (2003), has shown that a three-mass model can give a sufficiently accurate description of the robot dynamics for movements involving the first joint. It should however be noted that the results in Östring et al. (2003) are based on data from a smaller robot (ABB IRB1400), while here data will be collected from an experimental robot of essentially larger size. Broadband excitation indicates that a higher model order might be needed in some cases, see Figure 5. It is however assumed here that a three-mass

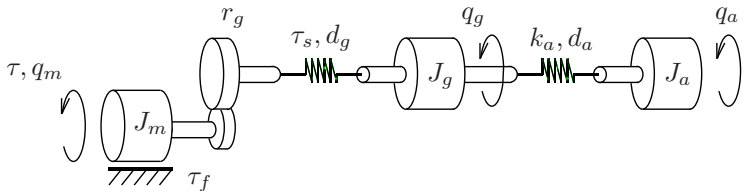


Figure 2: *The three-mass flexible model of the robot arm.*

model gives a sufficiently accurate description in the frequency range up to 20 Hz. The robot dynamics for movements involving the first joint will hence be described by the model in Figure 2. Gear ratio $r_g = 1$ is used in the model, which gives better numerical properties and easier notation. The true physical parameters can later be obtained by a simple scaling with the true gear ratio.

The input is the torque τ generated by the electrical motor, while the output is the motor velocity $\dot{q}_m$. In practice only the motor angle is measured, but since the noise level is low this signal can be differentiated with sufficiently good result. The velocities of the other masses, $\dot{q}_g$ and $\dot{q}_a$, are not measurable. Flexibility in the gearbox is modeled by a nonlinear spring, $\tau_s(\cdot)$, between the motor and the second mass. The spring between the second and third mass represents flexibilities in the arm structure. Finally, friction in the system is modeled by nonlinear friction, $\tau_f(\cdot)$, acting on the first mass. Applying torque balances for the three masses and introducing the states

$$x(t) = \begin{bmatrix} q_m(t) - q_g(t) \\ q_g(t) - q_a(t) \\ \dot{q}_m(t) \\ \dot{q}_g(t) \\ \dot{q}_a(t) \end{bmatrix}, \tag{1}$$

gives the nonlinear state-space model

$$\dot{x} = \begin{bmatrix} x_3 - x_4 \\ x_4 - x_5 \\ \frac{1}{J_m}\left(-\tau_s(x_1) - d_g(x_3 - x_4) - \tau_f(x_3) + u\right) \\ \frac{1}{J_g}\left(\tau_s(x_1) + d_g(x_3 - x_4) - k_a x_2 - d_a(x_4 - x_5)\right) \\ \frac{1}{J_a}\left(k_a x_2 + d_a(x_4 - x_5)\right) \end{bmatrix}, \tag{2}$$

$$y = x_3, \tag{3}$$

where J_m, J_g, and J_a are the moments of inertia of the motor, gear, and arm, respectively, $u = \tau$ is the motor torque, d_g and d_a are damping parameters, and k_a is the stiffness of the second spring. The spring and gear friction torques, τ_s and τ_f respectively, are often approximated by linear models (see, e.g., Östring et al., 2003). In this work, a nonlinear friction model from Feeny and Moon (1994) will be used to capture Coulomb friction and the Striebeck effect[1] as

$$\tau_f(x_3) = F_v x_3 + (F_c + F_{cs}\,\mathrm{sech}(\alpha x_3))\tanh(\beta x_3), \tag{4}$$

[1] Decreasing friction with increasing velocity at low velocity.

where $\operatorname{sech}(x) = 1/\cosh(x)$, F_v and F_c are the viscous and Coulomb friction coefficients, F_{cs} and α are used to model the Striebeck effect, and β is used to get a smooth model without discontinuity at zero velocity, which is more suitable for simulation. As a comparison, two simpler friction models

$$\tau_f(x_3) = F_v x_3 + F_c \tanh(\beta x_3), \tag{5}$$
$$\tau_f(x_3) = F_v \dot{x}_3 + F_c \operatorname{sign}(x_3), \tag{6}$$

will be used as well. For details on friction modeling, see Armstrong-Hélouvry et al. (1994). A recent friction model can also be found in Makkar et al. (2005). The torque of the spring is modeled as

$$\tau_s(x_1) = k_{g1} x_1 + k_{g3} x_1^3, \tag{7}$$

where k_{g1} and k_{g3} are the parameters of the spring.

3 Nonlinear Gray-Box Identification

The aim is to identify the parameters in the robot model, described in Section 2, using experimental data. The starting point for the nonlinear gray-box identification is the continuous-time state-space model structure

$$\dot{x}(t) = f(t, x(t), u(t), \theta), \tag{8a}$$
$$y(t) = h(t, x(t), u(t), \theta) + e(t), \tag{8b}$$

where f and h are nonlinear functions. $x(t)$ is the state vector, $u(t)$ and $y(t)$ are input and output signals, $e(t)$ a white measurement noise signal, and t denotes time. Finally θ is the vector of unknown parameters. Given a set of input/output-data the aim is to determine the parameter vector that minimizes a criterion like

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon^2(t, \theta), \tag{9}$$

where $\varepsilon(t)$ denotes the prediction error

$$\varepsilon(t, \theta) = y(t) - \hat{y}(t, \theta). \tag{10}$$

The experiments presented in this paper will utilize the nonlinear gray-box model structure IDNLGREY, available in the System Identification Toolbox (SITB) (Ljung, 2007). The model structure IDNLGREY is similar to the IDGREY model structure in SITB. The model can be either a discrete-time or continuous-time state-space model, and it is defined in a MATLAB m-file/mex-file. In the current version of the software, only OE-models can be used, i.e., only additive white noise, $e(t)$, on the output. The prediction $\hat{y}(t|\theta)$ then becomes the simulated output of the model (8) with the input $u(t)$, without $e(t)$, for the current parameter vector θ. The criterion (9) is minimized by an iterative numerical search algorithm which involves simulation of the system for different values of θ. The user specifies an initial parameter vector and it is also possible to fix some components in θ, as well as include min and max values for each parameter. To speed up the numerical

optimization, the simulation model is implemented in a mex-file (C-code). An inherent problem of the iterative search needed for minimization of the prediction error criterion (9) is that only convergence to a local minimum can be guaranteed. In order to converge to the global minimum, good initial parameter estimates are needed. Therefore a three-step identification procedure is proposed in the following section.

