Linköping Studies in Science and Technology. Dissertations

No. 566

Closed-loop Identification

Methods, Theory, and Applications

Urban Forssell

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

Linköping 1999

Closed-loop Identification: Methods, Theory, and Applications

ufo@isy.liu.se
http://www.control.isy.liu.se
Department of Electrical Engineering
Linköping University
SE–581 83 Linköping
Sweden

ISBN 91-7219-432-4 ISSN 0345-7524

Printed in Sweden by Linus & Linnea AB

To Susanne, Tea, and Hugo

Abstract

System identification deals with constructing mathematical models of dynamical systems from measured data. Such models have important applications in many technical and nontechnical areas, such as diagnosis, simulation, prediction, and control. The theme in this thesis is to study how the use of closed-loop data for identification of open-loop processes affects different identification methods. The focus is on prediction error methods for closed-loop identification and a main result is that we show that most common methods correspond to different parameterizations of the general prediction error method. This provides a unifying framework for analyzing the statistical properties of the different methods. Here we concentrate on asymptotic variance expressions for the resulting estimates and on explicit characterizations of the bias distribution for the different methods. Furthermore, we present and analyze a new method for closed-loop identification, called the projection method, which allows approximation of the open-loop dynamics in a fixed, user-specified frequency domain norm, even in the case of an unknown, nonlinear regulator.

In prediction error identification it is common to use some gradient-type search algorithm for the parameter estimation. A requirement is then that the predictor filters along with their derivatives are stable for all admissible values of the parameters. The standard output error and Box-Jenkins model structures cannot be used if the underlying system is unstable, since the predictor filters will generically be unstable under these circumstances. In the thesis, modified versions of these model structures are derived that are applicable also to unstable systems. Another way to handle the problems associated with output error identification of unstable systems is to implement the search algorithm using noncausal filtering. Several such approaches are also studied and compared.

Another topic covered in the thesis is the use of periodic excitation signals for time-domain identification of errors-in-variables systems. A number of compensation strategies for the least-squares and total least-squares methods are suggested. The main idea is to use a nonparametric noise model, estimated directly from data, to whiten the noise and to remove the bias in the estimates.

"Identification for Control" deals specifically with the problem of constructing models from data that are good for control. A main idea has been to try to match the identification and control criteria to obtain a control-relevant model fit. The use of closed-loop experiments has been an important tool for achieving this. We study a number of iterative methods for dealing with this problem and show how they can be implemented using the indirect method. Several problems with the iterative schemes are observed and it is argued that performing iterated identification experiments with the current controller in the loop is suboptimal. Related to this is the problem of designing the identification experiment so that the quality of the resulting model is maximized. Here we concentrate on minimizing the variance error and a main result is that we give explicit expressions for the optimal regulator and reference signal spectrum to use in the identification experiment in case both the input and the output variances are constrained.

Preface

System identification is a fascinating field for study and research with many challenging problems, both theoretical and practical ones. This thesis deals with various aspects of system identification using closed-loop data. The main focus is on analysis of the statistical properties of different prediction error methods for closed-loop identification. Other topics covered in the thesis include optimal experiment design, model structure selection, implementation of parameter estimation algorithms, and identification for control.

Some words on the organization of the thesis: This thesis is divided into two separate parts plus an introductory chapter, Chapter 1. The first part, entitled Topics in Identification and Control, is a unified survey of the results in the thesis while the second one is a collection of papers. This part is simply called Publications.

Part I can either be read separately or as an extended introduction to the whole thesis. This part contains some background material, an overview of the most important analysis results in the thesis, plus some material on identification for control that is not covered elsewhere in the thesis. The idea has been to keep the style rather casual in this first, introductory part, with most technical details left out. This choice has been made in order to make the results more easily accessible than, perhaps, in the papers in Part II which contain the detailed statements. An outline of Part I can be found in Chapter 1.

Part II contains edited versions of the following articles and conference papers.

> U. Forssell and L. Ljung. Closed-loop identification revisited. *Automatica*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-2021, Linköping University, Linköping, Sweden

> U. Forssell and L. Ljung. A projection method for closed-loop identification. *IEEE Transactions on Automatic Control*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-1984, Linköping University, Linköping, Sweden

> U. Forssell and C. T. Chou. Efficiency of prediction error and instrumental variable methods for closed-loop identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 1287–1288, Tampa, FL, 1998

> U. Forssell and L. Ljung. Identification of unstable systems using output error and Box-Jenkins model structures. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3932–3937, Tampa, FL, 1998. To appear in IEEE Transactions on Automatic Control

U. Forssell and H. Hjalmarsson. Maximum likelihood estimation of models with unstable dynamics and nonminimum phase noise zeros. In *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999

U. Forssell, F. Gustafsson, and T. McKelvey. Time-domain identification of dynamic errors-in-variables systems using periodic excitation signals. In *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999

U. Forssell and L. Ljung. Identification for control: Some results on optimal experiment design. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3384–3389, Tampa, FL, 1998. Submitted to Automatica

U. Forssell. Asymptotic variance expressions for identified black-box models. Technical Report LiTH-ISY-R-2089, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998. Submitted to Systems & Control Letters

These papers have all been written independently and hence there is some overlap between them. This is especially true for the basic prediction error theory which is covered in several of the papers. On the other hand, since the same notation has been used in all papers, most readers should be able to skim these introductory parts of the papers very quickly or skip them completely. Summaries of the papers can be found in Chapter 1.

Of related interest but not included in the thesis are the following papers:

U. Forssell and P. Lindskog. Combining semi-physical and neural network modeling: An example of its usfulness. In *Proceedings of the 11th IFAC Symposium on System Identification*, volume 4, pages 795–798, Fukuoka, Japan, 1997

L. Ljung and U. Forssell. Variance results for closed-loop identification methods. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 3, pages 2435–2440, San Diego, C.A., 1997

L. Ljung and U. Forssell. Bias, variance and optimal experiment design: Some comments on closed loop identification. In *Proceedings of the Colloqium on Control Problems in Honor of Prof. I.D. Landau*, Paris, June 1998. Springer Verlag

L. Ljung and U. Forssell. An alternative motivation for the indirect approach to closed-loop identification. *IEEE Transactions on Automatic Control*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-1989, Linköping University, Linköping, Sweden

Several persons have contributed to this thesis. First and foremost I would like to mention my supervisor, Professor Lennart Ljung, who introduced me to the subject and who has been a constant source of inspiration throughout this work. I would also like to thank my other co-authors: Drs. Tung Chou, Fredrik Gustafsson, Håkan Hjalmarsson, Peter Lindskog, and Tomas McKelvey for interesting discussions and good cooperation when preparing the manuscripts.

I have also benefitted a lot from working with the other people in the control group here in Linköping, both the faculty and my fellow graduate students. Special thanks goes to Fredrik Gunnarsson, Fredrik Tjärnström, and Jakob Roll who, together with Peter Lindskog, proofread parts of the manuscript before the printing.

I am also grateful to Professor Johan Schoukens who invited me to visit his research group last year and with whom I have had several enlightening discussions on various topics in system identification. Several other researchers in the identification field have also let me borrow some of their time to discuss identification and to ask stupid questions, which I am very grateful for. I especially would like to mention Professors Michel Gevers and Paul Van den Hof and Dr. Pierre Carrette.

Finally, I would like to express my warm gratitude to my wife Susanne who always is supportive and caring and to our children Tea and Hugo who really light up my days (and nights!).

Urban Forssell

Linköping, February 1999

Contents

Notation

Operators

$\arg\min_x f(x)$	Minimizing argument of $f(x)$
$\text{Prob}(x \leq C)$	Probability that the random variable x is less than C
$\text{Cov}\, x$	Covariance matrix of the random vector x
$\text{Var}\, x$	Variance matrix of the random variable x
Ex	Mathematical expectation of the random vector x
$\bar{E}x(t)$	$\lim\limits_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} E\ x(t)$
$\text{Re}\, z$, $\text{Im}\, z$	Real and imaginary parts of the complex number z
$\det A$	Determinant of matrix A
$\dim A$	Dimension of matrix A
$\text{tr}\, A$	Trace of matrix A
$\text{vec}\, A$	Vector consisting of the columns of A
A^T	Transpose of matrix A
A^*	Complex conjugate transpose of matrix A
A^{-1}	Inverse of matrix A
$\Pi_{\mathbf{A}}^{\perp}$	Projection matrix onto the orthogonal complement of $\mathbf{A}$: $\mathbf{A}\Pi_{\mathbf{A}}^{\perp} = 0,\ \Pi_{\mathbf{A}}^{\perp}\Pi_{\mathbf{A}}^{\perp} = \Pi_{\mathbf{A}}^{\perp}$
$\underline{\sigma}(A)$	Smallest singular value of matrix A
$\overline{\sigma}(A)$	Largest singular value of matrix A
μ	Structural singular value
q^{-1}	The delay operator: $q^{-1}x(t) = x(t-1)$

Symbols

This list contains symbols used in the text. Occasionally some symbols will have another meaning; this will then be stated in the text.

$x_N \in AsN(m, P)$	Sequence of random variables x_N that converges in distribution to the normal distribution with mean m and covariance matrix P
$A \subset B$	A is a subset of B
$\mathbb{C}$	The set of complex numbers
$D_{\mathcal{M}}$	Set of values over which θ ranges in a model structure
D_c	Set into which the θ-estimate converges
$e(t)$	Disturbance at time t; usually $\{e(t)\}$ is white noise with zero mean values and variances λ_0
$\varepsilon(t, \theta)$	Prediction error $y(t) - \hat{y}(t\|\theta)$
$G(q)$	Transfer function (matrix) from u to y
$G(q, \theta)$	Transfer function in a model structure, corresponding to the parameter value θ
$\hat{G}_N(q)$	Transfer function estimate: $\hat{G}_N(q) = G(q, \hat{\theta}_N)$
$G_0(q)$	"True" transfer function from u to y for a given system
$\mathcal{G}$	Set of transfer functions obtained in a given structure
$H(q)$	Transfer function from e to y
$H(q, \theta),\ \hat{H}_N(q)$	Analogous to G
$H_0(q)$	Analogous to G
$K(q)$	Linear feedback regulator
$\mathcal{M}$	Model structure
$\mathcal{M}(\theta)$	Particular model corresponding to the parameter value θ
$O(x)$	Ordo x: function tending to zero at the same rate as x
$o(x)$	Small ordo x: function tending to zero faster than x
P_θ	Asymptotic covariance matrix of θ
$\mathbb{R}$	The set of real numbers
$\mathbb{R}^d$	Euclidian d-dimensional space
$R_s(k)$	$\bar{E}s(t)s^T(t-k)$
$R_{sw}(k)$	$\bar{E}s(t)w^T(t-k)$
$S(q)$	Sensitivity function: $S(q) = (1 + G(q)K(q))^{-1}$
$S_0(q)$	Analogous to G
$\mathcal{S}$	The "true" system
$T(q)$	Transfer matrix $\text{vec}\big[G(q)\ H(q)\big]$
$T(q, \theta),\ \hat{T}_N(q)$	Analogous to G
$T_0(q)$	Analogous to G
$u(t)$	Input variable at time t
$\mathbf{U}$	$s \times N$ dimensional Hankel matrix with first row $[u(1), \ldots, u(N)]$
$V_N(\theta),\ V_N(\theta, Z^N)$	Criterion function to be minimized
$\mathbf{V}$	Analogous to $\mathbf{U}$

$y(t)$	Output variable at time t
$\mathbf{Y}$	Analogous to $\mathbf{U}$
$\hat{y}(t\vert\theta)$	Predicted output at time t using a model $\mathcal{M}(\theta)$ and based on Z^{t-1}
$\mathbb{Z}$	The set of integer numbers
Z^N	Data set $\{y(1), u(1), \ldots, y(N), u(N)\}$
θ	Parameter vector, dim $\theta = d$ $(= d \times 1)$
$\hat{\theta}_N$	Estimate of θ obtained using N data points
θ_0	"True" parameter value
$\varphi(t),\ \varphi(t,\theta)$	Regression vector at time t
$\zeta(t),\ \zeta(t,\theta)$	Correlation vector ("instruments") at time t
$\psi(t,\theta)$	Gradient of $\hat{y}(t\vert\theta)$ with respect to θ
$\Phi_s(\omega)$	Spectrum of $\{s(t)\}$ = Fourier transform of $R_s(k)$
$\Phi_{sw}(\omega)$	Cross-spectrum between $\{s(t)\}$ and $\{w(t)\}$ = Fourier transform of $R_{sw}(k)$
$\chi_0(t)$	Signal $\left[u^T(t)\ e^T(t)\right]^T$

Abbreviations and Acronyms

ARX	AutoRegressive with eXogenous input
ARMAX	AutoRegressive Moving Average with eXogenous input
BJ	Box-Jenkins
CLS	Compensated Least-Squares
EIV	Errors-In-Variables
FFT	Fast Fourier Transform
GMVC	Generalized Minimum Variance Control
IFT	Iterative Feedback Tuning
IMC	Internal Model Control
IV	Instrumental Variables
LQG	Linear Quadratic Gaussian
LTI	Linear Time-invariant
MIMO	Multiple Input Multiple Output
MRC	Model Reference Control
OE	Output Error
PEM	Prediction Error Method
SISO	Single Input Single Output
TLS	Total Least-Squares
w.p.	With probability

Chapter 1

Background and Scope

1.1 Introduction

Mathematical models of dynamical systems are of rapidly increasing importance in engineering and today all designs are more or less based on mathematical models. Models are also extensively used in other, nontechnical areas such as biology, ecology, and economy. If the physical laws governing the behavior of the system are known we can use these to construct so called *white-box models* of the system. In a white-box model, all parameters and variables can be interpreted in terms of physical entities and all constants are known a priori. At the other end of the modeling scale we have so called *black-box modeling* or *identification*. Black-box models are constructed from data using no physical insight whatsoever and the model parameters are simply knobs that can be turned to optimize the model fit. Despite the quite simplistic nature of many black-box models, they are frequently very efficient for modeling dynamical systems and require less engineering time to construct than white-box models.

In this thesis we will study methods for black-box identification of linear, time-invariant dynamical systems given discrete-time data. Such methods have a long history going back to Gauss [31] who invented the least-squares method to be able to predict the motions of the planets based on astronomical observations. By now, several methods exist for identifying discrete-time, linear models, for example, the *prediction error methods* (e.g., [71]), the *subspace approaches* (e.g., [114]), and the nonparametric *correlation and spectral analysis methods* (e.g., [13]). We will focus on the prediction error approach and especially on prediction error methods for identifying systems operating under output feedback, that is, in closed loop.

Closed-loop identification has often been suggested as a tool for identification of models that are suitable for control, so called *identification for control*. The main motivation has then been that by performing the identification experiments in closed loop it is possible to match the identification and control criteria so that the model is fit to the data in a control-relevant way. The second theme in the thesis is to study this area and to analyze methods for identification for control.

In the following section we will provide more background material on system identification, model based control, and the interaction between these two disciplines.

1.2 System Identification

System identification deals with the construction of models from data. This is an important subproblem in statistics and, indeed, many identification methods, as well as the tools for analyzing their properties, have their roots in statistics. As a separate field, system identification started to develop in the 1960's, see, for example, the much cited paper [6] by Åström and Bohlin. An important driving force for this development was the increased interest in model based control spurred by Kalman's work on optimal control (e.g., [54, 55]). Since then, the system identification field has experienced a rapid growth and is today a well established research area.

The system identification problem can be divided into a number of subproblems:

- Experiment design.
- Data collection.
- Model structure selection.
- Model estimation.
- Model validation.

Experiment design involves issues like choice of what signals to measure, choice of sampling time, and choice of excitation signals. Once these issues have been settled, the actual identification experiment can be performed and process data be collected. The next problem is to decide on a suitable model structure. This is a crucial step in the identification process and to obtain a good and useful model, this step must be done with care. Given a suitable model structure and measured data, we can turn to the actual estimation of the model parameters. For this, there exist special-purpose software tools that are very efficient and easy to use

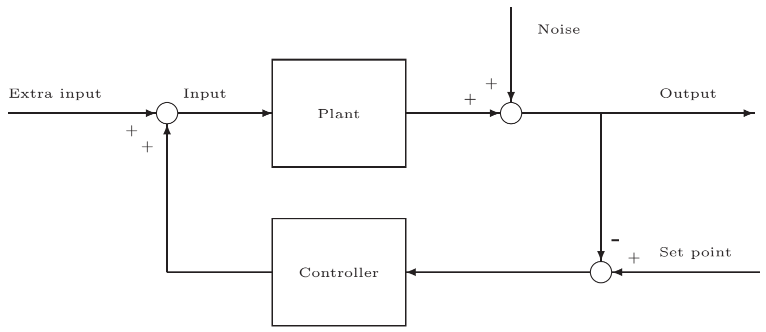

Figure 1.1 *A closed-loop system.*

and consequently this step is perhaps the most straightforward one. Before the model can be delivered to the user it has to pass some validation test. Model validation can loosely be said to deal with the question whether the best model is also "good enough" for its intended use. Common validation tools are *residual analysis* and, so called, *cross-validation*, where the model is simulated using "fresh" data and the output compared to the measured output. If the first model fails to pass the validation tests, some, or all, of the above steps have to be iterated until a model that passes the validations tests is found.

This is a brief outline of the different steps in the identification procedure. More material can be found in the text books [71, 76, 107]. Let us now discuss closed-loop identification.

Closed-loop identification results when the identification experiment is performed in closed loop, that is, with the output being fed back to the input by means of some feedback mechanism. A typical situation is when the system is being controlled by some feedback controller, as in Figure 1.1. Automatic controllers are frequently used for changing the behavior of dynamical systems. The objective may, for instance, be to increase the speed of response of the system or to make it less sensitive to disturbances, noise. In such cases it is possible to choose to perform the identification experiment in open loop, with the controller disconnected, or in closed loop. However, if the system is unstable or has to be controlled for economic or safety reasons, closed-loop experiments are unavoidable. This is also true if the feedback is inherent in the system and thus cannot be affected by us. Closed-loop experiments are also optimal in certain situations:

- With variance constraints on the output.

- For identification for control.