4 Three-Step Identification Procedure

The first step of the procedure considers identification of rigid body dynamics and friction. In the second step, initial values for flexibilities are obtained, and finally, in the third step, the nonlinear gray-box identification procedure is applied.

4.1 Step 1: Rigid Body Dynamics and Friction

There exists a vast amount of literature on the identification of rigid body dynamics (see, e.g., Gautier and Poignet, 2001; Grotjahn et al., 2001; Kozlowski, 1998; Swevers et al., 1997). The standard procedure includes a dynamic model

$$H_{rb}(q_m, \dot{q}_m, \ddot{q}_m)\theta_{rb} = \tau, \tag{11}$$

which is linear in the rigid body parameters θ_{rb}. Each link gives ten physical parameters. This representation is redundant, but there are methods to find a minimal dimensional parameter vector θ_b, called *base parameters*, that characterize the dynamic model like

$$H_b(q_m, \dot{q}_m, \ddot{q}_m)\theta_b = \tau. \tag{12}$$

The base parameters are nonlinear functions of the physical parameters like

$$\theta_b = \phi(\theta_{rb}). \tag{13}$$

In the presence of friction, τ in (12) should be replaced by $\tau - \tau_f(\dot{q}_m)$. Usually the friction model (6) is used, which gives two additional parameters per link to estimate, but still a linear regression. This model is not sufficient to correctly describe dynamic friction (see Armstrong-Hélouvry et al., 1994), but compensates the major frictional effects on the identification of rigid body dynamics when using a high amplitude excitation. The robot is moved along some (optimized) trajectory and applied torque and joint movements are recorded. The parameters are then estimated using linear regression.

Here, the more advanced friction model (4) will be used in order to more accurately model the low velocity behavior. For the robot model in Section 2, the rigid body dynamics and friction is

$$(J_m + J_g + J_a)\ddot{q}_m + \tau_f(\dot{q}_m) = \tau, \tag{14}$$

which can be written as a pseudo-linear regression

$$\hat{\tau}(t|\rho, \eta) = \rho^T \varphi(t, \eta), \tag{15}$$

when

$$\varphi(t,\eta) = \begin{bmatrix} \ddot{q}_m \\ \dot{q}_m \\ \tanh(\beta\dot{q}_m) \\ \operatorname{sech}(\alpha\dot{q}_m)\tanh(\beta\dot{q}_m) \end{bmatrix},$$
$$\rho = \begin{bmatrix} J & F_v & F_c & F_{cs} \end{bmatrix}^T, \quad \eta = \begin{bmatrix} \alpha & \beta \end{bmatrix}^T.$$

Here, J is the only base parameter and (13) simplifies to $J = J_m + J_g + J_a$.

Minimizing an identification criterion like

$$V_N(\rho,\eta) = \sum_{t=1}^{N} |\tau(t) - \rho^T\varphi(t,\eta)|^2 = |\boldsymbol{\tau} - \Phi(\eta)\rho|^2, \tag{16}$$

then is a separable least squares problem since the least squares part ρ can be separated out using

$$\hat{\rho} = [\Phi^T(\eta)\Phi(\eta)]^{-1}\Phi^T(\eta)\boldsymbol{\tau}, \tag{17}$$

and the problem is reduced to finding the optimal η. See Golub and Pereyra (1973) for a thorough treatment of this approach.

The result from the first step is estimates of the base parameters $\theta_b = J$ and the friction parameters $\theta_{fr} = \begin{bmatrix} F_v & F_c & F_{cs} & \alpha & \beta \end{bmatrix}^T$.

4.2 Step 2: Initial Values for Flexibilities

The major flexibilities in an industrial robot are normally located at the joint level, due to the transmission. A two-mass model (or coupled two-mass models for multivariable cases) is then sufficient to describe the dynamics. Weaker (more compliant) robot structures will in addition introduce significant flexibilities in the links and their connections. Therefore higher-order models are sometimes needed in order to get a sufficient description of the system. Identification of flexibilities is more involved than the identification of rigid body dynamics. The main reason is that only a subset of the state variables now are measured and one can therefore not use linear regression. Many different methods are described in the literature to handle this problem, see for example Behi and Tesar (1991), Johansson et al. (2000), and Albu-Schäffer and Hirzinger (2001). They differ in, for example, assumed model structure, required measurement signals, and complexity of the identification method.

Here, a method described in Berglund and Hovland (2000) and Hovland et al. (2001) will be used for the identification of masses, springs and dampers, only using applied torque and joint movements. The identification is based on an estimated nonparametric Frequency Response Function (FRF) together with the solution of an inverse eigenvalue problem. Consider an nth order spring-mass system, which can be modeled as

$$M\ddot{q}_m + Kq_m = \tau, \tag{18}$$

where M is a diagonal inertia matrix, q_m is a vector of the position of the masses, and K is a tridiagonal matrix built up from the $n-1$ spring constants. The eigenvalues of the

system are given by using $\ddot{q}_m = -\lambda q_m$, $\tau = 0$, as

$$Bu = \lambda u, \tag{19}$$

where $B = L^{-1}KL^{-T}$, $u = L^T q_m$, and $L = M^{1/2}$. Using the nonparametric FRF gives information about the resonance and anti-resonance frequencies. The first eigenvector u_1 of B can be derived from the squared resonance and anti-resonance frequencies, assuming that the damping is negligible. The B matrix can then be computed recursively by the Lanczos algorithm using u_1 and the squared resonance frequencies. Finally, the elements of M and K can be reconstructed by using their special structure. A damping matrix can be estimated afterwards by minimizing the discrepancy between the nonparametric FRF and the parametric model at the resonance and anti-resonance frequencies, using some optimization routine. For details about the method, see Berglund and Hovland (2000) and Hovland et al. (2001) and the references therein.

The result from the second step gives initial estimates of (some) rigid body parameters, θ_{rb}, and parameters describing the flexibilities, θ_{fl}, typically springs and dampers. As an optional step, one could of course refine all these estimated parameters by curve fitting between the nonparametric FRF and the parametric model as well.

4.3 Step 3: Nonlinear Gray-Box Identification

Combining the estimates from steps 1 and 2 gives an initial parameter estimate, and the nonlinear gray-box identification method described in Section 3 can now be applied. To reduce the complexity, parameters estimated with high accuracy in the previous steps can be fixed in this step, leading to a lower dimensional iterative search. Keeping $\hat{\theta}_b$ and $\hat{\theta}_{fr}$ from step 1 fixed will result in a modified criterion

$$\min_{\theta} V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon^2(t,\theta), \tag{20}$$

$$\text{subject to } \phi(\theta_{rb}) - \hat{\theta}_b = 0, \tag{21}$$

$$\theta_{fr} - \hat{\theta}_{fr} = 0, \tag{22}$$

where $\theta = \begin{bmatrix}\theta_{fr}^T & \theta_{rb}^T & \theta_{fl}^T\end{bmatrix}^T$.

5 Data Collection

The data used for identification are real data collected from an experimental robot. The open-loop system to be identified is unstable, which makes it necessary to collect data while the robot controller is running in closed loop. An experimental control system is used, which makes it possible to use off-line computed reference signals for the controllers. For the different steps in the identification procedure, different excitation signals are needed.

In step 1, the rigid body dynamics and friction parameters should be excited without introducing any oscillations due to the flexibilities. Therefore a low frequency excitation is preferred.

Table 1: *Estimated parameters from step 1, including one standard deviation.*

J	0.0329 ± 0.0027	F_{cs}	3.9863 ± 4.5876
F_v	0.0099 ± 0.0074	α	3.2402 ± 0.5469
F_c	0.7430 ± 0.3774	β	0.7994 ± 0.2685

In step 2, on the other hand, the whole frequency band should be excited where notch and peak frequencies in the FRF are expected. Since nonparametric FRF estimates are based on an assumption of linearity, the influence of nonlinear friction should be reduced, so a broadband excitation with as few zero velocity crossings as possible is selected.

Finally, for step 3 a data set (or a combination of data sets) is needed that excite all free parameters in the model. The following periodic excitation signals will be used as reference speed, $\dot{q}_m^{ref}$, for the controller. They all have a period time of 10 s and one period of the steady-state response is collected, sampled at 2 kHz ($T_s = 0.5$ ms).

Data set 1: Multisine signal (sum of sinusoids) with frequencies 0.1, 0.3, and 0.5 Hz. Five different peak values have been applied: 1, 5, 20, 40, and 160 rad/s.

Data set 2: Multisine signal with a flat amplitude spectrum in the frequency interval 1–40 Hz with a peak value of 16 rad/s. The multisine signal is superimposed on a filtered square wave with amplitude 20 rad/s, frequency 0.1 Hz, and cut-off frequency 1 Hz.

Data set 3: Similar to data set 2, but without the square wave.

For details on the selection of excitation signals, see, e.g., Pintelon and Schoukens (2001) and Ljung (1999).

6 Results

The physical parameters in the robot model from Section 2 will here be identified by applying the proposed three-step identification procedure from Section 4, using the experimental data described in Section 5.

6.1 Step 1

Using data set 1, with one period of data for each of the five amplitudes, together with the separable least squares method, as described in Section 4.1, gives parameter estimates according to Table 1.

As can be seen, some of the friction parameters have a quite large standard deviation, which actually indicates that the model cannot fully describe the data sets used for identification. The parameters in the right column only affect the low velocity region, and in this region the five experiments in data set 1 with different amplitudes behave quite differently (for example smaller β for larger amplitudes). An even more advanced friction model which handles these differences is therefore needed in order to fully capture the friction dynamics.

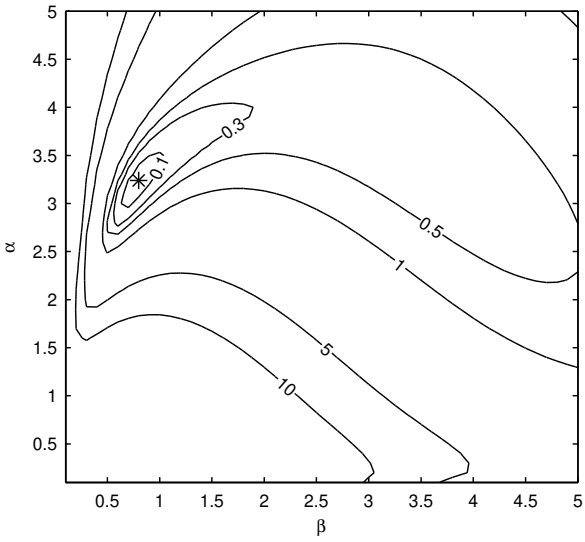


Figure 3: *Level curves showing the increase in the loss function (in percent) when moving away from the optimum.*

Table 2: *Loss function when evaluating the estimated friction models using data set 1 for different amplitudes.*

Amplitudes:	1	5	20	40
Model (4)	0.0794	0.0458	0.0287	0.0439
Model (5)	0.0897	0.0406	0.0292	0.0443
Model (6)	0.1603	0.0475	0.0275	0.0449

The separable least squares problem has no local minima in the area of interest, as can be seen in Figure 3. The loss function (9) is also relatively insensitive to different values of the two parameters since such differences can be somewhat compensated by the F_{cs} parameter.