A problem with closed-loop data is that many of the common identification methods, that work well in open loop, fail when applied directly to measured input-output data. This is true for instrumental variables, spectral analysis, and many subspace methods, for instance. The reason why these methods fail is the nonzero correlation between the input and the unmeasurable output noise (cf. Figure 1.1) that is inevitable in a closed-loop situation. Unlike the other methods, the prediction error method can be applied directly to closed-loop data and is probably the best choice among the methods that work in closed loop.

Another problem with closed-loop data is that identifiability may be lost. This means that it becomes impossible to uniquely determine the system parameters from measured input-output data. This problem has attracted a lot of interest in the literature, see, for example, the survey papers [3, 40].

These issues will be further discussed below. Let us now give some background material on model based control.

1.3 Model Based Control

Most modern methods for synthesis of linear controllers require some kind of model of the system to be controlled. The models may be used explicity in the design schemes or as tools for assessing the performance of the controller before applying it to the real plant, so called *controller validation*. To understand some of the questions that we will deal with in connection to identification for control, it is necessary to have some basic knowledge of the history of the control field. Therefore we here give a brief overview of the evolution of the (linear) control field.

The period 1930-1960 is often called the *classical control* period. Researchers like Bode and Nyquist developed tools to analyze and design control systems in terms of stability, robustness, and performance for single input single output systems. The chief paradigm during this period was the frequency domain and many of the design methods of this period are graphical, based on, for instance, Bode and Nyquist plots, Nichols charts, and root locus plots. The controllers that were built were PI and PID controllers. Gain and phase margins were used as tools to analyze the robustness of the designs.

With Kalman's seminal work on state space models in the late 1950's and early 1960's a new era in control theory begun. This is frequently referred to as the *optimal control* period. Now the focus shifted from frequency domain to time domain methods. Important concepts like optimal state feedback, optimal state estimation, controllability and observability were introduced. The control synthesis methods were mainly LQG and model reference control. The graphical techniques were thus replaced by *certainty equivalence* control design.

The problem with the optimal control methods is that model uncertainties are not easily incorporated in the designs. This was one of the motivations for the introduction of H_∞ methods in the beginning of the 1980's. The paper by Zames [120] marks the starting point of the *modern control* or *robust control* era. In robust control, model uncertainties specified in the frequency domain are again brought back into center-stage, just like in classical control. Concepts like robust stability and robust performance are central in modern robust control methods like the H_∞-loop shaping technique of Glover and McFarlane [79, 80] and μ-synthesis [124].

Parallel to these methods, the *adaptive control* methods have been developed. Adaptive control is another example of a model based control paradigm and, as we shall see, there is a close connection between adaptive control and the iterative approaches to identification for control that have been suggested in the literature. However, we will not deal directly with adaptive control in this thesis. The interested reader may consult the text book [8] for an introduction to and an overview of the adaptive control field.

1.4 The Interplay between Identification and Control

Since most control methods are model based and system identification deals with the construction of models from measured data, one may think that there should exist close connections between the fields. Initially, during the 1960's, this was indeed the case, but since then the two fields have developed quite independently.

In identification the emphasis has long been on methods that are consistent and statistically efficient given that the underlying system is linear, time-invariant, and of finite order and that the disturbances can be modeled as realizations of stochastic processes. This ties in nicely with the model demands in many optimal control methods, such as LQG and minimum variance control. However, in robust control it is assumed that both a nominal model and "hard", deterministic bounds on the model uncertainty are given a priori. The classical identification methods can deliver good nominal models but in the stochastic framework we cannot give hard bounds on the model uncertainty – we can only form confidence intervals for the models, so called "soft", probabilistic uncertainty bounds. "Hard" and "soft" uncertainty bounding are two completely different paradigms and, perhaps not surprisingly, there has been a fierce debate on the relative merits of these approaches. Critical and interesting discussions on this can be found in, for example, [47, 85, 119]. See also the survey paper [86]. In this thesis we will assume a stochastic framework and focus on methods for probabilistic uncertainty bounding.

Probabilistic uncertainty bounds, or confidence intervals, can be constructed from the covariance information delivered by the identification routine. A problem with this information is that it is only reliable when the bias due to undermodeling is negligible. With undermodeling, special measures have to be taken to obtain reliable covariance estimates [47, 51]. This is mainly a problem if we have bounds on the model order. If we can tolerate high-order models it should always be possible to get an unbiased estimate so that the covariance information is reliable. The problem is then that the variance will increase with the number of estimated parameters. This is known as the *bias/variance trade-off* (e.g., [71]). The limiting factor is the number of data points. With very long data records the importance of the bias/variance trade-off is diminished and we can quite safely use high-order models. With short data records, however, this is an important issue [86] and a tool for quantifying the size of the undermodeling would be very useful. Stochastic embedding [36, 37] is an approach for estimating the uncertainty in restricted complexity models that has been suggested for this purpose. Another example can be found in [44], see also [42], where a method for probabilistic uncertainty bounding is presented in which the bias and variance errors are explicitly evaluated. With this method we can therefore manually resolve the bias/variance trade-off problem, which can be advantageous. However, the method described in [44] is much more complicated than the straightforward approach of estimating a high-order ARX model, say, with small bias and computing the confidence regions for this model using the available covariance estimates. This will be further discussed later in the thesis.

"Model error modeling" [69, 70] can be seen as a way to present the information contained in the standard residual tests for model validation in a control-oriented fashion. The model validation/unfalsification approach but in a purely deterministic setting has also been adopted in for example [60, 91, 103, 104]. The idea is to trade-off the possibilities that the model uncertainties are due to unmodeled dynamics or due to disturbances. The problem formulation is closely related to robust control theory. Other deterministic approaches to model error bounding include the so called *worst-case identification* methods (e.g., [78, 83, 119]) and the *set membership identification* methods (e.g., [81, 82, 88]). A problem with deterministic error bounding techniques is that the bounds can be conservative. This is illustrated in [42] where both deterministic and probabilistic uncertainty bounding methods are studied.

The worst-case and set membership identification methods generally give some nominal model as part of the model set description. These methods could therefore in principle be used for the whole process of constructing a suitable model and computing uncertainty bounds. The main benefit would then be that the model can be fit to the data in a control-relevant norm, like the H_∞-norm. A problem with the worst-case and set membership identification methods, though, is that the nominal models tend to be unsuitable for high-performance control design. A reason for this is that they are chosen to give the smallest upper bounds on the uncertainty rather than to be the best possible approximations of the true system.

A continued discussion on identification for robust control will be given later in the thesis. Let us return to the question of how to identify good nominal models.

What is a good model? In the linear case, the answer is that the model fit has to be good in certain frequency ranges, typically around the cross-over frequency. This of course a quite vague statement but it is hard to give a more quantitative answer that is universally applicable. In general one can say that desired control performance dictates the required quality of the model fit: The higher the demands on control performance are, the better model fit is required.

If the important frequency ranges are known, or can be identified, we can use prefiltering and other techniques to obtain a good fit in these ranges. However, in some cases the important frequency ranges are not known and then the situation is less clear. A similar situation occurs when the achievable bandwidth is not known a priori but part of the information we learn from the identification experiments. A possible solution would then be to use repeated experiments, probing higher and higher frequencies, to gain more and more information about the system that is to be controlled; the "windsurfer" approach [4, 65]. Variants of this are used in many of the identification-for-control-schemes that have been suggested in the literature. A main feature of these schemes is that closed-loop data, generated with the current controller in the loop, is used for updating the model which in turn is used to update the controller, and so on. This is the same paradigm as adaptive control (e.g., [8]). The only difference is that the model is updated using a batch of data instead of at each iteration. The survey papers [32, 33, 112] give nice overviews of such methods. Unfortunately, this approach may lead to increasingly poor control performance which is unacceptable. Later in the thesis we will discuss this in some detail and also point at alternative directions for identification for control.

1.5 Outline of Part I

Part I consists of Chapters 2-6. In Chapter 2 we explain the basics of some alternative identification methods, including the instrumental variable method, the subspace approaches, and spectral analysis, and explain why these fail when applied directly to closed-loop data. This chapter also discusses different approaches for avoiding the problems associated with closed-loop identification.

Chapter 3 deals with the prediction error method. We describe the algorithm format and discuss different model structures and how to compute the estimates. In addition, we briefly state some of the most important analysis results for the statistical properties of this method. Two alternative prediction error methods for closed-loop identification are also presented and analyzed: the indirect method and the projection method.

Chapter 4 discusses the advantages of using periodic excitation signals in time-domain identification. A number of simple and efficient methods that can handle the general errors-in-variables situation are presented.

In Chapter 5 we return to the question of how to identify models that are suitable for control. A number of iterative approaches to identification for control will be studied. Identification for control is also the topic in Chapter 6. Here we present some results on optimal experiment design and suggest a novel approach to control-relevant identification.

1.6 Outline of Part II

As mentioned, the second part of the thesis consists of eight different papers which will be summarized in this section. The paper outlines below clearly indicate the scope of the thesis and also gives an idea of the contributions. A summary of the main contributions will be given in the next section.

Paper A: Closed-loop Identification Revisited. In this paper a comprehensive study of closed-loop identification in the prediction error framework is presented. Most of the existing methods are placed in a common framework by treating them as different parameterizations of *the* prediction error method. This facilitates a simultaneous analysis of the different methods using the basic statements on the asymptotic statistical properties of the general prediction error method: convergence and bias distribution of the limit transfer function estimates; asymptotic variance of the transfer function estimates (as the model orders tend to infinity); asymptotic variance and distribution of the parameter estimates.

Paper B: A Projection Method for Closed-loop Identification. This paper presents a novel method for closed-loop identification that allows fitting the model to the data with arbitrary frequency weighting. The method is called the projection method. The projection method is in form similar to the two-stage method [111], but the underlying ideas and the properties are different. The projection method is applicable to systems with arbitrary feedback mechanisms, just as the direct method. A drawback with the projection method is the suboptimal accuracy.

Paper C: Efficiency of Prediction Error and Instrumental Variable Methods for Closed-loop Identification The asymptotic variance properties of different prediction error and instrumental variable methods for closed-loop identification are compared. To a certain extent, the results can also be extrapolated to the subspace methods. One of the key observations is that the suboptimal accuracy of the indirect and a number of other methods is due to the fact that the whole input spectrum is not utilized in reducing the variance, as with the direct method.

Paper D: Identification of Unstable Systems Using Output Error and Box-Jenkins Model Structures. The prediction error method requires the predictors to be stable. If the underlying system is unstable and an output error or a Box-Jenkins model structure is used, this condition will not be satisfied. In this paper it is shown how to modify these model structures to make them applicable also to unstable systems.

Paper E: Maximum Likelihood Estimation of Models with Unstable Dynamics and Nonminimum Phase Noise Zeros. This paper can be seen as a continuation of Paper D and discusses how to implement maximum likelihood estimation of unstable systems or systems with nonminimum phase noise zeros. It is shown that, by using noncausal filtering, maximum likelihood estimation of such systems is possible.

Paper F: Time-domain Identification of Dynamic Errors-in-variables Systems Using Periodic Excitation Signals. The use of periodic excitation signals in time-domain identification is advocated. It is shown that it is possible to construct simple and noniterative methods for errors-in-variables estimation if periodic excitation signals are used. The idea is to include a bias correction step in the standard least-squares and total least-squares methods. Several such methods are suggested and compared.

Paper G: Identification for Control: Some Results on Optimal Experiment Design. In this paper it is shown that the optimal design for identification experiments with constrained output variance is to perform the experiment in closed loop with a certain LQ regulator controlling the plant. Explicit formulas for the optimal controller and the optimal reference signal spectrum are given. This result has important applications in identification for control, for instance. In the paper various other optimal experiment design problem formulations are also studied.

Paper H: Asymptotic Variance Expressions for Identified Black-box Models. From classical results on prediction error identification we know that it is possible to derive expressions for the asymptotic variance of identified transfer function models that are asymptotic both in the number of data and in the model order. These expressions tend to be simpler than the corresponding ones that hold for fixed model orders and yet they approximate the true covariance well in many cases. In Paper H the corresponding (doubly) asymptotic variance expressions for the real and imaginary parts of the identified transfer function model are computed. As illustrated in the paper, these results can be used to compute uncertainty regions for the frequency response of the identified model.

1.7 Main Contributions

The main contributions of this thesis are:

- The analysis of the statistical properties of the closed-loop identification methods presented in Paper A.
- The projection method introduced in Paper B.
- The modified versions of the common output error and Box-Jenkins model structures that are applicable also if the underlying system is unstable, which were presented in Paper D.
- The results on optimal experiment design in Paper G dealing with the situation that the output variance is constrained.

Part I

Topics in Identification and Control

Chapter 2

Problems with Closed-loop Experiments

As mentioned in Chapter 1, it is sometimes advantageous or even necessary to perform the identification experiment with the system operating in closed loop. Here we will begin our study of the problems associated with this and, among other things, we will explain why some well known identification methods cannot be used for closed-loop identification. The chapter starts with an introductory section that contains some technical assumptions on the data set and introduces the notation that will be used in the thesis.

2.1 Technical Assumptions and Notation

We will consider the following set-up. The system is given by

$$y(t) = G_0(q)u(t) + v(t), \quad v(t) = H_0(q)e(t) \tag{2.1}$$

Here $y(t)$ is the output, $u(t)$ the input, and $e(t)$ a white noise signal with zero mean and variance λ_0. The symbol q denotes the discrete-time shift operator:

$$q^{-1}u(t) = u(t-1) \tag{2.2}$$

Without loss of generality, we assume that $G_0(q)$ contains a delay. The noise model $H_0(q)$ is assumed monic and inversely stable. We further assume that the input is generated as

$$u(t) = k(t, y^t, u^{t-1}, r(t)) \tag{2.3}$$

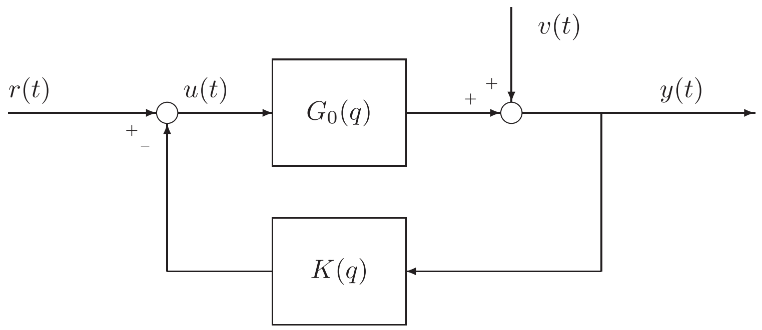

Figure 2.1 *System controlled by a linear regulator.*

where $y^t = [y(1), \ldots, y(t)]$, etc., and where the reference signal $r(t)$ is a given quasi-stationary signal, independent of $v(t)$ and k is a given deterministic function such that the closed-loop system is well posed and exponentially stable in the sense defined in [67]. See also Paper A.

For some of the analytic treatment we will assume that the feedback is linear and given by

$$u(t) = r(t) - K(q)y(t) \tag{2.4}$$

This set-up is depicted in Figure 2.1. With linear feedback the condition on exponential stability is equivalent to internal stability of the closed-loop system (2.1), (2.4).

By combining the equations (2.1) and (2.4) we get the closed-loop relations

$$y(t) = S_0(q)G_0(q)r(t) + S_0(q)v(t) \tag{2.5}$$

$$u(t) = S_0(q)r(t) - K(q)S_0(q)v(t) \tag{2.6}$$

where $S_0(q)$ is the sensitivity function

$$S_0(q) = \frac{1}{1 + G_0(q)K(q)} \tag{2.7}$$

For future use we also introduce

$$G_0^c(q) = S_0(q)G_0(q) \tag{2.8}$$

and

$$H_0^c(q) = S_0(q)H_0(q) \tag{2.9}$$

so that we can rewrite the closed-loop system (2.5) as

$$y(t) = G_0^c(q)r(t) + v_c(t), \quad v_c(t) = H_0^c(q)e(t) \tag{2.10}$$

We will frequently use the "total expectation" operator $\bar{E}$ (cf. [71]):

$$\bar{E}f(t) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} Ef(t) \tag{2.11}$$

The (power) spectrum of a signal $s(t)$ will be denoted by $\Phi_s(\omega)$ and the cross-spectrum between two signals $s(t)$ and $w(t)$ will be denoted by $\Phi_{sw}(\omega)$.

2.2 Examples of Methods that May Fail in Closed Loop

2.2.1 Instrumental Variable Methods

Consider the linear regression model

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \tag{2.12}$$

where $\varphi(t)$ is a regression vector and θ a parameter vector. With the standard instrumental variable method (e.g., [106]) the parameter estimate is found as

$$\hat{\theta}_N = \left[\frac{1}{N}\sum_{t=1}^{N}\zeta(t)\varphi^T(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\zeta(t)y(t) \tag{2.13}$$

Here $\zeta(t)$ is an instrumental variable vector whose elements are the instruments. (We can note that the celebrated least-squares method is obtained if $\zeta(t) = \varphi(t)$.)

Suppose that the true system is

$$y(t) = \varphi^T(t)\theta_0 + e(t) \tag{2.14}$$

where $e(t)$ is a zero mean, white noise signal. Under mild conditions on the data set we then have that

$$\hat{\theta}_N \to \theta_0 + \left[\bar{E}\zeta(t)\varphi^T(t)\right]^{-1}\bar{E}\zeta(t)e(t) \text{ w.p. 1 as } N\to\infty \tag{2.15}$$

See, for example, [106]. (The abbreviation w.p. stands for "with probability".) It follows that the instrumental variable method is consistent if the following two conditions hold:

$$\bar{E}\zeta(t)e(t) = 0 \tag{2.16}$$

$$\bar{E}\zeta(t)\varphi^T(t) \text{ has full rank} \tag{2.17}$$

In loose terms, the instruments should thus be well correlated with lagged inputs and outputs but uncorrelated with the noise sequence. In open loop it is common to let the instruments be filtered and delayed versions of the input. However, this straightforward approach will fail in closed loop since the input is correlated with the noise in this case. There are basically two ways to circumvent this problem: Either we construct the instruments from the reference signal which is uncorrelated with the noise (see, e.g., [108]) or we let them be delayed versions of the regressors, where the delay is chosen large enough so that (2.16) holds (at least approximately). With both these solutions we have to be careful to satisfy the rank condition (2.17) which imposes further constraints on the instruments. Furthermore, the choice of instruments will also determine the statistical efficiency of the method as the accuracy depends on the inverse of the matrix $\bar{E}\zeta(t)\varphi^T(t)$, which hence must be well conditioned. See Paper C.