To see the differences, the two simpler friction models (5) and (6) are estimated as well. Figure 4 shows the resulting friction curves. To evaluate the models, the loss function for some of the amplitudes are shown in Table 2. The advanced model (4) is slightly better than the 3 parameter model (5). For the 2 parameter model (6), the loss function for amplitude 1 is drastically increased. For large amplitudes, no major differences can be noted.

6.2 Step 2

In this step, first the nonparametric FRF must be estimated. Before the estimation, the input signal should be modified according to the estimated friction model from step 1. This gives minor differences for data set 2, but for data set 3 the resonances get more

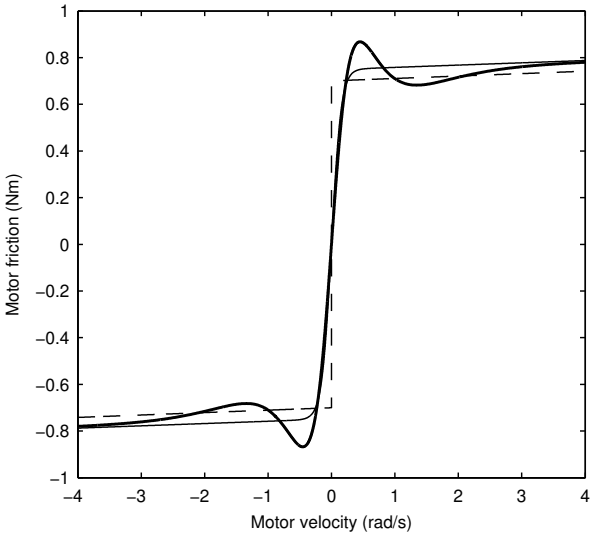


Figure 4: *Estimated friction models. Thick line:* (4)*, thin line:* (5)*, dashed line:* (6)*.*

visible. The FRF for data set 2 from motor torque to motor acceleration can be seen in Figure 5 on page 198. Solving the inverse eigenvalue problem according to Section 4.2 gives initial estimates of $\theta_{rb} = \begin{bmatrix} J_m & J_g & J_a \end{bmatrix}^T$ and $\theta_{fl} = \begin{bmatrix} k_{g1} & k_{g3} & k_a & d_g & d_a \end{bmatrix}^T$, except for k_{g3} which is kept zero at this step due to the assumption about linearity. These values can be seen in Table 3, model `m5init`;

6.3 Step 3

Combining the estimates from steps 1 and 2 gives the initial model `m5init` in Table 3. The nonlinear gray-box identification according to Section 4.3 is then applied, using data set 2. First, $k_{g3} = 0$ is used, giving the model `m5`. Releasing k_{g3} gives the model `m5kg3`. For comparison, models are also estimated using the simpler friction models from Section 6.1, giving models `m3` and `m2`, where the number indicates the number of friction parameters. The rigid body parameters, θ_{rb}, and the second damper are estimated with a standard deviation less than 10 %. For the remaining parameters in θ_{fl}, the standard deviation is less than 2 %.

To validate the estimated models, four additional realizations of the multisine signal have been applied, both for data set 2 and 3. The loss function for each of the models in Table 3, including the rigid model from step 1 have then been calculated for each of the realizations. Their values have first been normalized by the loss using the rigid model, and secondly averaged over the realizations to get the average loss, shown in Table 4. Comparing `m5init` and `m5` shows the big improvement by using step 3 in the procedure. For data set 2, the more advanced friction model (4) reduces the loss function, compared to using (6). For data set 3, on the other hand, there are no major differences. This can partly be understood by the fact that a correct friction model is more important

Table 3: *Estimated parameters, where* `m5init` *comes from the first two steps and the other models are estimated in step 3. Parameters grouped as* θ_{fr}, θ_{rb}, *and* θ_{fl}.

	m5init	m5	m3	m2	m5kg3	Scaling
F_v	0.986	0.986	0.979	1.04	0.986	$(\cdot 10^{-2})$
F_c	0.743	0.743	0.748	0.700	0.743	
F_{cs}	3.99	3.99	0	0	3.99	
α	3.24	3.24	0	0	3.24	
β	0.799	0.799	6.67	∞	0.799	
J_m	5.90	8.36	8.36	8.63	8.29	$(\cdot 10^{-3})$
J_g	20.1	20.5	20.5	21.6	20.6	$(\cdot 10^{-3})$
J_a	6.87	4.03	4.06	2.65	4.06	$(\cdot 10^{-3})$
k_{g1}	20.6	21.3	21.3	20.5	20.3	
k_{g3}	0	0	0	0	13.1	
k_a	20.2	15	15.1	11.1	15.2	
d_g	62.4	50.9	51.5	50.9	46.2	$(\cdot 10^{-3})$
d_a	9.88	5.15	5.06	1.69	4.54	$(\cdot 10^{-3})$

Table 4: *Average loss when evaluating the models using new realizations of data sets 2 and 3.*

	m5init	m5	m3	m2	m5kg3
Data set 2	0.5574	0.2727	0.2746	0.3133	0.2705
Data set 3	0.4441	0.1658	0.1661	0.1656	0.1638

for open-loop simulations if one must track a slowly varying signal, like the square wave. Using acceleration as output from the model would probably give a different result. The nonlinear spring stiffness reduces the loss function, as can be seen by comparing the loss for `m5` and `m5kg3`. Using a data set with a different amplitude would probably show an even bigger difference.

The discontinuous friction model (6) gives problems in simulation and therefore (5) with a large β is used instead. This is actually another reason for using the more advanced friction model (4), besides that it gives better performance.

Finally, a Bode diagram of the estimated model `m5` can be seen in Figure 5 together with the estimated nonparametric FRF for data set 2.

6.4 Discussion

Using the first step, where the rigid body dynamics and friction are estimated, is really important. The total inertia, $J = J_m + J_g + J_a$, as well as all the friction parameters are kept fixed during the last step when applying the nonlinear gray-box identification. Another option is to include data set 1 in the last step and allow all parameters to be estimated, which has been treated in Wernholt and Gunnarsson (2005).

The method described in Section 4.2, where an inverse eigenvalue problem is solved, is quite sensitive to changes in the FRF. Large variations in the identified parameters have

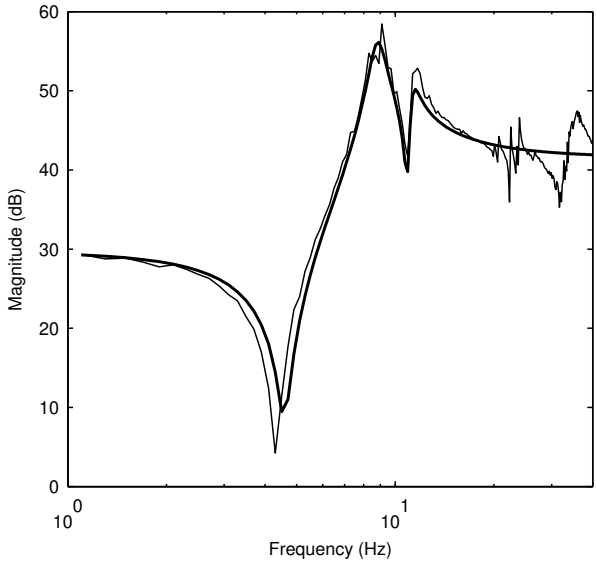


Figure 5: *Magnitude of the FRF for data set 2 from motor torque to motor acceleration (thin line) together with the estimated model* `m5` *(thick line).*

been observed, even for a fairly small perturbation of the location of the resonances. This method should therefore be replaced by, or complemented by, curve fitting the parametric model and the estimated nonparametric FRF.

During the iterative numerical search in the last step, it is important to ensure that every model that is tried by the optimization method is stable. Therefore, a constraint that all parameters must be positive is used.

As future work remains to apply this procedure to a multivariable system, or at least to an axis affected by gravity. One big problem to tackle then is that the model is unstable, which requires some changes to ensure a stable predictor. There can also be numerical problems since the dimension of the problem is larger, both in terms of states and parameters.

7 Conclusion

A three-step identification procedure has been used for the identification of rigid body dynamics, friction, and flexibilities, only using measurements on the motor side. The procedure has been exemplified using experimental data from an industrial robot together with a flexible three-mass model where nonlinear spring stiffness and a nonlinear friction model have been added. The estimated physical parameters have realistic numerical values and give a model with good correspondence to FRF measurements. To further improve the results the friction could be modeled even more accurately.

Acknowledgments

This work was supported by VINNOVA's Center of Excellence ISIS at Linköping University and the Swedish Research Council (VR).

References

Albu-Schäffer, A. and Hirzinger, G. (2001). Parameter identification and passivity based joint control for a 7DOF torque controlled light weight robot. In *Proc. 2001 IEEE International Conference on Robotics and Automation*, pages 2852–2858, Seoul, Korea.

Armstrong-Hélouvry, B., Dupont, P., and Canudas de Wit, C. (1994). A survey of models, analysis tools and compensation methods for the control of machines with friction. *Automatica*, 30(7):1083–1138.

Behi, F. and Tesar, D. (1991). Parametric identification for industrial manipulators using experimental modal analysis. *IEEE Transactions on Robotics and Automation*, 7(5):642–652.

Berglund, E. and Hovland, G. E. (2000). Automatic elasticity tuning of industrial robot manipulators. In *Proc. 39th IEEE Conference on Decision and Control*, pages 5091–5096, Sydney, Australia.

Feeny, B. and Moon, F. (1994). Chaos in a forced dry-friction oscillator: experiments and numerical modelling. *Journal of Sound and Vibration*, 170(3):303–323.

Gautier, M. and Poignet, P. (2001). Extended kalman filtering and weighted least squares dynamic identification of robot. *Control Engineering Practice*, 9(12):1361–1372.

Golub, G. and Pereyra, V. (1973). The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2):413–432.

Grotjahn, M., Daemi, M., and Heimann, B. (2001). Friction and rigid body identification of robot dynamics. *International Journal of Solids and Structures*, 38:1889–1902.

Gunnar, J., Wernholt, E., Hovland, G., and Brogårdh, T. (2006). Nonlinear grey-box identification of linear actuators containing hysteresis. In *Proc. 2006 IEEE International Conference on Robotics and Automation*, pages 1818–1823, Orlando, Florida.

Hovland, G. E., Berglund, E., and Hanssen, S. (2001). Identification of coupled elastic dynamics using inverse eigenvalue theory. In *Proc. 32nd International Symposium on Robotics (ISR)*, pages 1392–1397, Seoul, Korea.

Isaksson, A., Lindkvist, R., Zhang, X., Nordin, M., and Tallfors, M. (2003). Identification of mechanical parameters in drive train systems. In *Proc. 13th IFAC Symposium on System Identification*, pages 1542–1547, Rotterdam, The Netherlands.

Johansson, R., Robertsson, A., Nilsson, K., and Verhaegen, M. (2000). State-space system identification of robot manipulator dynamics. *Mechatronics*, 10(3):403–418.