Thus it is not completely straightforward to apply the instrumental variable method with closed-loop data. The reason is basically that it is hard to find instruments that simultaneously satisfy (2.16) and (2.17) in closed loop. A similar situation occurs with the subspace methods, as will be shown next.

2.2.2 Subspace Methods

A discrete-time, linear, and time-invariant system can always be written as

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) + v(t) \end{aligned} \tag{2.18}$$

where $v(t)$ is a lumped noise term. For simplicity we will here only discuss scalar systems but the discussion below also holds for the multivariable case with some minor notational changes.

Subspace methods can be used to estimate the order n of the system, that is, the dimension of the state vector $x(t)$, and the system matrices (A, B, C, D) up to a similarity transformation.

If we organize $\{y(t)\}$ into a Hankel matrix as follows:

$$\mathbf{Y} = \begin{bmatrix} y(1) & y(2) & \cdots & y(N) \\ y(2) & y(3) & \cdots & y(N+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(s) & y(s+1) & \cdots & y(N+s-1) \end{bmatrix} \tag{2.19}$$

where $s \geq n$ and store $\{u(t)\}$ and the noise sequence $\{v(t)\}$ in exactly the same format in the block Hankel matrices $\mathbf{U}$ and $\mathbf{V}$, respectively, we can write

$$\mathbf{Y} = \Gamma X + H\mathbf{U} + \mathbf{V} \tag{2.20}$$

with

$$\Gamma = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{s-1} \end{bmatrix}, \quad H = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ CA^{s-2}B & CA^{s-3}B & \cdots & D \end{bmatrix} \tag{2.21}$$

$$X = \begin{bmatrix} x(1) & \ldots & x(N) \end{bmatrix} \tag{2.22}$$

The matrix Γ is often called the extended observability matrix. An important observation is that the term ΓX in (2.20) will be rank deficient (rank $\Gamma X = n < s$) while the others will have full rank. If we somehow could determine n and consistently estimate the column space of Γ – this is the subspace step – we can get consistent estimates of the matrices (A, C). Given these it is then possible to estimate (B, D) by solving a standard least-squares problem (see, e.g., [71]).

If the $\mathbf{U}$- and $\mathbf{V}$-terms were zero, the column space of Γ could be found from the column space of $\mathbf{Y}$. To use this idea also with nonzero $\mathbf{U}$- and $\mathbf{V}$-terms we first need to eliminate the effects of these. Define the projection matrix

$$\Pi_{\mathbf{U}}^{\perp} = I - \mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1}\mathbf{U} \tag{2.23}$$

If we multiply (2.20) from the right by this matrix we get

$$\mathbf{Y}\Pi_{\mathbf{U}}^{\perp} = \Gamma X \Pi_{\mathbf{U}}^{\perp} + \mathbf{V}\Pi_{\mathbf{U}}^{\perp} \tag{2.24}$$

The next problem is to remove the last term. Since this term is made up of noise contributions, the idea is to correlate it away with a suitable matrix $\mathbf{W}$ of dimension $s \times N$. Multiplying (2.24) from the right by $\mathbf{W}^T$ and normalizing by N gives

$$\tilde{Y}_N = \frac{1}{N}\mathbf{Y}\Pi_{\mathbf{U}}^{\perp}\mathbf{W}^T = \Gamma\frac{1}{N}X\Pi_{\mathbf{U}}^{\perp}\mathbf{W}^T + \frac{1}{N}\mathbf{V}\Pi_{\mathbf{U}}^{\perp}\mathbf{W}^T \triangleq \Gamma\tilde{X}_N + \tilde{V}_N \tag{2.25}$$

Now, if the matrix $\mathbf{W}$ is such that

$$\lim_{N\to\infty} \tilde{V}_N = 0 \tag{2.26}$$

$$\lim_{N\to\infty} \tilde{X}_N \triangleq \tilde{X} \text{ has full rank } n \tag{2.27}$$

we will have that

$$\tilde{Y}_N = \Gamma \tilde{X} + E_N \tag{2.28}$$

$$E_N = \Gamma(\tilde{X}_N - \tilde{X}) + \tilde{V}_N \to 0 \text{ as } N \to \infty \tag{2.29}$$

From this it follows that we can get a consistent estimate of the column space of Γ by computing a singular value decomposition of $\tilde{Y}_N$ and extracting the left singular vectors corresponding to the n largest singular values of $\tilde{Y}_N$.

This is a short description of the subspace approach to system identification. More details can be found in, for example, the text book [114] and the papers [61, 117, 118].

The key to good results is to choose the matrix $\mathbf{W}$ in such a way that it is uncorrelated with the noise (condition (2.26)) and results in a full rank matrix $\tilde{X}$ (condition (2.27)). (Notice the similarity with the conditions (2.16) and (2.17) for the instrumental variable method!) In open loop we can, for instance, build up $\mathbf{W}$ using delayed inputs but, just as with the instrumental variable method, this approach will fail in closed loop due to the correlation between the input and the noise. A feasible approach would be to build up $\mathbf{W}$ using delayed versions of the reference signal which would automatically ensure (2.26). In practice, it is also important that the rank condition (2.27) is satisfied with some margin. Otherwise the accuracy can be very poor, as is discussed in Paper C.

Other approaches that have been suggested in the literature include the methods by [115, 116]. The idea proposed in [116] is to identify an augmented system with the input and the output as outputs and the reference signal as input. The system and controller parameters are then retrieved in a second, algebraic step. In [115] it is assumed that a number of Markov parameters for the controller are known a priori. As shown in the paper, it is then possible to consistently estimate the open-loop system from closed-loop data using subspace techniques.

Here we should also mention that there are alternative forms of the subspace methods that are directly applicable to closed-loop data, see, for example, [14, 15, 90]. Some of these methods are yet not fully developed, but the ideas show promise.

2.2.3 Correlation and Spectral Analysis Methods

Correlation and spectral analysis (e.g., [13, 71]) are examples of nonparametric identification methods. Correlation analysis uses time-domain measurements while spectral analysis is a frequency-domain method (or can at least be interpreted as such).

Consider the system description

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k) + v(t) \tag{2.30}$$

where $u(t)$ is quasi-stationary [71] and $u(t)$ and $v(t)$ are uncorrelated. Given finite sample estimates $\hat{R}_u^N(k)$ and $\hat{R}_{yu}^N(k)$ of the correlation functions

$$R_u(k) = \bar{E}u(t)u(t-k) \tag{2.31}$$

$$R_{yu}(k) = \bar{E}y(t)u(t-k) \tag{2.32}$$

respectively, we can estimate the impulse response coefficients $g(k)$, $k = 1, \dots, M$ by solving

$$\hat{R}_{yu}^N(\tau) = \sum_{k=1}^{M} g(k)\hat{R}_u^N(k-\tau) \tag{2.33}$$

This is the idea behind the correlation approach. This clearly has close connections to the least-squares method, but less emphasis is put on the model structure (regressors) used. Often some kind of prefiltering operation is used to simplify the solution of (2.33), for example prewhitening of the input signal.

In spectral analysis a similar idea is used: Given spectral estimates $\hat{\Phi}_u^N(\omega)$ and $\hat{\Phi}_{yu}^N(\omega)$ of $\Phi_u(\omega)$ and $\Phi_{yu}(\omega)$, respectively, we may estimate the true transfer function (cf. (2.30))

$$G_0(e^{i\omega}) = \sum_{k=-\infty}^{\infty} g(k)e^{-i\omega k} \tag{2.34}$$

by

$$\hat{G}_N(e^{i\omega}) = \frac{\hat{\Phi}_{yu}^N(\omega)}{\hat{\Phi}_u^N(\omega)} \tag{2.35}$$

This assumes that the cross-spectrum $\Phi_{uv}(\omega)$ is zero, that is, that $u(t)$ and $v(t)$ are uncorrelated. Spectral analysis can thus be seen as a Fourier domain counterpart of correlation analysis.

A classical result for spectral analysis is that under (2.1) and (2.4) the standard spectral analysis estimate (2.35) will converge to

$$G^*(e^{i\omega}) = \frac{G_0(e^{i\omega})\Phi_r(\omega) - K(e^{-i\omega})\Phi_v(\omega)}{\Phi_r(\omega) + |K(e^{i\omega})^2|\Phi_v(\omega)} \tag{2.36}$$

as the number of data tends to infinity. This was shown in [1]. Obviously the limit estimate $G^*(e^{i\omega}) \neq G_0(e^{i\omega})$, unless the data is noise free ($\Phi_v(\omega) = 0$). Hence spectral analysis in its standard form also fails in closed loop, just as the standard instrumental variable and subspace methods discussed above. Also note that with no external reference signal present (i.e., $\Phi_r(\omega) = 0$), the limit estimate will equal the negative inverse of the controller.

A closed-loop solution for spectral analysis was given in [1]. The idea is as follows: Compute the spectral estimates

$$\hat{G}_N^{yx}(e^{i\omega}) = \frac{\hat{\Phi}_{yx}^N(\omega)}{\hat{\Phi}_x^N(\omega)} \quad \text{and} \quad \hat{G}_N^{ux}(e^{i\omega}) = \frac{\hat{\Phi}_{ux}^N(\omega)}{\hat{\Phi}_x^N(\omega)} \tag{2.37}$$

where the signal $x(t)$ is correlated with $y(t)$ and $u(t)$, but uncorrelated with the noise $e(t)$ (a standard choice is $x(t) = r(t)$). The open-loop system may then be consistently estimated as

$$\hat{G}_N^{yu}(e^{i\omega}) = \frac{\hat{G}_N^{yx}(e^{i\omega})}{\hat{G}_N^{ux}(e^{i\omega})} = \frac{\hat{\Phi}_{yx}^N(\omega)}{\hat{\Phi}_{ux}^N(\omega)} \tag{2.38}$$

2.3 Approaches to Closed-loop Identification

In the preceding section we gave some examples of methods that cannot be applied directly to closed-loop data. The prediction error method, on the other hand, works fine as long as the parameterization is flexible enough. This method will be thoroughly studied in the next chapter. In the above section we also mentioned some alternative ways to implement the methods to make them applicable also in the closed-loop situation. In fact it turns out that a large number of such "methods"/parameterizations exist, in particular a number of prediction error "methods" for closed-loop identification. It is convenient to consider the following classification of the different types of methods, which basically is due to [40] (see also [107] and Paper A): Depending on what assumptions are made on the nature of the feedback, all closed-loop identification methods can be classified as *direct*, *indirect*, or *joint input-output* methods.

In the direct approach, the method is applied directly to measured input-output data and no assumptions whatsoever are made on how the data was generated. The indirect methods assume perfect knowledge of the feedback controller used in the identification experiment and the idea is to identify the closed-loop system and to compute the open-loop parameters from this estimate, using the knowledge of the controller. The third approach, the joint input-output approach, amounts to modeling the input and the output jointly as outputs from an augmented system driven by the reference signal and the unmeasurable noise. Given an estimate of this augmented system, the open-loop model (and the controller) can be solved for.

In the joint input-output approach it is thus not required to know the controller, but the controller must be known to have a certain structure.

The indirect and joint input-output approaches are typically only used when the feedback law is linear, as in (2.4), but can also be applied with nonlinear feedback. The price is of course that the estimation problems then become much more involved.

It should be noted that the direct approach is only applicable with the prediction error method and some of the subspace methods. The reason for this is the unavoidable correlation between the input and the noise, which, as we have seen, rules out most other methods. With the indirect and joint input-output approaches the closed-loop problem is converted into an open-loop one since the reference signal (which plays the role of an input) is uncorrelated with the noise. Hence these approaches can be used together with all open-loop methods, for example spectral analysis and the instrumental variable method.

In the next chapter we will study different prediction error methods that fall into these categories. Of the methods discussed in the previous section, the method described in [115] falls into the indirect category while the methods in [1, 108, 116] should be regarded as joint input-output methods.

2.4 Generating Informative Data in Closed Loop

Above we saw examples of how to identify systems operating in closed loop. An underlying assumption is that the data is sufficiently rich so that it is indeed possible to uniquely determine the system parameters. This is a much studied subject in the (closed-loop) identification literature. See, among many references, [3, 40, 71, 105].

In open loop the situation is rather transparent and it can be shown that the data set is *informative enough* with respect to the chosen model structure if, and only if, the input is *persistently exciting* of sufficiently high order. See [71] for more details. In closed loop the situation is more complicated. The following standard example shows what can happen.

Example 1 *Consider the system*

$$y(t) + ay(t-1) = bu(t-1) + e(t) \tag{2.39}$$

where $e(t)$ is a white noise signal. Suppose that a proportional regulator

$$u(t) = -ky(t) \tag{2.40}$$

is used during the identification experiment. Inserting the feedback law into the system equation gives

$$y(t) + (a + bk)y(t - 1) = e(t) \tag{2.41}$$

This shows that all models $(\hat{a}, \hat{b})$ *of the form*

$$\hat{a} = a - \gamma k, \qquad \hat{b} = b + \gamma \tag{2.42}$$

give the same input-output description of the system (2.39) *under the feedback* (2.40) *and thus there is no way to distinguish between all these models based on the measured input-output data.*

As this example shows, it is not sufficient in closed loop that the input is persistently exciting to be able to guarantee informative data. The reason is basically that the feedback law used in the example is too simple. With a time-varying, nonlinear, or noisy regulator, or if we switch between several different linear, time-invariant regulators, the data will, in general, be informative (enough). In the next chapter we will return to this question and derive an explicit condition which must be satisfied for the data to be informative.

2.5 Summary

In this chapter we have studied the closed-loop identification problem and we have given examples of methods that fail when applied directly to closed-loop data. We have also studied different ways to avoid the problems associated with closed-loop identification which lead to a characterization of the possible approaches into three different categories: the direct, the indirect, and the joint input-output approaches. With the direct approach one should apply the identification method directly to input-output data, ignoring possible feedback. This is the simplest approach, but cannot be used with all methods. An exception is the prediction error method which will be studied in the next chapter. The other two approaches aim at identifying some closed-loop transfer function from which the open-loop estimate can be determined. These approaches can be applied with all open-loop identification methods.

Chapter 3

Closed-loop Identification in the Prediction Error Framework

3.1 Further Notation

Consider the *model structure* $\mathcal{M}$ (cf. (2.1)):

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \tag{3.1}$$

$G(q, \theta)$ will be called the *dynamics model* and $H(q, \theta)$ the *noise model*. $H(q, \theta)$ is assumed monic. The parameter vector θ ranges over a set $D_{\mathcal{M}} \subset \mathbb{R}^d$ ($d = \dim \theta$) which is assumed compact and connected.

The model structure (3.1), where $\theta \in D_{\mathcal{M}}$, describes a *model set*. We say that the true system (2.1) is contained in the model set if, for some $\theta_0 \in D_{\mathcal{M}}$,

$$G(q, \theta_0) = G_0(q), \;\; H(q, \theta_0) = H_0(q) \tag{3.2}$$

This will also be written $\mathcal{S} \in \mathcal{M}$. The case when the true noise properties cannot be correctly described within the model set, but where there exists a $\theta_0 \in D_{\mathcal{M}}$ such that

$$G(q, \theta_0) = G_0(q) \tag{3.3}$$

will be denoted $G_0 \in \mathcal{G}$.

We will use the notation

$$T(q,\theta) = \begin{bmatrix} G(q,\theta) \\ H(q,\theta) \end{bmatrix} \tag{3.4}$$

and, in the same vein, we also define

$$T_0(q) = \begin{bmatrix} G_0(q) \\ H_0(q) \end{bmatrix} \tag{3.5}$$

For future use we further introduce the signal

$$\chi_0(t) = \begin{bmatrix} u(t) & e(t) \end{bmatrix}^T \tag{3.6}$$

The spectrum for $\chi_0(t)$ is

$$\Phi_{\chi_0}(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{eu}(\omega) & \lambda_0 \end{bmatrix} \tag{3.7}$$

We will also frequently use

$$\Phi_u^r(\omega) = \Phi_u(\omega) - |\Phi_{ue}(\omega)|^2/\lambda_0 \tag{3.8}$$

which we recognize as the Schur complement of λ_0 in the matrix $\Phi_{\chi_0}(\omega)$ (see, e.g., [53]). The matrix $\Phi_u^r(\omega)$ defined in (3.8) can be seen as that part of the input spectrum that originates from the reference signal $r(t)$. Using (2.6) we can easily derive the following expression for $\Phi_u^r(\omega)$ under (2.4):

$$\Phi_u^r(\omega) = |S_0(e^{i\omega})|^2 \Phi_r(\omega) \tag{3.9}$$

Here $\Phi_r(\omega)$ is the spectrum of the reference signal. If $u(t)$ is generated as in (2.4) the cross-spectrum between $u(t)$ and $e(t)$ is

$$\Phi_{ue}(\omega) = -K(e^{i\omega})S_0(e^{i\omega})H_0(e^{i\omega})\lambda_0 \tag{3.10}$$

The spectrum of $v(t)$ is

$$\Phi_v(\omega) = |H_0(e^{i\omega})|^2 \lambda_0 \tag{3.11}$$

Let us finally introduce the quantity

$$\Phi_e^r(\omega) = \lambda_0 - |\Phi_{ue}(\omega)|^2/\Phi_u(\omega) \tag{3.12}$$

which is the Schur complement of $\Phi_u(\omega)$ in the matrix $\Phi_{\chi_0}(\omega)$.