Kozlowski, K. (1998). *Modelling and identification in robotics.* Advances in Industrial Control. Springer, London.

Ljung, L. (1999). *System Identification: Theory for the User.* Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition.

Ljung, L. (2007). *System Identification Toolbox 7 – User's Guide.* The MathWorks, Inc., Natick, MA, USA.

Makkar, C., Dixon, W., Sawyer, W., and Hu, G. (2005). A new continuously differentiable friction model for control systems design. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 600–605, Monterey, CA.

Östring, M., Gunnarsson, S., and Norrlöf, M. (2003). Closed-loop identification of an industrial robot containing flexibilities. *Control Engineering Practice*, 11:291–300.

Pintelon, R. and Schoukens, J. (2001). *System identification: a frequency domain approach.* IEEE Press, New York.

Swevers, J., Ganseman, C., Tükel, D. B., De Schutter, J., and Van Brussel, H. (1997). Optimal robot excitation and identification. *IEEE Transactions on Robotics and Automation*, 13(5):730–740.

Wernholt, E. and Gunnarsson, S. (2005). Nonlinear grey-box identification of industrial robots containing flexibilities. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic.

Wernholt, E. and Gunnarsson, S. (2006). Nonlinear identification of a physically parameterized robot model. In *Proc. 14th IFAC Symposium on System Identification*, pages 143–148, Newcastle, Australia.

Notation

Note that all vectors are column vectors. The same symbol can be used for different purposes. The principal notation is listed below, any deviations from this is explained in the text.

Symbols and Operators

$\mathbb{Z}$	The set of integers
$\mathbb{N}$	The set of natural numbers ($0 \in \mathbb{N}$)
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
t	Time variable
s	Laplace transform variable
z	z-transform variable
p	Differentiation operator, $pu(t) = \frac{du(t)}{dt}$
q	Shift operator, $qu(t) = u(t + T_s)$
j	Imaginary unit
N	Number of samples
N_P	Number of samples in each period
P	Number of periods
M	Number of experiments (or blocks of experiments for MIMO systems)
N_f	Number of frequencies
Q	Number of positions
Q_c	Number of candidate positions (for experiment design)
n_y	Number of outputs
n_u	Number of inputs

n_e	Number of experiments
T_s	Sampling period
T_0	Period, $T_0 = T_s N_P$
$u(t)$	Input signal at time t
$y(t)$	Output signal at time t
$r(t)$	Reference signal at time t
$v(t)$	Disturbance signal at time t
$x(t)$	State vector at time t
$\dot{x}(t)$	Time derivative of $x(t)$
$F(\cdot)$	Controller
$G(\cdot)$	Transfer function from input to output
$U(\omega_k)$	DFT of input signal $u(t)$
$Y(\omega_k)$	DFT of output signal $y(t)$
$R(\omega_k)$	DFT of reference signal $r(t)$
$V(\omega_k)$	DFT of disturbance signal $v(t)$
ω_k	DFT frequency, see (3.21)
$\mathbf{U}(\omega_k)$	Input DFT matrix of size $n_u \times n_e$, see (3.40), and similarly for $\mathbf{Y}(\omega_k)$, $\mathbf{R}(\omega_k)$, and $\mathbf{V}(\omega_k)$
$\mathbf{U}^{[m]}(\omega_k)$	Block m of size $n_u \times n_u$ of the DFT matrix $\mathbf{U}(\omega_k)$, see (3.43)
$\mathbf{U}^{[m]}_{\text{diag}}(\omega_k)$	Diagonal matrix of multisine signals, $\mathbf{U}^{[m]}(\omega_k) = \mathbf{U}^{[m]}_{\text{diag}}(\omega_k)\mathbf{T}$, see (3.67)
$\mathbf{T}$	Permutation matrix, see (3.67)
$\widehat{G}^{(i)}(\omega_k)$	Nonparametric FRF estimate in position i
$G^{(i)}(\omega_k, \theta)$	FRF of parametric model in position i, see Section 4.4.4
$\widehat{G}^{[m,p]}(\omega_k)$	Nonparametric FRF estimate from block m and period p
$\widehat{G}^{[m]}(\omega_k)$	Nonparametric FRF estimate from block m, see (3.45)
$G_R(\omega_k)$	Best linear approximation, see (3.71)
$G_S(\omega_k)$	Stochastic nonlinear contribution, see (3.71)
$N_G(\omega_k)$	Measurement noise contribution, see (3.71)
q_a	Vector of arm joint coordinates
q_m	Vector of motor joint coordinates
q_e	Vector of elastic coordinates, see (2.11)
τ	Vector of motor torques
τ_{fa}	Arm friction, see (2.5)
τ_{fm}	Motor friction, see (2.9)
F_v, F_c	Viscous and Coulomb friction parameters
$M_a(q_a)$	Inertia matrix in the dynamic equations (2.5)

$c_a(q_a, \dot{q}_a)$	Velocity dependent term in the dynamic equations (2.5)
$g_a(q_a)$	Gravitational term in the dynamic equations (2.5)
J_m, J_a	Moments of inertia of the motor and arm
r_g	Gear ratio
k_g, d_g	Spring stiffness and damping of elastic gear
k_e, d_e	Spring stiffness and damping of elastic arm structure, see (2.11)
θ	Vector of (unknown) parameters
$\hat{\theta}$	Estimated parameters
$\hat{y}(t\|\theta)$	A model's prediction of $y(t)$ given θ and data up to time $t - T_s$
$\varepsilon(t, \theta)$	Prediction error, $y(t) - \hat{y}(t\|\theta)$
$V_N(\theta)$	Prediction error criterion, see (3.32)
$\hat{\theta}_N$	Parameter estimator, see (3.32)
$\mathrm{E}(x)$	Expectation of the random variable x
$\mathrm{Cov}(x)$	Covariance matrix of the random variable x
$\Phi_u(\omega)$	Spectrum for the signal $u(t)$
$\Phi_{yu}(\omega)$	Cross spectrum between $y(t)$ and $u(t)$
P_θ	Covariance matrix, see (3.58b)
$\mathcal{N}(m, P)$	Normal (Gaussian) distribution with mean value m and covariance matrix P
σ_X^2	Covariance matrix, $\sigma_X^2 = \mathrm{Cov}(\mathrm{vec}(X))$
$\sim$	Denotes "is distributed according to"
$\propto$	Proportional to
$\in$	Belongs to
$\forall$	For all
$\arg\min_x f(x)$	The value of x that minimizes $f(x)$
$\mathrm{sign}(x)$	Sign function
$\|z\|$	Absolute value, $\|z\| = \sqrt{x^2 + y^2}$, with $z = x + jy$ and $x, y \in \mathbb{R}$
$\arg(z)$	Argument or phase ϕ of complex variable $z = \|z\|e^{j\phi}$
∂	Partial derivative
$\frac{\partial f}{\partial x}(x_0)$	Jacobian matrix, where element ij is defined as $[\frac{\partial f}{\partial x}(x_0)]_{ij} = \left.\frac{\partial f_i(x)}{\partial x_j}\right\|_{x=x_0}$
$\det A$	Determinant of matrix A
$\mathrm{Tr}\, A$	Trace of matrix A
A^T	Transpose of matrix A
A^H	Complex conjugate transpose of matrix A
A^{-1}	Inverse of matrix A
$\lambda_{max}(A)$	Largest eigenvalue of matrix A

$\mathrm{vec}(A)$	Stack the columns of matrix $A = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix}$ on top of each other as $\begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}^T$
$\mathrm{diag}\,\{a_k\}_{k=1}^n$	$n \times n$ diagonal matrix with elements $a_1, a_2, \ldots, a_n$ on the diagonal
$\mathrm{diag}\,\{A\}$	Will return the diagonal of the square matrix A as a column vector
$\mathrm{diag}\,\{a\}$	A diagonal matrix with the vector a on the diagonal
$\mathcal{O}(\cdot)$	Denotes "order of", e.g., $f(x) = \mathcal{O}(g(x))$ means that there exist x_0 and k such that $\lvert f(x)\rvert \leq k\lvert g(x)\rvert$ for $x > x_0$

Abbreviations and Acronyms

a.s.lim	almost sure limit
e.g.	for example
i.e.	in other words
ARI	Arithmetic mean (estimator)
CAD/CAM	Computer Aided Design/Computer Aided Manufacture
CRLB	Cramér-Rao lower bound
DAE	Differential-Algebraic Equation
DFT	Discrete Fourier Transform
DOF	Degrees Of Freedom
EIV	Errors-in-variables (estimator)
ETFE	Empirical Transfer Function Estimate
FFT	Fast Fourier Transform
FRF	Frequency Response Function
HAR	Harmonic mean (estimator)
IRB	Industrial RoBot, used in names for ABB robots
JIO	Joint input-output (estimator)
LLS	Logarithmic Least Squares
LOG	Logarithmic mean (estimator)
LTI	Linear Time Invariant
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
NLS	Nonlinear Least Squares
ODE	Ordinary Differential Equation
RV	Rotary Vector (gear)
SCARA	Selective Compliant Assembly Robot Arm
SISO	Single Input Single Output
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TCP	Tool Center Point
YALMIP	Yet Another LMI Parser

PhD Dissertations
Division of Automatic Control
Linköping University

M. Millnert: Identification and control of systems subject to abrupt changes. Thesis No. 82, 1982. ISBN 91-7372-542-0.
A. J. M. van Overbeek: On-line structure selection for the identification of multivariable systems. Thesis No. 86, 1982. ISBN 91-7372-586-2.
B. Bengtsson: On some control problems for queues. Thesis No. 87, 1982. ISBN 91-7372-593-5.
S. Ljung: Fast algorithms for integral equations and least squares identification problems. Thesis No. 93, 1983. ISBN 91-7372-641-9.
H. Jonson: A Newton method for solving non-linear optimal control problems with general constraints. Thesis No. 104, 1983. ISBN 91-7372-718-0.
E. Trulsson: Adaptive control based on explicit criterion minimization. Thesis No. 106, 1983. ISBN 91-7372-728-8.
K. Nordström: Uncertainty, robustness and sensitivity reduction in the design of single input control systems. Thesis No. 162, 1987. ISBN 91-7870-170-8.
B. Wahlberg: On the identification and approximation of linear systems. Thesis No. 163, 1987. ISBN 91-7870-175-9.
S. Gunnarsson: Frequency domain aspects of modeling and control in adaptive systems. Thesis No. 194, 1988. ISBN 91-7870-380-8.
A. Isaksson: On system identification in one and two dimensions with signal processing applications. Thesis No. 196, 1988. ISBN 91-7870-383-2.
M. Viberg: Subspace fitting concepts in sensor array processing. Thesis No. 217, 1989. ISBN 91-7870-529-0.
K. Forsman: Constructive commutative algebra in nonlinear control theory. Thesis No. 261, 1991. ISBN 91-7870-827-3.
F. Gustafsson: Estimation of discrete parameters in linear systems. Thesis No. 271, 1992. ISBN 91-7870-876-1.
P. Nagy: Tools for knowledge-based signal processing with applications to system identification. Thesis No. 280, 1992. ISBN 91-7870-962-8.
T. Svensson: Mathematical tools and software for analysis and design of nonlinear control systems. Thesis No. 285, 1992. ISBN 91-7870-989-X.
S. Andersson: On dimension reduction in sensor array signal processing. Thesis No. 290, 1992. ISBN 91-7871-015-4.
H. Hjalmarsson: Aspects on incomplete modeling in system identification. Thesis No. 298, 1993. ISBN 91-7871-070-7.
I. Klein: Automatic synthesis of sequential control schemes. Thesis No. 305, 1993. ISBN 91-7871-090-1.
J.-E. Strömberg: A mode switching modelling philosophy. Thesis No. 353, 1994. ISBN 91-7871-430-3.
K. Wang Chen: Transformation and symbolic calculations in filtering and control. Thesis No. 361, 1994. ISBN 91-7871-467-2.
T. McKelvey: Identification of state-space models from time and frequency data. Thesis No. 380, 1995. ISBN 91-7871-531-8.
J. Sjöberg: Non-linear system identification with neural networks. Thesis No. 381, 1995. ISBN 91-7871-534-2.
R. Germundsson: Symbolic systems – theory, computation and applications. Thesis No. 389, 1995. ISBN 91-7871-578-4.
P. Pucar: Modeling and segmentation using multiple models. Thesis No. 405, 1995. ISBN 91-7871-627-6.
H. Fortell: Algebraic approaches to normal forms and zero dynamics. Thesis No. 407, 1995. ISBN 91-7871-629-2.