3.2 The Method

The one-step-ahead predictor for the model structure (3.1) is

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t) \tag{3.13}$$

The prediction error is

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{3.14}$$

Given the model (3.13) and measured data

$$Z^N = \{y(1),\ u(1),\ \ldots,\ y(N),\ u(N)\} \tag{3.15}$$

we determine the prediction error estimate through

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N) \tag{3.16}$$

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \ell(\varepsilon(t,\theta)) \tag{3.17}$$

Here $\ell(\cdot)$ is a suitably chosen positive (norm) function. The standard choice is

$$\ell(\varepsilon(t,\theta)) = \frac{1}{2}\varepsilon^2(t,\theta) \tag{3.18}$$

possibly combined with some linear, monic, and possibly parameterized prefilter $L(q,\theta)$:

$$\ell(\varepsilon(t,\theta)) = \frac{1}{2}\varepsilon_F^2(t,\theta) \tag{3.19}$$

$$\varepsilon_F(t,\theta) = L(q,\theta)\varepsilon(t,\theta) \tag{3.20}$$

However, since

$$\varepsilon_F(t,\theta) = L(q,\theta)H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{3.21}$$

the effect of the prefilter can be included in the noise model and $L(q,\theta) = 1$ can be assumed without loss of generality. This will be done in the sequel.

The above choice of $\ell(\cdot)$ gives the least-squares method. If we let the criterion function $\ell(\cdot)$ be equal to the negative logarithm of $f_e(\cdot)$, the probability density function of $e(t)$, we obtain the maximum likelihood method. These and other choices of norms are further discussed in [71]. In this thesis we will for simplicity assume that a quadratic criterion is used.

The following notation will be used for the estimates $G(q,\hat{\theta}_N)$, $H(q,\hat{\theta}_N)$, etc.:

$$\hat{G}_N(q) = G(q,\hat{\theta}_N) \text{ and } \hat{H}_N(q) = H(q,\hat{\theta}_N) \tag{3.22}$$

3.3 Some Common Model Structures

With the prediction error approach, the models may have arbitrary parameterizations. However, in most cases it is convenient and also sufficient to use some standard model structure. Consider the following generalized model structure [71]:

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{3.23}$$

where

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \tag{3.24}$$

and similarly for the C-, D-, and F-polynomials, while

$$B(q) = b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{n_b} q^{-n_b} \tag{3.25}$$

It is also possible to redefine the B-polynomial to include extra delays. For simplicity, this is not done here.

The model structure (3.23) contains several common special cases, some of which are listed in Table 3.1.

Table 3.1 *Some Common Model Structures as Special Cases of* (3.23).

Polynomials Used in (3.23)	Name of Model Structure
B	FIR
A,B	ARX
A,B,C	ARMAX
B,F	OE
B,C,D,F	BJ

In Table 3.1 the acronyms FIR, ARX, ARMAX, OE and BJ denote Finite Impulse Response, AutoRegressive with eXogenous (or eXternal) input, AutoRegressive Moving Average with eXogenous input, Output Error and Box-Jenkins, respectively.

3.4 Computing the Estimate

If an FIR or an ARX model is used together with a quadratic criterion function, we obtain the standard least-squares method. To find the estimate $\hat{\theta}_N$ we then only have to solve a standard least-squares problem, which can be done without iterations. For other parameterizations and criteria we typically have to rely on some numerical search scheme to find the estimate. The standard choice is to use a search routine of the form

$$\hat{\theta}_N^{(i+1)} = \hat{\theta}_N^{(i)} - \mu_N^{(i)} [R_N^{(i)}]^{-1} V_N'(\hat{\theta}_N^{(i)}, Z^N) \tag{3.26}$$

where $V_N'(\hat{\theta}_N^{(i)}, Z^N)$ denotes the gradient of the criterion function $V_N(\hat{\theta}_N^{(i)}, Z^N)$ with respect to θ, $R_N^{(i)}$ is a matrix that modifies the search direction, and $\mu_N^{(i)}$ is a scaling factor that determines the step length.

With the criterion

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N}\frac{1}{2}\varepsilon^2(t,\theta) \tag{3.27}$$

we have

$$V_N'(\hat{\theta}_N^{(i)}, Z^N) = -\frac{1}{N}\sum_{t=1}^{N}\psi(t,\theta)\varepsilon(t,\theta) \tag{3.28}$$

where $\psi(t,\theta)$, of dimension $d \times 1$, denotes the negative gradient of the prediction error:

$$\psi(t,\theta) = -\frac{d}{d\theta}\varepsilon(t,\theta) \ (= \frac{d}{d\theta}\hat{y}(t|\theta)) \tag{3.29}$$

A common choice for the matrix $R_N^{(i)}$ is

$$R_N^{(i)} = \frac{1}{N}\sum_{t=1}^{N}\psi(t,\theta)\psi^T(t,\theta) \tag{3.30}$$

which gives the Gauss-Newton method (e.g., [20]).

Here we can note that to be able to implement the algorithm (3.26), the filters $H^{-1}(q,\theta)G(q,\theta)$ and $H^{-1}(q,\theta)$ along with their derivatives, need to be uniformly stable for $\theta \in D_{\mathcal{M}}$ [71]. This is automatically satisfied with FIR and ARX models and can also be ensured with ARMAX models by restricting the noise model to be minimum phase which typically is no restriction (cf. Paper E). However, when the underlying system is unstable and an output error or a Box-Jenkins model is used the situation is more problematic. Consider the standard output error predictor

$$\hat{y}(t|\theta) = \frac{B(q)}{F(q)}u(t) \tag{3.31}$$

If the underlying system is unstable, this is bound to become unstable at some point in the iterations. The standard way to resolve this problem is simply to project θ into the region of stability. In some cases this is acceptable, but if the true system indeed is unstable we will of course get an erroneous estimate with this approach. A more satisfying solution is provided in Paper D. The idea is to use the following modified output error model structure:

$$y(t) = \frac{B(q)}{F(q)}u(t) + \frac{F_a^*(q)}{F_a(q)}e(t) \tag{3.32}$$

where $1/F_a(q)$ is the anti-stable part of $1/F(q)$ and where $1/F_a^*(q)$ is the anti-stable part with the poles reflected into the unit disc. Since the "noise model" $F_a^*(q)/F_a(q)$ is an all-pass filter whose poles are exactly the unstable poles of the dynamics model, the predictor for the modified structure (3.32) is guaranteed to be stable. Furthermore, since $F_a^*(q)/F_a(q)$ is all-pass the resulting prediction error estimate will asymptotically be the same as with the standard output error model structure. The details are given in Paper D.

Another way to handle the problem with unstable predictor filters is to consider noncausal filtering. The idea is then to divide the filters into their stable and anti-stable parts and to filter the data backwards through the anti-stable part. This leads to conceptually very simple solutions with comparable performances as the approach above. See Paper E.

3.5 Analysis Results for the Prediction Error Method

3.5.1 Consistency and Identifiability

Consider the prediction error method applied with a quadratic criterion

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N}\frac{1}{2}\varepsilon^2(t,\theta) \tag{3.33}$$

Under mild conditions we have that

$$V_N(\theta, Z^N) \to \bar{V}(\theta) = \bar{E}\frac{1}{2}\varepsilon^2(t,\theta) \quad \text{w.p. } 1 \text{ as } N \to \infty \tag{3.34}$$

and

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \bar{V}(\theta) \quad \text{w.p. } 1 \text{ as } N \to \infty \tag{3.35}$$

See [67].

Using Parseval's relationship, we can write

$$\bar{V}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\Phi_\varepsilon(\omega)d\omega \tag{3.36}$$

where $\Phi_\varepsilon(\omega)$ is the spectrum of the prediction error. If the data was generated by

$$y(t) = G_0(q)u(t) + H_0(q)e(t) \tag{3.37}$$

we see that

$$\begin{aligned}\varepsilon(t,\theta) &= H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \\ &= H^{-1}(q,\theta)\big[(G_0(q) - G(q,\theta))u(t) + (H_0(q) - H(q,\theta))e(t)\big] + e(t) \\ &= H^{-1}(q,\theta)\tilde{T}^T(q,\theta)\chi_0(t) + e(t) \end{aligned} \tag{3.38}$$

In the last step we introduced $\tilde{T}(q,\theta) = T_0(q) - T(q,\theta)$.

Suppose now that $G(q,\theta)u(t)$ (as well as $G_0(q)u(t)$) depends only on $e(s)$ for $s < t$ and that $H(q,\theta)$ (and $H_0(q)$) is monic, then the last term is independent of the rest and

$$\Phi_\varepsilon(\omega) = \frac{1}{|H(e^{i\omega},\theta)|^2}\tilde{T}^T(e^{i\omega},\theta)\Phi_{\chi_0}(\omega)\tilde{T}(e^{-i\omega},\theta) + \lambda_0 \tag{3.39}$$

Together with (3.35) and (3.36) this shows that

$$\hat{\theta}_N \to D_c = \arg\min_{\theta\in D_{\mathcal{M}}} \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{T}^T(e^{i\omega},\theta)\Phi_{\chi_0}(\omega)\tilde{T}(e^{-i\omega},\theta)\frac{1}{|H(e^{i\omega},\theta)|^2}d\omega \quad \text{w.p. 1 as } N\to\infty \tag{3.40}$$

Here we have dispensed with the scaling 1/2 in (3.36) which is immaterial for the result.

The result (3.40) shows that the prediction error estimate $T(q,\hat{\theta}_N)$ will converge to the "true" transfer function $T_0(q)$ if the parameterization is flexible enough so that $\mathcal{S}\in\mathcal{M}$ and if the structure $\tilde{T}^T(e^{i\omega},\theta)$ does not lie in the (left) null space of $\Phi_{\chi_0}(\omega)$. In [71] the latter is called that the data is *informative enough* with respect to the chosen model structure. An *informative* data set is one for which the matrix $\Phi_{\chi_0}(\omega)$ is positive definite for almost all frequencies. In this case, it follows that the limit estimate will always be such that $\tilde{T}^T(e^{i\omega},\theta) = 0$, that is, $\hat{T}_N(e^{i\omega})$ will tend to $T_0(q)$ as the number of data tends to infinity.

In fact, we can be more specific about when the data will be informative. Consider the factorization

$$\Phi_{\chi_0}(\omega) = \begin{bmatrix}1 & \Phi_{ue}(\omega)/\lambda_0 \\ 0 & 1\end{bmatrix}\begin{bmatrix}\Phi_u^r(\omega) & 0 \\ 0 & \lambda_0\end{bmatrix}\begin{bmatrix}1 & 0 \\ \Phi_{eu}(\omega)/\lambda_0 & 1\end{bmatrix} \tag{3.41}$$

From this it follows that $\Phi_{\chi_0}(\omega)$ is positive definite (for almost all frequencies) if, and only if, $\Phi_u^r(\omega)$ is positive (for almost all frequencies). The general condition for this to be true is that there should *not* be a linear, noise-free relation between the input and the output, see [71] and Section 2.4 above. With a linear feedback law (2.4) we have that $\Phi_u^r(\omega) = |S_0(e^{i\omega})|^2\Phi_r(\omega)$ and since the analytic function $S_0(e^{i\omega})$ only can have finitely many zeros, the condition in this case becomes $\Phi_r(\omega) > 0$ (for almost all frequencies), that is, that the reference signal $r(t)$ should be persistently exciting [71].

3.5.2 Bias Distribution

Inserting the factorization

$$\Phi_{\chi_0}(\omega) = \begin{bmatrix} 1 & 0 \\ \Phi_{eu}(\omega)/\Phi_u(\omega) & 1 \end{bmatrix} \begin{bmatrix} \Phi_u(\omega) & 0 \\ 0 & \Phi_e^r(\omega) \end{bmatrix} \begin{bmatrix} 1 & \Phi_{ue}(\omega)/\Phi_u(\omega) \\ 0 & 1 \end{bmatrix} \tag{3.42}$$

into (3.40) immediately gives the following characterization of the limit estimate as the number of data tends to infinity:

$$\begin{aligned} \hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \big\{ |G_0(e^{i\omega}) + B(e^{i\omega}, \theta) - G(e^{i\omega}, \theta)|^2 \Phi_u(\omega) \\ + |H_0(e^{i\omega}) - H(e^{i\omega}, \theta)|^2 \Phi_e^r(\omega) \big\} \frac{1}{|H(e^{i\omega}, \theta)|^2} d\omega \quad \text{w.p. 1 as } N \to \infty \end{aligned} \tag{3.43}$$

where

$$B(e^{i\omega}, \theta) = (H_0(e^{i\omega}) - H(e^{i\omega}, \theta)) \Phi_{eu}(\omega)/\Phi_u(\omega) \tag{3.44}$$

A variant of this result is given in Paper A. It follows that the bias inclination in the G-estimate will be small in frequency ranges where either the noise model is good ($|H_0(e^{i\omega}) - H(e^{i\omega}, \theta)|$ is small) or the input spectrum $\Phi_u(\omega)$ dominates the cross-spectrum $\Phi_{eu}(\omega)$. Also notice that the bias term $B(e^{i\omega}, \theta)$ is always identically zero in open loop since then $\Phi_{eu}(\omega) = 0$. For more details on these issues, see Paper A.

3.5.3 Asymptotic Variance of Parameter Estimates

If $\mathcal{S} \in \mathcal{M}$ and the data is sufficiently rich we have that (see, e.g., [72])

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, P_\theta) \tag{3.45}$$

$$P_\theta = \lambda_0 R^{-1} \tag{3.46}$$

$$R = \bar{E}\psi(t, \theta_0)\psi^T(t, \theta_0) \tag{3.47}$$

If the noise $e(t)$ has a normal distribution, the matrix R equals the, so called, Fischer information matrix which shows that the prediction error method is efficient and meets the Cramér-Rao lower bound. By noting that

$$\psi(t, \theta) = -\frac{d}{d\theta}\varepsilon(t, \theta) = \frac{1}{H(q, \theta)} (T'(q, \theta))^T \begin{bmatrix} u(t) \\ \varepsilon(t, \theta) \end{bmatrix} \tag{3.48}$$

where $T'(q, \theta)$ of dimension $2 \times d$ denotes the derivative of $T(q, \theta)$ with respect to θ, we see that the matrix R can alternatively be written as

$$R = \frac{1}{2\pi} \int_{-\pi}^{\pi} (T'(e^{i\omega}, \theta_0))^T \Phi_{\chi_0}(\omega) T'(e^{-i\omega}, \theta_0) \frac{1}{|H_0(e^{i\omega})|^2} d\omega \tag{3.49}$$

If we specialize to the linear feedback case (2.4) and assume that the model is parameterized as

$$y(t) = G(q,\rho)u(t) + H(q,\beta)e(t), \quad \theta = \begin{bmatrix} \rho \\ \beta \end{bmatrix} \tag{3.50}$$

we get (cf. Paper A):

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P_\rho^{\mathrm{D}}) \tag{3.51a}$$

where

$$P_\rho^{\mathrm{D}} = \lambda_0 (R_\rho^r + \Delta)^{-1} \tag{3.51b}$$

$$R_\rho^r = \frac{1}{2\pi}\int_{-\pi}^{\pi} G'_\rho(e^{i\omega},\rho_0)(G'_\rho(e^{-i\omega},\rho_0))^T \frac{\Phi_u^r(\omega)}{|H_0(e^{i\omega})|^2} d\omega \tag{3.51c}$$

$$\Delta = R_\rho^e - R_{\rho\beta}^e (R_\beta^e)^{-1} R_{\beta\rho}^e \tag{3.51d}$$

$$R_\rho^e = \frac{1}{2\pi}\int_{-\pi}^{\pi} \lambda_0 |K(e^{i\omega})|^2 |S_0(e^{i\omega})|^2 G'_\rho(e^{i\omega},\rho_0)(G'_\rho(e^{-i\omega},\rho_0))^T d\omega \tag{3.51e}$$

$$R_{\rho\beta}^e = -\frac{1}{2\pi}\int_{-\pi}^{\pi} G'_\rho(e^{i\omega},\rho_0)(H'_\beta(e^{-i\omega},\beta_0))^T \frac{\lambda_0}{|H_0(e^{i\omega})|^2} K(e^{i\omega}) S_0(e^{i\omega}) H_0(e^{i\omega}) d\omega \tag{3.51f}$$

$$R_\beta^e = \frac{1}{2\pi}\int_{-\pi}^{\pi} H'_\beta(e^{i\omega},\beta_0)(H'_\beta(e^{-i\omega},\beta_0))^T \frac{\lambda_0}{|H_0(e^{i\omega})|^2} d\omega \tag{3.51g}$$

Here $G'_\rho(e^{i\omega},\rho_0)$ denotes the $(\dim\rho)\times 1$ dimensional derivative of $G(e^{i\omega},\rho)$ with respect to ρ, evaluated for $\rho = \rho_0$, etc.

Furthermore,

$$0 \le \Delta \le R_\rho^e \tag{3.52}$$

and Δ approaches zero as the order of the noise model tends to infinity, while $\Delta = R_\rho^e$ if the noise model is fixed and equal to the true one, that is, if $H(q,\beta) = H_*(q) = H_0(q)$.

These results will be further commented upon below.

3.5.4 Asymptotic Variance for Identified High-order Black-box Models

Suppose that the vector θ can be decomposed so that

$$\theta = \left[\theta_1^T,\ \theta_2^T,\ \ldots,\ \theta_n^T\right]^T \quad \dim\theta_k = s \quad \dim\theta = n\cdot s \tag{3.53}$$

and that $T(q,\theta)$ has the following shift structure:

$$\frac{\partial}{\partial \theta_k} T(q,\theta) = q^{-k+1} \frac{\partial}{\partial \theta_1} T(q,\theta) \tag{3.54}$$

(This condition is satisfied by most polynomial-type model structures, including (3.23). Hence this is a rather mild assumption.) Then it can be shown that the transfer function estimate $T(e^{i\omega}, \hat{\theta}_N)$ asymptotically, as N as well as the model order n (i.e., the dimensionality of θ) tends to infinity, has a normal distribution with covariance matrix

$$\operatorname{Cov} T(e^{i\omega}, \hat{\theta}_N) \approx \frac{n}{N} \Phi_v(\omega) \Phi_{\chi_0}^{-1}(-\omega) \text{ as } N,\ n \to \infty \tag{3.55}$$

This holds regardless of the nature of the feedback as long as certain technical conditions hold. (For brevity these conditions are not listed here; instead these can be found in Paper A. See also Paper H and [68, 87, 125].) If we concentrate on the asymptotic variance for the estimate $\hat{G}_N(e^{i\omega})$ we have that

$$\operatorname{Var} \hat{G}_N(e^{i\omega}) \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u^r(\omega)} \text{ as } N,\ n \to \infty \tag{3.56}$$

This result follows immediately from (3.55) by noting that

$$\Phi_{\chi_0}^{-1}(-\omega) = \begin{bmatrix} \frac{1}{\Phi_u^r(\omega)} & -\frac{\Phi_{eu}(\omega)}{\lambda_0 \Phi_u^r(\omega)} \\ -\frac{\Phi_{ue}(\omega)}{\lambda_0 \Phi_u^r(\omega)} & \frac{1}{\Phi_e^r(\omega)} \end{bmatrix} \tag{3.57}$$

3.5.5 Discussion

The direct method, that is, the standard prediction error method applied directly to input-output data, gives consistency and optimal accuracy provided that the parameterization contains the true system. Furthermore, since no assumptions are made on how the data was generated, this method is applicable to systems with arbitrary feedback mechanisms. The direct method should therefore be seen as the first choice of methods for closed-loop identification. However, in the literature several other "methods"/parameterizations have also been suggested. In the following sections we will study two of these, first the indirect and then the projection method.

3.6 The Indirect Method

If the control law is known we can use the indirect method. For simplicity, we will here restrict to the linear feedback case (2.4). The idea can then be explained as follows: Identify the closed-loop system $G_0^c(q)$ defined in (2.8) (or some other closed-loop transfer function) and compute an open-loop model from the resulting estimate $\hat{G}_N^c(q)$ using the knowledge of the regulator $K(q)$. With the prediction error method (that allows arbitrary parameterizations), it is natural to parameterize the model of the closed-loop system in terms of the open-loop model and use

$$y(t) = \frac{G(q,\theta)}{1+G(q,\theta)K(q)}r(t) + H_*(q)e(t) \tag{3.58}$$

Here we have included a fixed noise model $H_*(q)$ which is common in this approach. With this parameterization the somewhat problematic second step, of computing the open-loop estimate, is avoided and the open-loop model is delivered directly.

It can be noted that the indirect method is equivalent to the direct method with the noise model $H(q,\theta)$ parameterized as $H(q,\theta) = (1 + G(q,\theta)K(q))H_*(q)$. This is further discussed in [75] and is also covered in Paper A. Let us also mention that another, very interesting, parameterization idea is to use the, so called, dual Youla-Kucera parameterization which parameterizes the set of all models that are stabilized by a certain linear controller. A large number of papers have been written on how to use this idea for identification (see, e.g., [2, 17, 45, 46, 95, 110]), but as shown in Paper A this is equivalent to using the standard indirect method, at least as far as the statistical properties of the resulting estimates are concerned.

Let us now turn to the statistical properties of the indirect method. If the model structure is (3.58) we have that

$$\hat{\theta}_N \to D_c = \arg\min_{\theta\in D_{\mathcal{M}}} \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{G_0(e^{i\omega}) - G(e^{i\omega},\theta)}{1+G(e^{i\omega},\theta)K(e^{i\omega})}\right|^2 \frac{|S_0(e^{i\omega})|^2\Phi_r(\omega)}{|H_*(e^{i\omega})|^2}d\omega$$
$$\text{w.p. } 1 \text{ as } N\to\infty \tag{3.59}$$

This means that the indirect method can give consistent estimates of $G_0(e^{i\omega})$ if the parameterization of $G(e^{i\omega},\theta)$ is flexible enough, that is, if $G_0 \in \mathcal{G}$, even with a fixed noise model. This is the main advantage of the indirect method. In case of undermodeling the resulting G-estimate will try to minimize the mismatch between $G_0(e^{i\omega})$ and $G(e^{i\omega},\theta)$ and at the same time try to minimize the model sensitivity function $S(e^{i\omega},\theta) = (1 + G(e^{i\omega},\theta)K(e^{i\omega}))^{-1}$. There will thus be a "bias-pull" towards transfer functions that give a small sensitivity for the given regulator, which can be advantageous, for instance, if the model is to be used for control [32, 33, 112]. Unlike (3.43) it is not easy to quantify this bias error in the open-loop model, though.

Now suppose that the model structure is changed to

$$y(t) = \frac{G(q,\rho)}{1+G(q,\rho)K(q)}r(t) + H^c(q,\beta)e(t), \quad \theta = \begin{bmatrix}\rho\\ \beta\end{bmatrix} \tag{3.60}$$

and that $\mathcal{S} \in \mathcal{M}$ holds for this model. Under mild conditions, we then have that

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P_\rho^{\mathrm{I}}) \tag{3.61a}$$

where

$$P_\rho^{\mathrm{I}} = \lambda_0 (R_\rho^r)^{-1} \tag{3.61b}$$

with R_ρ^r given by (3.51c). Furthermore, if the noise model in (3.60) is fixed $H^c(q, \beta) = H_*(q)$ (or, equivalently, if (3.58) is used) and the model structure now satisfies $G_0 \in \mathcal{G}$ we have that

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, \bar{P}_\rho^{\mathrm{I}}) \tag{3.62a}$$

where

$$\bar{P}_\rho^{\mathrm{I}} = R^{-1} Q R^{-1} \tag{3.62b}$$

$$R = \frac{1}{2\pi} \int_{-\pi}^{\pi} G_\rho'(e^{i\omega}, \rho_0)(G_\rho'(e^{-i\omega}, \rho_0))^T \frac{|S_0(e^{i\omega})|^2 \Phi_u^r(\omega)}{|H_*(e^{i\omega})|^2} d\omega \tag{3.62c}$$

$$Q = \frac{1}{2\pi} \int_{-\pi}^{\pi} G_\rho'(e^{i\omega}, \rho_0)(G_\rho'(e^{-i\omega}, \rho_0))^T \frac{|S_0(e^{i\omega})|^4 \Phi_v(\omega) \Phi_u^r(\omega)}{|H_*(e^{i\omega})|^4} d\omega \tag{3.62d}$$

It is interesting to compare these results with the corresponding ones that hold for the direct method, see (3.51). First of all, we have the following ranking between the asymptotical covariance matrices for the methods:

$$P_\rho^{\mathrm{D}} \leq P_\rho^{\mathrm{I}} \leq \bar{P}_\rho^{\mathrm{I}} \tag{3.63}$$

This shows that the indirect method typically gives worse accuracy than the optimal direct method. The first inequality is due to the term Δ in (3.51) that is missing in (3.61). An important observation regarding the term Δ is that it is entirely due to the noise part of the input spectrum. Since $\Delta \geq 0$, this contribution has a positive effect on the accuracy, contrarily to what one might have guessed. We conclude that in the direct method the noise in the loop is utilized in reducing the variance. For the indirect methods this contribution is zero as can be seen from (3.61). The reason for this is that in the indirect method only that part of the input that originates from the reference signal is used in the identification. Another comment regarding the term Δ is that it vanishes as more and more parameters are used to estimate the noise color. It is as large as it can be if the noise model is fixed and correct, see Paper A. We can also note that the second inequality becomes an equality if $H_*(q) = H_0^c(q) = S_0(q)H_0(q)$, that is, if a fixed and correct noise model is used.

A final comment on the statistical properties of the indirect methods is that, if the feedback is linear and of the form (2.4), the result (3.56) holds also for this method. Here we should remember that $\Phi_u^r(\omega) = |S_0(e^{i\omega})|^2 \Phi_r(\omega)$ in case of linear feedback.

3.7 The Projection Method

Paper B describes the projection method. This method can be seen as a joint input-output method, as is argued in Paper A, and can be explained using the following steps:

1. Estimate the parameters s_k in the noncausal FIR model

$$u(t) = S(q)r(t) + e(t) = \sum_{k=-M_1}^{M_2} s_k r(t-k) + e(t) \tag{3.64}$$

   where M_1 and M_2 are chosen so large that any correlation between $u(t)$ and $r(s)$ for $t - s \notin [-M_1, \ldots, M_2]$ can be ignored, and simulate the signal $\hat{u}(t) = \hat{S}(q)r(t)$.

2. Identify the open-loop system using a model of the kind

$$y(t) = G(q, \theta)\hat{u}(t) + H_*(q)e(t) \tag{3.65}$$

Since $\hat{u}(t)$ is constructed from $r(t)$, which is uncorrelated with the noise, we have converted the original closed-loop identification problem into an open-loop one. If the feedback is linear, the signal $\hat{u}(t)$ will, ideally, be exactly equal to the noise-free part of the input, $S_0(q)r(t)$. This idea is thus similar to the idea used in the instrumental variable approach: Make the instruments approximations of the noise-free parts of the regressors.

The projection method is in form very similar to the, so called, two-stage method of Van den Hof and Schrama [111]. The difference is the noncausal FIR model used in the projection method; the two-stage method only considers causal models in the first step. As shown in Paper B, it is exactly this difference that makes the projection method applicable to systems with arbitrary, nonlinear feedback mechanisms, whereas the two-stage method can only be used with linear feedback.

For the projection method,

$$y(t) = G(q, \theta)\hat{u}(t) + H_*(q)e(t), \quad \hat{u}(t) = \hat{S}_N(q)r(t), \; \tilde{u}(t) = u(t) - \hat{u}(t) \tag{3.66}$$

we have that

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \frac{1}{2\pi} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) + \tilde{B}(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 \frac{\Phi_{\hat{u}}(\omega)}{|H_*(e^{i\omega})|^2} d\omega$$
$$\text{w.p. 1 as } N \to \infty \tag{3.67}$$

where

$$\tilde{B}(e^{i\omega}) = G_0(e^{i\omega})\Phi_{\tilde{u}\hat{u}}(\omega)\Phi_{\hat{u}}^{-1}(\omega) \tag{3.68}$$

The result (3.67) holds for both the two-stage and the projection method and for arbitrary feedback mechanisms. It is clear that any correlation between $\hat{u}$ and $\tilde{u}$ will deteriorate the G-estimate in these methods. The error is quantified by the term $\tilde{B}(e^{i\omega})$ defined in (3.68). With the projection method, $\Phi_{\tilde{u}\hat{u}}(\omega) = 0$ is (asymptotically) achieved by construction, which implies that $\tilde{B}(e^{i\omega}) = 0$. Hence, the model can be fit to the data with a fixed, user-specified frequency weighting with this method by changing the noise model $H_*(q)$. (Observe that $\Phi_{\hat{u}}(\omega)$ is known when the second step of the algorithm is applied.) This is a clear advantage compared to the direct method, which requires a flexible, parameterized noise model to give unbiased results and thus does not allow a customized fit (cf. (3.43)). Also note that this holds regardless of the feedback. The projection method is therefore applicable to systems with arbitrary feedback mechanisms, just as the direct method.

A drawback with the projection method is that it gives suboptimal accuracy. The reason for this is basically that only the "noise-free part" $\Phi_{\hat{u}}(\omega)$ of the total input spectrum $\Phi_u(\omega)$ is utilized for reducing the variance, not $\Phi_u(\omega)$ itself as with the direct method. This is the price we have to pay for obtaining a method that gives an explicitly tunable criterion for the model fit. For more details on the projection method, see Paper B.

Chapter 4

Utilizing Periodic Excitation in System Identification

4.1 Introduction

The use of periodic excitation signals in identification is an often overlooked possibility in time-domain identification. With periodic excitation a number of things are gained, the most obvious being the data reduction due to the averaging of the measured signals over the periods. Another advantage is that with periodic excitation it is possible to separate the "signal" from the "noise" and hence the noise properties can be estimated separately. This information can be used to improve the quality of the estimates (cf. minimum variance estimation) or to construct identification methods that can handle special situations, like when both the input and the output are corrupted by noise. This is called an *errors-in-variables problem* and in this chapter we will study methods that explicitly use the periodicity of the excitation signals to handle such problems. The main reference to the results presented here is Paper F.

4.2 Notation and Technical Assumptions

Consider the linear regression model

$$\begin{aligned} y(t) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) \\ + b_0 u(t-n_k) + \cdots + b_{n_b} u(t-n_k-n_b) \end{aligned} \tag{4.1}$$

The problem we will study is how to identify the model parameters $\{a_i\}$ and $\{b_i\}$ using noisy measurements of $y(t)$ and $u(t)$.

Our measured data can be described by

$$Z_m^N = \{z_m(1), \ldots, z_m(N)\} \tag{4.2}$$

$$z_m(t) = z(t) + w(t) \tag{4.3}$$

$$z(t) = \begin{bmatrix} y(t) & u(t) \end{bmatrix}^T, \; w(t) = \begin{bmatrix} w_y(t) & w_u(t) \end{bmatrix}^T \tag{4.4}$$

We will also use the notation $y_m(t) = y(t) + w_y(t)$ and $u_m(t) = u(t) + w_u(t)$. The unknown signals $w_y(t)$ and $w_u(t)$ act as noise sources on the measured output and input, respectively. It is assumed that $u(t)$ is periodic with period P and that an integer number of periods are measured: $N = M \cdot P$. We further assume that $w(t)$ is a zero mean, quasi-stationary signal and that $w(t)$ is uncorrelated with $z(t)$.

4.3 Estimating the Noise Statistics

The first operation we should do once we have obtained the measured data is to average it over the periods:

$$\bar{z}(t) = \frac{1}{M} \sum_{k=0}^{M-1} z_m(t+kP), \; t \in [1, P] \tag{4.5}$$

By periodically continuing $\bar{z}(t)$ outside $t = [1, P]$ we can estimate the noise $w(t)$ as

$$\hat{w}(t) = z_m(t) - \bar{z}(t), \; t \in [1, N] \tag{4.6}$$

A consistent estimate of the covariance function

$$R_{ww}(k) = E w(t) w^T(t+k) \tag{4.7}$$

can now be computed as

$$\hat{R}_{ww}(k) = \frac{1}{(M-1)P} \sum_{t=1}^{MP} \hat{w}(t) \hat{w}^T(t+k) \tag{4.8}$$

where the convention is that all signals outside the interval $t = [1, MP]$ are replaced by 0. In practice, for large data sets, the covariance function should be computed using FFT, see [94]. It is important to note that we have used P degrees of freedom for estimating the mean, so the proper normalization to get an unbiased estimate is $MP - P = (M-1)P$. How many periods do we need then? The rather precise answer provided in [93] is $M \geq 4$. The asymptotic properties ($N = MP \to \infty$) of the estimate are then independent of how the excitation is divided into M and P.

An unbiased estimate of the spectrum of $w(t)$ is obtained by the periodogram

$$\hat{\Phi}_w(\omega) = \sum_{k=-MP+1}^{MP-1} \hat{R}_{ww}(k) e^{-i\omega k} \tag{4.9}$$

This can be used for prewhitening of $w(t)$ prior to the estimation. It turns out that the poor variance properties of (4.9) does not diminish its usefulness for prewhitening, see Paper F and [39]. We also mention that $\hat{\Phi}_w(\omega)$ can be estimated very efficiently directly from the original data using FFT.

4.4 Least-squares Estimation Using Periodic Excitation

Introduce the notation

$$\theta = \left[a_1, \ldots, a_{n_a}, b_0, \ldots, b_{n_b}\right]^T \tag{4.10}$$

$$\varphi_m(t) = \left[-y_m(t-1), \ldots, -y_m(t-n_a), u_m(t-n_k), \ldots, u_m(t-n_k-n_b)\right]^T \tag{4.11}$$

$$\varphi_z(t) = \left[-y(t-1), \ldots, -y(t-n_a), u(t-n_k), \ldots, u(t-n_k-n_b)\right]^T \tag{4.12}$$

$$\varphi_w(t) = \left[-w_y(t-1), \ldots, -w_y(t-n_a), w_u(t-n_k), \ldots, w_u(t-n_k-n_b)\right]^T \tag{4.13}$$

The least-squares estimate of θ using N data samples can be written

$$\hat{\theta}_N = R_m^{-1}(N) f_m(N) \tag{4.14}$$

$$R_m(N) = \frac{1}{N}\sum_{t=1}^{N} \varphi_m(t)\varphi_m^T(t) \tag{4.15}$$

$$f_m(N) = \frac{1}{N}\sum_{t=1}^{N} \varphi_m(t) y_m(t) \tag{4.16}$$

An important observation is now that both $R_m(N)$ and $f_m(N)$ originate from two separate and uncorrelated sources: $z(t)$ and $w(t)$. Thus

$$\lim_{N\to\infty} R(N) = R_m = R_z + R_w \tag{4.17}$$

$$\lim_{N\to\infty} f(N) = f_m = f_z + f_w \tag{4.18}$$

$$R_z = \bar{E}\varphi_z(t)\varphi_z^T(t), \quad R_w = \bar{E}\varphi_w(t)\varphi_w^T(t) \tag{4.19}$$

$$f_z = \bar{E}\varphi_z(t)y(t), \qquad f_w = \bar{E}\varphi_w(t)w_y(t) \tag{4.20}$$

In the ideal situation where the prediction error

$$\varepsilon(t,\theta) = y_m(t) - \varphi_m^T(t)\theta \tag{4.21}$$

is a white noise sequence for the "true" parameter value $\theta = \theta_0$, the least-squares estimate (4.14) is consistent with asymptotic covariance matrix (cf. (3.45))

$$N \operatorname{Cov} \hat{\theta}_N \approx \lambda_0 R_m^{-1} \tag{4.22}$$

where λ_0 is the variance of $\varepsilon(t,\theta_0)$.

Let $\bar{R}(P)$ and $\bar{f}(P)$ be defined similar to $R_m(P)$ and $f_m(P)$, respectively, except that averaged data is used. We then have that

$$\lim_{P\to\infty} \bar{R}(P) = \bar{R} = R_z + \frac{1}{M}R_w \tag{4.23}$$

$$\lim_{P\to\infty} \bar{f}(P) = \bar{f} = f_z + \frac{1}{M}f_w \tag{4.24}$$

The M normalization is due to the averaging which decreases the noise variance with a factor of M. The least-squares estimate based on averaged data will still be unbiased if the prediction error (4.21) is white, but the asymptotic covariance matrix changes to

$$N \operatorname{Cov} \hat{\theta}_P \approx \lambda_0 \bar{R}^{-1} \tag{4.25}$$

The scaling factor is thus the same, but R_m is replaced by $\bar{R}$, and since $R_m \geq \bar{R}$ it follows that the asymptotic covariance increases with averaged data. However, as shown in [39] and also in Paper F, it is possible to compensate for this loss in accuracy using the nonparametric noise model (4.8).