A. Helmersson: Methods for robust gain scheduling. Thesis No. 406, 1995. ISBN 91-7871-628-4.
P. Lindskog: Methods, algorithms and tools for system identification based on prior knowledge. Thesis No. 436, 1996. ISBN 91-7871-424-8.
J. Gunnarsson: Symbolic methods and tools for discrete event dynamic systems. Thesis No. 477, 1997. ISBN 91-7871-917-8.
M. Jirstrand: Constructive methods for inequality constraints in control. Thesis No. 527, 1998. ISBN 91-7219-187-2.
U. Forssell: Closed-loop identification: Methods, theory, and applications. Thesis No. 566, 1999. ISBN 91-7219-432-4.
A. Stenman: Model on demand: Algorithms, analysis and applications. Thesis No. 571, 1999. ISBN 91-7219-450-2.
N. Bergman: Recursive Bayesian estimation: Navigation and tracking applications. Thesis No. 579, 1999. ISBN 91-7219-473-1.
K. Edström: Switched bond graphs: Simulation and analysis. Thesis No. 586, 1999. ISBN 91-7219-493-6.
M. Larsson: Behavioral and structural model based approaches to discrete diagnosis. Thesis No. 608, 1999. ISBN 91-7219-615-5.
F. Gunnarsson: Power control in cellular radio systems: Analysis, design and estimation. Thesis No. 623, 2000. ISBN 91-7219-689-0.
V. Einarsson: Model checking methods for mode switching systems. Thesis No. 652, 2000. ISBN 91-7219-836-2.
M. Norrlöf: Iterative learning control: Analysis, design, and experiments. Thesis No. 653, 2000. ISBN 91-7219-837-0.
F. Tjärnström: Variance expressions and model reduction in system identification. Thesis No. 730, 2002. ISBN 91-7373-253-2.
J. Löfberg: Minimax approaches to robust model predictive control. Thesis No. 812, 2003. ISBN 91-7373-622-8.
J. Roll: Local and piecewise affine approaches to system identification. Thesis No. 802, 2003. ISBN 91-7373-608-2.
J. Elbornsson: Analysis, estimation and compensation of mismatch effects in A/D converters. Thesis No. 811, 2003. ISBN 91-7373-621-X.
O. Härkegård: Backstepping and control allocation with applications to flight control. Thesis No. 820, 2003. ISBN 91-7373-647-3.
R. Wallin: Optimization algorithms for system analysis and identification. Thesis No. 919, 2004. ISBN 91-85297-19-4.
D. Lindgren: Projection methods for classification and identification. Thesis No. 915, 2005. ISBN 91-85297-06-2.
R. Karlsson: Particle Filtering for Positioning and Tracking Applications. Thesis No. 924, 2005. ISBN 91-85297-34-8.
J. Jansson: Collision Avoidance Theory with Applications to Automotive Collision Mitigation. Thesis No. 950, 2005. ISBN 91-85299-45-6.
E. Geijer Lundin: Uplink Load in CDMA Cellular Radio Systems. Thesis No. 977, 2005. ISBN 91-85457-49-3.
M. Enqvist: Linear Models of Nonlinear Systems. Thesis No. 985, 2005. ISBN 91-85457-64-7.
T. B. Schön: Estimation of Nonlinear Dynamic Systems — Theory and Applications. Thesis No. 998, 2006. ISBN 91-85497-03-7.
I. Lind: Regressor and Structure Selection — Uses of ANOVA in System Identification. Thesis No. 1012, 2006. ISBN 91-85523-98-4.
J. Gillberg: Frequency Domain Identification of Continuous-Time Systems Reconstruction and Robustness. Thesis No. 1031, 2006. ISBN 91-85523-34-8.
M. Gerdin: Identification and Estimation for Models Described by Differential-Algebraic Equations. Thesis No. 1046, 2006. ISBN 91-85643-87-4.

C. Grönwall: Ground Object Recognition using Laser Radar Data – Geometric Fitting, Performance Analysis, and Applications. Thesis No. 1055, 2006. ISBN 91-85643-53-X.

A. Eidehall: Tracking and threat assessment for automotive collision avoidance. Thesis No. 1066, 2007. ISBN 91-85643-10-6.

F. Eng: Non-Uniform Sampling in Statistical Signal Processing. Thesis No. 1082, 2007. ISBN 978-91-85715-49-7.

Multivariable Frequency-Domain Identification of Industrial Robots

Erik Wernholt, Linköping 2007

Department of Electrical Engineering
Linköping University
SE–581 83 Linköping, Sweden

Linköping University
INSTITUTE OF TECHNOLOGY