Here we will go directly to the main result in Paper F, namely that it is possible to compensate for the bias in the least-squares estimate incurred if the data is not generated as (cf. (4.21))

$$y_m(t) = \varphi_m^T(t)\theta_0 + e(t) \tag{4.26}$$

where $e(t)$ is a white noise signal and where θ_0 denotes the true parameter vector, as usual.

The main idea is to construct nonparametric estimates $\hat{R}_w^{np}$ and $\hat{f}_w^{np}$ of R_w and f_w, respectively, from $\hat{R}_{ww}(k),\ k = 0, 1, \ldots$ and to use these to correct for the bias in the least-squares estimate due to the (colored) noise $w(t)$. We will here describe an approach where the aim is to remove all effects of the noise $w(t)$. Form the estimates

$$\hat{R}_z(P) = \bar{R}(P) - \frac{1}{M-1}\hat{R}_w^{np}, \quad \hat{f}_z(P) = \bar{f}(P) - \frac{1}{M-1}\hat{f}_w^{np} \tag{4.27}$$

of R_z and f_z, respectively. By equating the number of degrees of freedom, it can be shown that these estimates are consistent, see Paper F, and the result of the compensation (4.27) is thus that we have removed all effects of the noise $w(t)$ in $\bar{R}(P)$ and $\bar{f}(P)$. Hence the resulting least-squares estimate

$$\hat{\theta}_P = \hat{R}_z^{-1}\hat{f}_z \tag{4.28}$$

will be consistent regardless of (the color of) $w(t)$.

In Paper F the method (4.27)-(4.28) is referred to as the compensated least-squares (CLS) method. Due to its simplicity and general applicability, this method is a very interesting alternative to other methods that are applicable in the errors-in-variables situation. Note that with the compensated least-squares method no iterations are required to find the estimate – a clear advantage compared to most other errors-in-variables methods, which frequently use singular value decompositions, and to most other time-domain identification schemes, which often use Gauss-Newton type of search algorithms for finding the estimates.

The above idea for bias correction can be combined with prefiltering of the data through a prewhitening filter constructed directly from data. See Paper F.

4.5 Compensation Methods for Total Least-squares Estimation

We will now describe some similar techniques for removing the bias in the total least-squares method when applied to data corrupted by colored input and output noises.

Take $n_a = n_b = n$ and $n_k = 0$, for simplicity, and define

$$T^N = \begin{bmatrix} -Y^N & U^N \end{bmatrix} \tag{4.29}$$

$$Y^N = \begin{bmatrix} y(n+1) & \dots & y(1) \\ \vdots & \ddots & \vdots \\ y(N) & \dots & y(N-n) \end{bmatrix} \tag{4.30}$$

$$U^N = \begin{bmatrix} u(n+1) & \dots & u(1) \\ \vdots & \ddots & \vdots \\ u(N) & \dots & u(N-n) \end{bmatrix} \tag{4.31}$$

$$\theta' = \begin{bmatrix} 1, a_1, \dots, a_n, b_0, \dots, b_n \end{bmatrix}^T \tag{4.32}$$

Denote the noisy variant of T^N by T_m^N. The total least-squares solution $\hat{\theta}'_{TLS}$ can then be stated as

$$T_{TLS}^N = \arg\min_T \|T_m^N - T\|_F^2 \tag{4.33}$$

subject to

$$T_{TLS}^N \hat{\theta}'_{TLS} = 0 \tag{4.34}$$

The solution can easily be calculated using a singular value decomposition of T_m^N [35].

Introduce the notation

$$R'_m(N) = \frac{1}{N}(T_m^N)^T T_m^N \tag{4.35}$$

$$R'_z = \bar{E}\varphi'_z(t)(\varphi'_z(t))^T \tag{4.36}$$

$$\varphi'_z(t) = \begin{bmatrix} -y(t), \dots, -y(t-n), u(t), \dots, u(t-n) \end{bmatrix}^T \tag{4.37}$$

$$R'_w = \bar{E}\varphi'_w(t)(\varphi'_w(t))^T \tag{4.38}$$

$$\varphi'_w(t) = \begin{bmatrix} -w_y(t), \dots, -w_y(t-n), w_u(t), \dots, w_u(t-n) \end{bmatrix}^T \tag{4.39}$$

The solution to the total least-squares problem is given by the right null space of T_m^N or, alternatively, by the null space of $R'_m(N)$. Since $z(t)$ and $w(t)$ are uncorrelated we have that

$$R'_m(N) \to R'_z + R'_w, \quad \text{as } N \to \infty \tag{4.40}$$

If R'_w is the identity matrix it will not affect the directions of the singular vectors of R'_z and the total least-squares estimate will be consistent.

If we use periodic data we can estimate R'_w and compensate for any color in $w(t)$. Let $\hat{R}'_w$ be the nonparametric estimate of R'_w computed from (4.8). One possibility is then to subtract $\hat{R}'_w$ from $R'_m(N)$, similarly to what we did in the previous section. Another idea is to scale $R'_m(N)$ by the inverse of $\hat{R}'_w$:

$$(\hat{R}'_w)^{-T/2} R'_m(N)(\hat{R}'_w)^{-1/2} \tag{4.41}$$

In this way the term R'_w in (4.40) will be replaced by the identity matrix and consistency can be restored. This can be seen as a way to whiten the noise $w(t)$ which also can be done more directly in the frequency domain using the nonparametric noise model (4.9). Obviously we can also combine the other approaches above with prefiltering. The benefits of doing so are discussed in Paper F.

Chapter 5

Identification for Control

What is a good model for control? A general answer is that the model should be such that the controller design based on this model gives robust/high performance when applied to the real system. Different design methods impose different constraints on the model. The challenge in identification for control is to understand what constraints are imposed by different design methods and to device identification methods that deliver models meeting these constraints. This is the topic of this and the next chapter. This chapter is focused on the model demands imposed by the design method and on iterative schemes for identification for control. Here we will also discuss how to assess the quality of a given model. The next chapter deals with optimal experiment design aimed at minimizing the performance degradation due to model errors. A new approach to identification for control is also proposed.

5.1 Models and Robust Control

Introduction

As is well known from the control literature, a wealth of linear control design methods exist, ranging from very simple PID tuning rules (e.g., [126]) to μ-synthesis (e.g., [124]). Consequently the model demands can be very different depending on what type of design method is considered. Take the Ziegler-Nichols rule for PI tuning, for example. Here the controller parameters are tuned using an estimate of one single point on the Nyquist curve: at the phase cross-over frequency. At the other extreme we have μ-synthesis with structured (parametric) uncertainties.

Here it is assumed that a nominal model and a detailed uncertainty description, including both structural information and hard (norm) bounds on the uncertainties, are given a priori.

Although the uncertainty description may not be used explicitly in the design, as in μ-synthesis, we always need some knowledge of the quality of the model in order to evaluate the control performance before applying the new controller to the real plant, so called controller validation. Sometimes this is handled by using prior information/assumption, for instance expressed as: "The amplitude of the system's frequency response is monotonically decreasing so we can achieve good control using a PI regulator", but in most cases we need some mathematical description of the uncertainty to do the controller validation. We therefore need a suitable format for representing the model plus its uncertainties. Here we will mostly deal with unstructured uncertainty descriptions, like the following additive one

$$G_\Delta = G_\Delta(G, W) = G + W \cdot \Delta, \quad \|\Delta\|_\infty < 1 \tag{5.1}$$

where G denotes the nominal model and W represents the uncertainty bound, see below. (Here we have omitted the arguments $e^{i\omega}$ and ω which also frequently will be done in the sequel for notational simplicity.) Other variants, like multiplicative uncertainty descriptions, are also possible, but in the scalar case these are essentially equivalent to (5.1). Therefore we will only consider this format in the sequel. An important observation is that (5.1) really is a model set description representing a whole family of models centered around the nominal model G.

What are the demands on the nominal model?

As already mentioned, the requirements on the model will strongly depend on the chosen design method. To limit the scope somewhat we will here focus on model based approaches, like certainty equivalence control (e.g., LQG, model reference, and pole placement) and robust control (e.g., H_∞ and μ). In both these approaches the model is used as a tool for designing high-performance controllers and any plant/model mismatch will limit the achievable performance. Other limiting factors include nonminimum phase zeros and poles and nonlinear effects in the plant (cf. [100]). However, we have little/no control over these so what we can do is to minimize the performance degradation due to a poor model fit. This does not imply that the model has to be equally good in all frequency ranges, though. The model fit needs only to be good in frequency ranges where this is important. (Notice the difference!) The problem is of course to know what frequency ranges are important. Consider, for example, a case where the controller for some reason is required to have a built-in integral action. In such a case the model fit at low frequencies is less important. It is also easy to construct examples where two models with seemingly identical Nyquist curves give rise to controllers with vastly different control performance (see, e.g., [96, 97, 102]). What typically has been (purposely) ignored in these examples is the importance of a good model fit around the cross-over frequency which roughly coincides with the closed-loop bandwidth. This can be said to be part of the folk wisdom in control and a topic that has been much studied in the literature. To mention just one reference, see [109]. That the model has to be good around the cross-over frequency is a quite straightforward message that applies to a wide range of design methods, although not all. An exception is pole placement where it is not obvious that a poor model fit in this frequency region automatically leads to poor control designs. It might also be that the achievable bandwidth of the system is unknown a priori. In such cases the situation is less clear.

What are the demands on the uncertainty description?

If we consider robust control, a general answer is that the uncertainty description should be reliable, nonconservative, and lend itself easily to robust control design. If we concentrate on scalar, unstructured uncertainty descriptions, the standard format (5.1) clearly satisfies these requirements provided that the frequency weighting W is chosen properly. Note that in (5.1) the uncertainty is bounded using the H_∞-norm (in the scalar case $\|\cdot\|_\infty = \sup_\omega |\cdot|$). This means that the statement/assumption $G_0 \in G_\Delta$ with G_Δ given by (5.1) is equivalent to saying that

$$|G_0 - G| < |W|, \ \forall\ \omega \tag{5.2}$$

This shows that the key issue is that $|W|$ constitutes a reliable error bound for the model and that, by using the H_∞-norm, any phase information provided about W is obsolete. For robust control (H_∞ or μ), however, W still needs to be represented as a (stable) transfer function. It is then important that W is expressed using a low order filter since the order of the resulting controller typically will be equal to the sum of the orders of G and W (at least).

Summary

To summarize the discussion so far, we can draw the following general conclusions:

1. The model format needs to be well suited for control design. Robust control requires both a nominal model and an uncertainty description.

2. The model fit must be good in certain frequency ranges – typically around the cross-over frequency.

3. It is often sufficient if the uncertainty description is given in terms of an upper bound on the size of the model uncertainty. The bound should be tight and be given in a suitable norm.

In the next section we will further discuss methods for uncertainty bounding and the role of model validation. After that, we will turn to the question of how to obtain a customized model fit.

5.2 Uncertainty Bounding and Model Validation

In the stochastic framework we have adopted in this thesis the disturbances are assumed to be realizations of stochastic processes with certain (known) properties. As mentioned in Chapter 1, this implies that it is impossible to construct uncertainty regions for the identified model that hold with certainty; we can only guarantee that they will hold with some level of probability. This is a problem since most robust control methods assume that hard uncertainty bounds are given a priori. However, as is vividly described in [119], given only measured data we can never be 100% positive about the uncertainty bounds. We can merely say that there is no evidence in the data that the true system does not lie within the bounds. Under these premises, any control engineer should therefore, in our opinion, be able to accept an uncertainty bound that hold with 99.9% certainty, say, although the design method requires hard bounds.

There exist several, more or less involved, algorithms for constructing probabilistic uncertainty bounds from data, see for example [44, 86]. A crucial issue is the so called *bias/variance trade-off*, that is, finding a good compromise between minimizing the bias error and the variance error. Recall, from Chapter 1, that we can only use estimated covariance information to compute reliable confidence regions for the identified model if the bias error is negligible. If we are allowed to increase the model order until the validation tests do not indicate any systematic error due to undermodeling this is no problem, as is illustrated in Paper H. However, with bounds on the model order we have to be more careful. Both [44] and [86] discuss methods in which the bias error due to undermodeling is explicitly estimated and accounted for when constructing the uncertainty regions. In the iterative identification for control schemes that will be discussed in the next section this question is avoided completely by simply postulating that all model errors are due to undermodeling. However, this is unrealistic since in practice all identified model will be corrupted by both bias errors due to undermodeling and variance errors due to noise. The key to good results is to find the model which gives the smallest total error.

The above discussion illustrates the importance of the model validation procedure for successful uncertainty bounding. In fact, the construction of the uncertainty bounds serves itself as a model validation tool, since a model whose uncertainty bounds are excessively large can hardly be said to be useful in any application. (This of course calls for a nonconservative method for constructing the bounds.) The problem with most traditional validation tools, like residual analysis, is that they deliver the results in a format that is hard to incorporate in the control design; recall that for robust control design it is important to have frequency domain descriptions of the model uncertainty. In order to meet these requirements we can consider using the concept "model error modeling" [69, 70] where the idea is to fit a model to input-residual data and to compute reliable confidence regions for this "model error model" in the frequency domain. If these confidence regions contain zero for all frequencies there is no evidence in the current data that the data was not generated by the original model. Any systematic error due to undermodeling will also be detected by this method. It should be noted that model error modeling is suitable if a fixed model structure is used to identify the system. Without this constraint it may be better to re-identify the system using a higher-order model, with small bias, and to compute the confidence regions for this models. In fact, this is what we recommend for most practical problems. An exception is when a low-order, linear model is sought that should approximate a given process with complex dynamics, as well as possible. In this case one could use a more flexible structure for the model error model than the original model. In [26] this idea is used for the modeling of a highly nonlinear plant using a combination of a low-order linear regression model and neural network model.

5.3 Iterative Approaches to Identification for Control

A number of iterative approaches to identification for control have been suggested in the literature. A major part of them assume a fixed, and typically low order, model structure. Since low order models are used, the focus has been on how to shape the bias error to do the least harm in the subsequent design stage. The main idea has been to try to match the identification and control criteria to obtain a control-relevant model fit and the key tool for achieving this has been the use of closed-loop experiments.

Suppose we have a prechosen control objective function $F(G_0, K)$ in some Banach (normed) space. The design objective is to minimize the control performance cost measured by $\|F(G_0, K)\|$ with respect to K, that is

$$\min_K \|F(G_0, K)\| \tag{5.3}$$

The objective function $F(G_0, K)$ can be, for example, a weighted sensitivity function and the norm is typically the H_2-norm or the H_∞-norm. In practice a controller is found through minimization of $\|F(G,K)\|$ where G is a model. For a given model G and a given controller K we can bound $\|F(G_0,K)\|$ using the following inequalities (cf. [96, 97])

$$\begin{aligned} \Big| \|F(G,K)\| - \|F(G_0,K) - F(G,K)\| \Big| &\le \\ \le \|F(G_0,K)\| &\le \\ \le \|F(G,K)\| + \|F(G_0,K) - F(G,K)\| & \end{aligned} \tag{5.4}$$

From (5.4) it is clear that $\|F(G_0,K)\|$ is small if

(i) The designed performance cost $\|F(G,K)\|$ is small, and

(ii) $\|F(G_0,K) - F(G,K)\| \ll \|F(G,K)\|$.

These conditions can be interpreted as demands for high nominal performance and robust performance, respectively. The nominal performance cost $\|F(G,K)\|$ is minimized in the control design. For robust performance we should minimize $\|F(G_0,K) - F(G,K)\|$ which, for a fixed K, is an identification problem.

To minimize $\|F(G_0,K)\|$ it has thus been suggested to iterate on

$$G_{i+1} = \arg\min_G \|F(G_0,K_i) - F(G,K_i)\| \tag{5.5a}$$

$$K_{i+1} = \arg\min_K \|F(G_{i+1},K)\| \tag{5.5b}$$

A number of observations can be made:

1. The criterion used in the identification step (5.5a) is dictated by the control criterion, that is, by F and the chosen norm. The plant/model mismatch is thus minimized in a control-relevant way (for the current controller K_i).
2. The control problem (5.5b) is handled using standard design tools. To solve (5.5a) we may have to construct a customized identification method which can be highly nontrivial depending on the chosen norm and the control objective function F.
3. With the H_2-norm it is possible to tune the prediction error method (3.16), (3.17) to approximate (5.5a) for a wide range of control objective functions $F(G_0,K)$ using experimental data from the current closed-loop system (G_0,K_i). The approximation will become exact as the number of data points tends to infinity.
4. If we succeed in satisfying $\|F(G_0,K_i) - F(G_i,K_i)\| \ll \|F(G_i,K_i)\|$ we have tight upper and lower bounds on the achieved performance $\|F(G_0,K_i)\|$. If we fail we have to reconsider either the identification step or the design step or both.

5. As we have formulated the problem here the control objective function F is fixed a priori. In practice it may be desirable to adjust this as we learn more and more about the achievable performance. This is also standard in control synthesis problems. Some of the proposed iterative schemes include this feature.

5.4 Examples

As mentioned, most iterative schemes focus on shaping the bias distribution in a control-relevant way. A key feature has been the use of closed-loop experiments and the methods used have often been different variants of the indirect method. Here we will go through a number of examples and show that with the indirect method it is actually possible to match the identification and control criteria perfectly for a wide range of control design paradigms.

Consider the indirect method applied with the closed-loop model structure

$$y(t) = \frac{G(q,\theta)}{1+G(q,\theta)K(q)}r(t) + H_*(q)e(t) \tag{5.6}$$

where $G(q,\theta)$ is the open-loop model and $H_*(q)$ a fixed noise model (prefilter). An expression for the limit model, as the number of measurements tends to infinity was given in (3.59). A convenient, but not formally correct, interpretation of this result is that the estimate $\hat{G}_N(e^{i\omega})$ converges to

$$G^*(e^{i\omega}) = \arg \min_{G(e^{i\omega},\theta)\in\mathcal{G}} \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{G_0(e^{i\omega}) - G(e^{i\omega},\theta)}{1+G(e^{i\omega},\theta)K(e^{i\omega})}\right|^2 \frac{|S_0(e^{i\omega})|^2\Phi_r(\omega)}{|H_*(e^{i\omega})|^2}d\omega \tag{5.7}$$

as N tends to infinity. By choosing the design variables Φ_r and H_* properly, this expression is exactly equal to the criterion (5.5a) for almost all interesting control criteria (5.5b) expressed in the H_2-norm. In the following we will illustrate this on a number of examples taken from the literature. As a result we will obtain a unifying framework for most existing iterative schemes.

Model Reference Control/Pole Placement/Internal Model Control

One possible approach is to use model reference control (MRC) or the related pole placement design method, see [5] for instance. In model reference control the desired closed-loop system is specified as $y = G_d r$ and a 1-degree-of-freedom controller K ($u = K(r-y)$) is computed from

$$G_d = \frac{GK}{1+GK} \tag{5.8}$$

where G is the model. (Here we have neglected the arguments which will also be done in the sequel for notational simplicity.) The control objective function $F(G_0, K)$ is

$$F(G_0, K) = \frac{G_0K}{1+G_0K} - G_d \tag{5.9}$$

If we measure the deviation with the H_2-norm the identification problem (5.5a) at iteration i becomes

$$\begin{aligned} G_i &= \arg\min_G \|F(G_0, K_i) - F(G, K_i)\|_2^2 \\ &= \arg\min_G \left\| \frac{G_0K_i}{1+G_0K_i} - \frac{GK_i}{1+GK_i} \right\|_2^2 \\ &= \arg\min_G \left\| \frac{(G_0 - G)K_i}{(1+G_0K_i)(1+GK_i)} \right\|_2^2 \\ &= \arg\min_G \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{G_0 - G}{1+GK_i} \right|^2 |S_0|^2 |K_i|^2 d\omega \end{aligned} \tag{5.10}$$

Comparing this with expression (5.7) we see that the proper design choices are a filter H_* and a reference spectrum Φ_r such that $\Phi_r/|H_*|^2 = |K_i|^2$.

A variant of this approach is used in the iterative scheme known as the windsurfer approach to adaptive robust control [4, 62, 63, 64, 65, 66]. Here the control design is based on the so called internal model control (IMC) design paradigm which is closely related to model reference control. In IMC the desired closed-loop system is parameterized using a low-pass filter whose bandwidth is increased gradually as the model is refined experimentally. In each iteration a similar identification problem as in (5.10) is solved.

Basically the same idea has also been utilized in a number of papers by Keviczky and Banyasz [9, 56, 57, 58, 59]. They use a modified internal model principle which is claimed to simplify the calculations and simplify a good tuning of the controller parameters. However, the necessary re-parameterizations of the different controller blocks are not trivial.

Optimization Based Methods

The second main approach has been to consider some optimization based control paradigm, for instance LQG as in [41, 43]. LQG has also been used in the so called

Zangscheme [121, 122, 123], see also [12, 89]. The design problem is

$$\min_K \|F(G_0, K)\|_2^2 = \min_K \lim_{N\to\infty} \sum_{t=1}^{N} [(y(t) - r(t))^2 + \lambda^2 u^2(t)] \tag{5.11}$$

We are thus lead to the following identification problem for a given controller K_i

$$\begin{aligned} \min_G \|F(G_0, K_i) - F(G, K_i)\|_2^2 = \min_G \lim_{N\to\infty} \sum_{t=1}^{N} [(y(t) - \frac{GK_i}{1+GK_i} r(t))^2 \\ + \lambda^2 (u(t) - \frac{K_i}{1+GK_i} r(t))^2] \end{aligned} \tag{5.12}$$

Now, it can be shown (cf. [121]) that

$$\begin{aligned} \|F(G_0, K_i) - F(G, K_i)\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{G_0 - G}{1+GK_i} \right|^2 |S_0|^2 (1 + \lambda^2 |K_i|^2) |K_i|^2 \Phi_r d\omega \\ + \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_0|^2 (1 + \lambda^2 |K_i|^2) \Phi_v d\omega \end{aligned} \tag{5.13}$$

A comparison with expression (5.7) tells us that in this case the noise model H_* should be chosen such that $1/|H_*|^2 = (1 + \lambda^2 |K_i|^2)|K_i|^2$.

The coprime factor identification scheme (e.g., [113]) has been motivated by considering the following control problem (H_∞-loop shaping, e.g., [80]):

$$\min_K \|F(G_0, K)\|_\infty \tag{5.14}$$

where now

$$F(G_0, K) = \begin{bmatrix} G_0 \\ 1 \end{bmatrix} (1 + G_0 K)^{-1} \begin{bmatrix} 1 & K \end{bmatrix} \tag{5.15}$$

If we consider the iterative identification and control scheme (5.5) we should for robust performance try to minimize

$$\|F(G_0, K_i) - F(G, K_i)\|_\infty \tag{5.16}$$

It is shown in [113] how $F(G_0, K_i) - F(G, K_i)$ can be expressed using coprime factors which, when replacing the H_∞-norm by the H_2-norm, leads to the coprime factor identification scheme. See also [96, 97, 98, 99]. Here we will choose a different and more straight-forward route: We have

$$\begin{aligned} \|F(G_0, K_i) - F(G, K_i)\|_2^2 &= \left\| \left\{ \begin{bmatrix} G_0 \\ 1 \end{bmatrix} \frac{1}{1+G_0K_i} - \begin{bmatrix} G \\ 1 \end{bmatrix} \frac{1}{1+GK_i} \right\} \begin{bmatrix} 1 & K_i \end{bmatrix} \right\|_2^2 \\ &= \left\| \begin{bmatrix} 1 \\ -K_i \end{bmatrix} \frac{G_0 - G}{(1+G_0K_i)(1+GK_i)} \begin{bmatrix} 1 & K_i \end{bmatrix} \right\|_2^2 \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{G_0 - G}{1+GK_i} \right|^2 |S_0|^2 (1 + |K_i|^2)^2 d\omega \end{aligned} \tag{5.17}$$

With the indirect method, the proper design choices becomes to tune H_* and Φ_r so that $\Phi_r/|H_*|^2 = (1 + |K_i|^2)^2$.

The iterative scheme just presented involving H_∞-loop shaping has recently been further developed to include robustness considerations, see [16, 18, 19]. Here the problem formulation is slightly more involved and the control design is done using μ-synthesis, but the same identification problem is considered.

5.5 Evaluation of the Iterative Schemes

As we just have seen, several of the existing iterative schemes can be seen as variants of a single, "generic" scheme using the indirect method for the identification step. Therefore the comments given below apply to all of the methods mentioned above.

A number of problems with iterative designs have been noted and we will try to high-light some of them here. The discussion will be focused on the following three questions:

1. Is it relevant to consider a fixed, low-order model class and to concentrate on the bias error, neglecting the variance error?
2. Is it optimal to perform identification experiments with the current controller operating in the loop to obtain more relevant information about the system?
3. Will this type of iterative schemes converge and, in that case, what will the convergence points be?

Fixed Model Structures

A typical identification session involves collecting the data, estimating a model, and finally checking the quality of the model (model validation). By now, very efficient software for performing the last two steps exists, so the estimation and validation step can be performed at virtually no cost given the experimental data. To find a suitable model one often estimates a number of models of different types (ARX, ARMAX, etc.) and of different orders. The final choice is then made using model validation techniques, such as residual analysis and comparison of measured and simulated outputs using "fresh" data (cross-validation). For control purposes it is important to ensure that the model fit is good enough around the cross-over frequency, as we have seen. The validation step must hence also include some verification of this before the model can be accepted. Typically, one starts by checking if the model of lowest order is sufficiently good. The model order is then increased until a model that passes the validation tests is found. As a general rule of thumb we then know that the variance error will dominate the bias error for this model [71]. If the resulting model order is too high we can consider model reduction or repeating the estimation phase using some suitable prefilter (cf. [11]).

With this as a background we think that it is questionable to handle the identification step as is typically done in the identification-for-control-literature, namely to fix the model structure a priori and to validate the model by evaluating the control performance achieved with the new controller designed using the identified model. It is true that if we use a fixed, low model order the bias error will probably be the dominating source of error. Therefore it is also reasonable to tune the identification criterion to match the control ditto, as described above. However, if the model is biased, the existing validation tools will give information about the modeling error that is hard to incorporate in the control design, or they will fail completely. Thus it can be hard/impossible to ensure that the best model is also good enough which really is the key question.

In the iterative schemes the model is validated through the new controller it gives rise to. The main problem with this is that, instead of learning more about the true system, we basically only get information of what model suits the current controller best. This model can be very different from the true system and from the model that results in the best controller in the subsequent design step! The bottom line is that fixing the model order a priori limits the chances of a successful identification since we are then not utilizing the full potential of system identification as a modeling tool and the model validation becomes difficult to perform in a reasonable way.

Maximally Informative Experiments

The core of the identification-for-control-problem is to device an identification procedure that gives a model which results in the best possible controller for the given plant. This includes designing the identification experiments. A key feature of the existing schemes has been the use of closed-loop experiments and the main argument has been that one then gets a control-relevant frequency weighting with the sensitivity function “for free”, see expression (5.7). Again, this applies if we neglect the variance error or, alternatively, assume infinite data records, so that the model error is only due to undermodeling. However, any frequency weighting that is possible with closed-loop experiments is also possible with open-loop experiments provided the prefilter is chosen properly. (To achieve a perfect match we may have to know the true system. However, in practice we can never achieve a perfect match with the indirect method either, so this drawback of the direct method is not unique.)

Closed-loop experiments are only optimal when the experiment design has to be done with constraints on the output variance, confer Section 6.1, or when the misfit in both the dynamics model and the noise model is penalized, see, e.g., [34, 71] and Paper G. Another comment on the iterative schemes is that they use the current closed-loop to generate new data for the identification step in the next iteration. We question this for two reasons. First, updating the model based on data from only the current loop implies that we discard information gathered during the preceding experiments on the system. If the system is time-invariant this is clearly suboptimal. Secondly, we have already mentioned that collecting data with the current controller operating in the loop may not give the proper information about the system necessary to improve the model quality. In fact, this can be completely opposite to the optimal experiment design especially if the current controller is reasonably good, achieving good disturbance rejection in the closed-loop. In this case, the power that we put in the reference signal to probe the system may be filtered out by the controller so that the actual input seen by the system is very smooth and brings us very little useful information about the true system. This shows that good identification experiments and good control action are conflicting requirements that counteracts each other. This problem has been studied in detail in dual control (see, e.g., [8, 21]), but so far no practically feasible solution has been found. In the identification for control literature this issue has largely been neglected but, as the above discussion shows, this is definitely one of the key issues that needs to be addressed in the future.

Convergence Analysis

In this section we will present some small steps towards a complete analysis of the convergence properties of the iterative schemes. As mentioned, it has been common to use a fixed (low-order) model structure in the identification step (5.5a). This is also standard in adaptive control which is the same paradigm as iterative design. If we introduce the notation

$$J(G,K) = \|F(G_0,K) - F(G,K)\|, \quad J'_G(G,K) = \frac{\partial}{\partial G}J(G,K), \quad K'_G = \frac{\partial}{\partial G}K(G) \tag{5.18}$$

the basic adaptive algorithm will be (with a somewhat symbolic notation)

$$G_{i+1} = G_i + \mu J'_G(G_i, K(G_i)) \tag{5.19}$$

The actual convergence analyses for the adaptive and the iterative methods (5.19) and (5.5) are not easy in general and here we shall just comment on possible convergence points, the fix points of the schemes. The analysis below has its roots in Section 7.3.2 of [77]. Similar results have been proven in [50, 70] for the identification-for-control-application.

A fix point for (5.19) is a model G^* such that

$$0 = J'_G(G^*, K(G^*)) \tag{5.20}$$

However, note that K in $J(G, K)$ depends on G: $K = K(G)$. The best models for control are therefore those that minimize $J(G, K(G))$ with respect to G, that is, a model G^* such that

$$0 = J'_G(G^*, K(G^*)) + J'_K(G^*, K(G^*))K'_G(G^*) \tag{5.21}$$

But (5.21) is not the same as (5.20). They are equal if, and only if, $J'_K(G^*, K(G^*)) = 0$. This means that the criterion $J(G, K)$ must not depend on the controller $K = K(G)$ at G^*, which in turn implies that the model G^* must be essentially correct, that is $G^* \approx G_0$.

Summary

The three issues discussed above illustrate some of the problems associated with the iterative schemes. We will here collect some of the main issues in this section.

1. In the literature it has been common to concentrate on the bias error with the motivation that only low-order models are used. However, in practice we can never neglect the variance error, especially without a priori bounds on the model order.
2. If we concentrate on low-order modeling and use the new controller as a validation tool we are not utilizing the full power of system identification. Before performing a new experiment we should exhaust all possibilities of obtaining a good model through repeated estimation and prefiltering, etc.
3. In connection to the last remark we should also notice the requirement $\|F(G_0, K_i) - F(G_i, K_i)\| \ll \|F(G_i, K_i)\|$, which is a kind of robustness criterion (cf. (5.4)). With a fixed and too low model order we might not be able to satisfy this at any iteration.
4. In the last section we saw that, for the possible convergence points to be equal to the desired ones, it was necessary that the model was essentially correct. On the other hand, a standing assumption in many of the iterative schemes is that the chosen model order is too low to admit a perfect match. Under these circumstances it is thus unlikely that a convergence proof can be given for the iterative schemes.
5. Collecting experimental data from the current closed-loop may be sub-optimal for obtaining informative data. Only using the current loop can actually be detrimental for the identification results since the closed-loop signals may have little information contents, especially if the current controller is good.

5.6 Other Approaches and Extensions

The above evaluation applies to iterative schemes of the same form as (5.5). However, many alternative iterative approaches have also been suggested.

An approach aiming at improved robust performance through iterative modeling and control design can be found in [10]. Here a joint criterion for identification and control is formulated. This leads to a difficult optimization problem and to obtain a feasible algorithm basically the same steps as in (5.5) is used. A restriction is that noise-free data is assumed.

In [38] an iterative scheme is proposed in which robust performance is monitored at each iteration. The scheme uses stochastic embedding for uncertainty quantification and an IMC design method. Robust performance monitoring is also a key issue in [16, 18, 19] although from the identification point of view this is the same paradigm as (5.5).

Several authors have also studied the possibility to construct control-relevant prefilters to apply to the process data, see, for instance, [5, 52, 92, 101]. Note that by iteratively updating the prefilter no new experiments are necessary if the first data set is informative enough (and the signal-to-noise-ratio is high).

The IFT method for direct controller tuning derived by Hjalmarsson and co-workers (e.g., [49]) avoids the modeling step completely. Instead the controller parameters are tuned directly using iterated closed-loop experiments.

Another, rather obvious, approach is to first identify the system using a high-order model and then to reduce this model to lower order using some suitable frequency weighting. A nice example, comparing this approach and direct, low-order identification is given in [11]. The dual of the model reduction approach is to apply the reduction techniques to the controller which may be constructed using the high-order model. Several such methods are described in [124].

In a sense all iterative schemes can be viewed as (batch) adaptive control algorithms where the controller is updated using a batch of data instead of at each sampling instant. This also means that all adaptive control algorithms can be seen as iterative schemes, but we will not elaborate on that any further here.

Chapter 6

Optimal Experiment Design

In this chapter we will continue the discussion on how to identify models that are good for control. We will briefly quote some results on optimal experiment design and propose a new method for identification for control.

6.1 Minimizing the Performance Degradation Due to Variance Errors

If the experimental conditions can be chosen freely it is wise to choose them to optimize the quality of the identified model. In this section we will consider experiment design problems where the goal is to minimize a weighted integral of the variance of the estimate $\hat{G}_N(e^{i\omega})$:

$$J = \int_{-\pi}^{\pi} \operatorname{Var} \hat{G}_N(e^{i\omega}) C(\omega) d\omega \tag{6.1}$$

Here $C(\omega)$ is a weighting function that should reflect the relative importance of a good model fit for different frequencies.

The minimization shall be carried out with respect to the design variables $K(e^{i\omega})$ (the regulator) and $\Phi_r(\omega)$ (the reference spectrum). To obtain a more explicit problem formulation we replace $\operatorname{Var} \hat{G}_N(e^{i\omega})$ by the asymptotic expression (3.56):

$$\operatorname{Var} \hat{G}_N(e^{i\omega}) \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u^r(\omega)} \tag{6.2}$$

The above problem formulation is studied in Paper G and here we will briefly state the main results. Other references dealing with similar problems include [34, 48].

The problem that is considered is to minimize J given by (6.1), (6.2) with respect to $K(e^{i\omega})$ and $\Phi_r(\omega)$. The minimization is to be carried out under the constraint

$$\int_{-\pi}^{\pi} \alpha\Phi_u(\omega) + \beta\Phi_y(\omega) d\omega \leq 1 \tag{6.3}$$

which represents a bound on a linear combination of the variances of the input and the output.

The solution is to select the regulator $u(t) = -K(q)y(t)$ that solves the LQG problem

$$\min_K \alpha E u^2(t) + \beta E y^2(t), \quad y(t) = G_0(q)u(t) + H_0(q)e(t) \tag{6.4}$$

The reference signal spectrum shall be chosen as

$$\Phi_r^{opt}(\omega) = \mu\sqrt{\Phi_v(\omega)C(\omega)}\frac{|1 + G_0(e^{i\omega})K(e^{i\omega})|^2}{\sqrt{\alpha + \beta|G_0(e^{i\omega})|^2}} \tag{6.5}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} \alpha\Phi_u(\omega) + \beta\Phi_y(\omega) d\omega = 1 \tag{6.6}$$

This is shown in Paper G where also some alternative problem formulations are considered.

6.2 A Novel Approach to Identification for Control

Introduction

With the discussion in this and the previous chapter as a background, we will in this section try to point out a new direction for research on identification for control. Both system identification of linear models and linear control design are well established fields with many efficient software tools available. Therefore, we will put emphasis on other aspects of this problem, like experiment design and model validation which we think have been little treated in the identification-for-control-literature in the past.

We would like to avoid unnecessary simplifying assumptions and instead try to approach the problem in an unprejudiced way. Thus, for the identification of a nominal model we would like to do as well as possible with the existing tools, without a priori constraints on the model order and without neglecting either bias or variance errors. The uncertainty bounding should be treated similarly using control-relevant model validation tests. The results will to a large part depend on the experiment design, which consequently has to be done with care. A problem with the experiment design problem formulation is that the resulting controller will strongly depend on the choices made in the experiment design but that the link is typically neither explicit nor simple. To get realistic experiment designs we will have to impose some constraints, for instance, an upper bound on the number of data N or bounds on the input and/or output power, as in the preceding section.

This is an outline of the main features of an ultimate identification for control scheme where we have emphasized the importance of a good experiment design and put less weight on the estimation of the models.

The inseparability of the control and modeling problems has been noted previously in the literature, see, among many references, [7, 96, 102]. Often the conclusion has been that an iterative solution is required. The alternative would be to treat these problems jointly, as was suggested in [10]. A problem with such an approach is the rather complicated nature of the optimization problem one then has to deal with. With an iterative scheme one can use simpler techniques for solving the joint problem, but one has to be careful in order to obtain a convergent scheme. In this section we will try to formulate a realistic goal for an identification-for-control-scheme based on a joint approach.

Problem Formulation

We will assume that we from the identification (and uncertainty bounding) can compute a model set description of the form (cf. (5.1) and (5.2))

$$G_\Delta = G + W \cdot \Delta, \quad \|\Delta\|_\infty < 1 \tag{6.7}$$

such that $G_0 \in G_\Delta$ with certainty. X will denote the set of all admissible experiment designs and a particular design will be denoted $\mathcal{X}$; if it is admissible we will write $\mathcal{X} \in X$. Clearly, since the nominal model G and the uncertainty bound W are estimated/computed from data, they will both depend on $\mathcal{X}$, that is, $G = G(\mathcal{X})$ and $W = W(\mathcal{X})$.

The robust control criterion will be to assure robust stability and that

$$\left\| \frac{W_S}{1 + G_\Delta K} \right\|_\infty \leq 1 \tag{6.8}$$

where W_S specifies the performance requirements. This is a robust performance problem that can be solved using μ-synthesis [124]. In the scalar case a necessary and sufficient condition for robust performance is (cf. [84])

$$\left|\frac{WK}{1+GK}\right| + \left|\frac{W_S}{1+GK}\right| < 1, \ \forall\omega \tag{6.9}$$

From this condition it is clear that the robust performance condition is satisfied if both the robust stability and the nominal performance requirements are satisfied with some margin.

In this framework, the identification and control problems can be formulated jointly as

Problem 1

$$\min_{\substack{\mathcal{X}\in X\\K}} \sup_{\omega} \left(\left|\frac{W(\mathcal{X})K}{1+G(\mathcal{X})K}\right| + \left|\frac{W_S}{1+G(\mathcal{X})K}\right|\right) \tag{6.10}$$

Here it is assumed that the performance weight W_S is fixed and given a priori. In practice we also would like to optimize over this variable. A simple way to include this feature in the criterion is to reformulate the joint problem as follows.

Problem 2

$$\max_{\substack{\mathcal{X}\in X\\K}} \rho \quad \text{subject to} \quad \sup_{\omega} \left(\left|\frac{W(\mathcal{X})K}{1+G(\mathcal{X})K}\right| + \rho\left|\frac{W_S}{1+G(\mathcal{X})K}\right|\right) \leq 1 \tag{6.11}$$

Clearly, a high value of the constant ρ corresponds to high demands on control performance, and vice versa. (A similar problem formulation has previously been suggested in [10].)

This can be seen as an attempt to formulate the ultimate identification-for-control-problem. Unfortunately it is not clear how to solve such problems in practice. A possible approach would be to use some iterative procedure.

A Relaxation Algorithm for Problem 1

A relaxation algorithm for solving Problem 1 can look as follows. At iteration i we have obtained the model G_i, the uncertainty bound W_i, and the controller K_i. For this triple we can compute the current performance index γ_i:

$$\gamma_i^{-1} = \sup_{\omega} \left(\left|\frac{W_iK_i}{1+G_iK_i}\right| + \left|\frac{W_S}{1+G_iK_i}\right|\right) \tag{6.12}$$

Iteration $i+1$ starts with an experiment design update:

$$\min_{\mathcal{X}\in X} \sup_{\omega} \left(\left| \frac{W(\mathcal{X})K_i}{1+G(\mathcal{X})K_i} \right| + \left| \frac{W_S}{1+G(\mathcal{X})K_i} \right| \right) \triangleq \bar{\gamma}_{i+1}^{-1} \tag{6.13}$$

which results in a new experiment design $\mathcal{X}_{i+1}$, and subsequently in a new model G_{i+1} and a new uncertainty bound W_{i+1}. The updated performance index $\bar{\gamma}_{i+1}$ satisfies $\bar{\gamma}_{i+1} \geq \gamma_i$. (If this was not satisfied we would not accept the new design $\mathcal{X}_{i+1}$ but retain the previous one $\mathcal{X}_i$.)

The iteration is completed with a controller update:

$$\min_{K} \sup_{\omega} (\left| \frac{W_{i+1}K}{1+G_{i+1}K} \right| + \left| \frac{W_S}{1+G_{i+1}K} \right|) \triangleq \gamma_{i+1}^{-1} \tag{6.14}$$

resulting in a new controller K_{i+1} and an updated performance index $\gamma_{i+1} \geq \bar{\gamma}_{i+1} \geq \gamma_i$. (The first inequality holds since the new controller must perform at least as good as the old one, otherwise we would have to reject it.)

Clearly this algorithm yields a monotonically increasing performance index γ. The controller update can be handled using existing design methods but it is not clear how to solve the experiment design problem. A possible remedy for this problem is discussed next.

A Pseudo-relaxation Algorithm for Problem 1

If we make the simplifying assumption that

$$\left| \frac{W_S}{1+G_iK_i} \right| \approx \left| \frac{W_S}{1+G_{i+1}K_i} \right| \tag{6.15}$$

the experiment design step in the above relaxation algorithm can be reduced to

$$\min_{\mathcal{X}\in X} \sup_{\omega} \left| \frac{W(\mathcal{X})K_i}{1+G_iK_i} \right| = \bar{\gamma}_{i+1}^{-1} \tag{6.16}$$

This is still not easy to solve in the framework we consider but, if we replace the supremum norm by a least mean squares measure and assume that W is entirely due to the variance error in the estimated model G, we arrive at the following design problem (cf. Section 6.1)

$$\min_{\Phi_r,\ K} \int_{-\pi}^{\pi} \operatorname{Var} G \cdot \left| \frac{K_i}{1+G_iK_i} \right|^2 d\omega \tag{6.17}$$

The design parameters are the reference spectrum and the controller to be used in the identification experiment. We can also include a constraint of the form (6.3), say, to make the design more realistic.

This would result in a feasible procedure but a problem is that since the supremum norm is replaced by a least mean squares measure we can never assure $\bar{\gamma}_{i+1} \geq \gamma_i$ after the model update step. This must be verified in each iteration and measures have to be taken if this is not satisfied. As an example we may have to consider using more input power or longer experiments. This is basically the same problem as with the iterative scheme (5.5), where we had to ensure $\|F(G_0, K_i) - F(G_i, K_i)\| \ll \|F(G_i, K_i)\|$ in each iteration, but the possible failure of the new method will depend on the experimental conditions rather than on constraints on the model order.

6.3 Summarizing Discussion on Identification for Control

For successful robust control design we need good nominal models plus tight uncertainty bounds. Models obtained using standard identification methods need not be suitable due to a poor model fit in important frequency ranges or due to the uncertainty descriptions that comes with the models. The prediction error method is a well developed method that can be tuned to give customized model fits. Unfortunately it is hard to compute reliable uncertainty bounds for such models, especially if a low-order model is sought that should approximate the true system in some given frequency weighted norm. High-order ARX modeling could then be an alternative and there are several alternative ways to construct such uncertainty bounding/model validation methods. Note, though, that the bounds will be "soft" since the prediction error method assumes stochastic disturbances.

We have advocated that the experiment design and controller synthesis problems should be treated jointly; otherwise it is not clear how to guarantee performance improvements through model and subsequent controller updates. A problem with such an approach is the complexity of the involved optimization problem. Earlier attempts to marry identification and (robust) control include the iterative identification-for-control-schemes. These are, however, questionable for a number of reasons. A key issue is that in these schemes the experiment design is not done properly which implies that the new information that can be gained by doing a new closed-loop experiment, with the current controller operating in the loop, may be obsolete and result in degraded control performance.

Bibliography

[1] H. Akaike. Some problems in the application of the cross-spectral method. In B. Harris, editor, *Spectral Analysis of Time Series*, pages 81–107. John Wiley & Sons, 1967.

[2] B. D. O. Anderson. From Youla-Kucera to identification, adaptive and nonlinear control. *Automatica*, 34:1485–1506, 1998.

[3] B. D. O. Anderson and M. Gevers. Identifiability of linear stochastic systems operating under linear feedback. *Automatica*, 18:195–213, 1982.

[4] B. D. O. Anderson and R. L. Kosut. Adaptive robust control: on-line learning. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 297–298, Brighton, U.K., 1991.

[5] K. J. Åström. Matching criteria for control and identification. In *Proceedings of the 2nd European Control Conference*, pages 248–251, Groningen, The Netherlands, 1993.

[6] K. J. Åström and T. Bohlin. Numerical identification of linear dynamic systems from normal operating records. In *IFAC Symposium on Self-Adaptive Systems*, Teddington, England, 1965. Also in P.H Hammond, ed., Theory of Self-Adaptive Control Systems.

[7] K. J. Åström and B. Wittenmark. Problems of identification and control. *Journal of Mathematical Analysis and Applications*, 34:90–113, 1971.

[8] K. J. Åström and B. Wittenmark. *Adaptive Control*. Addison-Wesley, 2nd edition, 1995.

[9] C. Banyasz and L. Keviczky. Convergence and robustness properties of a generic regulator refinement scheme. In *IFAC Robust Control Design (ROCOND'97)*, pages 455–460, Budapest, Hungary, 1997.

[10] D. S. Bayard, Y. Yam, and E. Mettler. A criterion for joint optimization of identification and robust control. *IEEE Transactions on Automatic Control*, 37:986–991, 1992.

[11] P. Bendotti, B. Codrons, C.-M. Falinower, and M. Gevers. Control oriented low order modelling of a complex PWR plant: a comparision between open loop and closed loop methods. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3390–3395, Tampa, FL, 1998.

[12] R. R. Bitmead, M. Gevers, and V. Wertz. *Adaptive Optimal Control, The Thinking Man's GPC*. Prentice-Hall, 1990.

[13] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Holden-Day, 1981.

[14] C. T. Chou and M. Verhaegen. Subspace algorithms for the identification of multivariable errors-in-variables models. *Automatica*, 33:1857–1869, 1997.

[15] C. T. Chou and M. Verhaegen. Closed-loop identification using canonical variate analysis. Technical report TUD/ET/SCE98.006, Electrical Engineering Dept., Delft University of Technology, 1998.

[16] R. A. de Callafon. *Feedback Oriented Identification for Enhanced and Robust Control*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 1998.

[17] R. A. de Callafon and P. M. J. Van den Hof. Multivariable closed-loop identification: From indirect identification to Dual-Youla parametrization. In *Proceedings of the 35th Conference on Decision and Control*, pages 1397–1402, Kobe, Japan, 1996.

[18] R. A. de Callafon and P. M. J. Van den Hof. Multivariable feedback relevant system identification of a wafer stepper system. In *Proceedings of DETC'97*, Sacramento, CA, 1997. American Society of Mechanical Engineers.

[19] R. A. de Callafon and P. M. J. Van den Hof. Suboptimal feedback control by a scheme of iterative identification and control design. *Mathematical Modelling of Systems*, 3:77–101, 1997.

[20] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, 1983.

[21] A. A. Feldbaum. *Optimal Control Theory*. Academic Press, 1965.

[22] U. Forssell. Asymptotic variance expressions for identified black-box models. Technical Report LiTH-ISY-R-2089, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998. Submitted to Systems & Control Letters.

[23] U. Forssell and C. T. Chou. Efficiency of prediction error and instrumental variable methods for closed-loop identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 1287–1288, Tampa, FL, 1998.

[24] U. Forssell, F. Gustafsson, and T. McKelvey. Time-domain identification of dynamic errors-in-variables systems using periodic excitation signals. In *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999.

[25] U. Forssell and H. Hjalmarsson. Maximum likelihood estimation of models with unstable dynamics and nonminimum phase noise zeros. In *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999.

[26] U. Forssell and P. Lindskog. Combining semi-physical and neural network modeling: An example of its usfulness. In *Proceedings of the 11th IFAC Symposium on System Identification*, volume 4, pages 795–798, Fukuoka, Japan, 1997.

[27] U. Forssell and L. Ljung. Identification for control: Some results on optimal experiment design. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3384–3389, Tampa, FL, 1998. Submitted to Automatica.

[28] U. Forssell and L. Ljung. Identification of unstable systems using output error and Box-Jenkins model structures. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3932–3937, Tampa, FL, 1998. To appear in IEEE Transactions on Automatic Control.

[29] U. Forssell and L. Ljung. Closed-loop identification revisited. *Automatica*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-2021, Linköping University, Linköping, Sweden.

[30] U. Forssell and L. Ljung. A projection method for closed-loop identification. *IEEE Transactions on Automatic Control*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-1984, Linköping University, Linköping, Sweden.

[31] K. F. Gauss. *Theoria Motus Corporum Celestium*. Dover, 1963. English translation: Theory of the Motion of the Heavenly Bodies.

[32] M. Gevers. Towards a joint design of identification and control. In H. L. Trentelman and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 111–151. Birkhäuser, 1993.

[33] M. Gevers. Identification for control. In *Proceedings of the 5th IFAC Symposium on Adaptive Control and Signal Processing*, pages 1–12, Budapest, Hungary, 1995.

[34] M. Gevers and L. Ljung. Optimal experiment design with respect to the intended model application. *Automatica*, 22:543–554, 1986.

[35] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 2nd edition, 1989.

[36] G. C. Goodwin, M. Gevers, and B. Ninness. Quantifying the error in estimated transfer functions with applications to model order selection. *IEEE Transactions on Automatic Control*, AC-37:913–928, 1992.

[37] G. C. Goodwin and M. E. Salgado. A stochastic embedding approach for qualifying uncertainty in the estimation of restricted complexity models. *International Journal of Adaptive Control and Signal Processing*, 3:333–356, 1990.

[38] S. F. Graebe, M. R. West, and G. C. Goodwin. A laboratory scale application of incremental estimation and stochastic embedding. In *Proceedings of the American Control Conference*, pages 3028–3032, San Francisco, CA, 1993.

[39] F. Gustafsson and J. Schoukens. Utilizing periodic excitation in prediction error based system identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3926–3931, Tampa, FL, 1998.

[40] I. Gustavsson, L. Ljung, and T. Söderström. Identification of processes in closed loop — Identifiability and accuracy aspects. *Automatica*, 13:59–75, 1977.

[41] R. G. Hakvoort. Optimal experiment design for prediction error identification in view of feedback design. *Selected Topics in Identification, Modelling and Control*, 2:71–78, 1990. Delft University Press.

[42] R. G. Hakvoort. *System Identification for Robust Process Control*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 1994.

[43] R. G. Hakvoort, R. J. P. Schrama, and P. M. J. Van den Hof. Approximate identification with closed-loop criterion and application to LQG feedback design. *Automatica*, 30:679–690, 1994.

[44] R. G. Hakvoort and P. M. J. Van den Hof. Identification of probabilistic uncertainty regions by explicit evaluation of bias and variance errors. *IEEE Transactions on Automatic Control*, 42:1516–1528, 1997.

[45] F. R. Hansen. *A Fractional Representation to Closed-loop System Identification and Experiment Design*. PhD thesis, Stanford University, Stanford, CA, USA, 1989.

[46] F. R. Hansen, G. F. Franklin, and R. Kosut. Closed-loop identification via the fractional representation: experiment design. In *Proceedings of the American Control Conference*, pages 1422–1427, Pittsburg, PA, 1989.

[47] H. Hjalmarsson. *Aspects on Incomplete Modeling in System Identification*. PhD thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, April 1993.

[48] H. Hjalmarsson, M. Gevers, and F. De Bruyne. For model-based control design, closed loop identification gives better performance. *Automatica*, 32:1659–1673, 1996.

[49] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin. Iterative feedback tuning: Theory and applications. *IEEE Control Systems*, 18:26–41, 1998.

[50] H. Hjalmarsson, S. Gunnarsson, and M. Gevers. Optimality and sub-optimality of iterative identification and control schemes. In *Proceedings of the American Control Conference*, pages 2559–2563, Seattle, 1995.

[51] H. Hjalmarsson and L. Ljung. Estimating model variance in the case of undermodeling. *IEEE Transactions on Automatic Control*, 37:1004–1008, 1992.

[52] K. S. Jun, D. E. Rivera, E. Elisante, and V. E. Sater. A computer-aided design tool for robustness analysis and control-relevant identification of horizon predicitve control with application to a binary distillation column. *Journal of Process Control*, 6:177–186, 1996.

[53] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

[54] R. E. Kalman. Contributions to the theory of optimal control. *Bol. Soc. Mat. Mex.*, 5:102–119, 1960.

[55] R. E. Kalman. When is a control system optimal? *American Society of Mechanical Engineers Transactions Series D: Journal of Basic Engineering*, 86:1–10, 1964.

[56] L. Keviczky. Combined identification and control: Another way. In *Proceedings of the 5th IFAC Symposium on Adaptive Systems in Control and Signal Processing*, pages 13–30, Budapest, Hungary, 1995.

[57] L. Keviczky and C. Banyasz. A new structure to design optimal control systems. In *IFAC Workshop on New Trends in Design of Control Systems*, pages 102–105, Smolenice, Slovak Republik, 1994.

[58] L. Keviczky and C. Banyasz. A new pole-placement regulator design method. In *Proceedings of the 34th IEEE Conference on Decision and Control*, pages 3408–3413, New Orleans, LA, 1995.

[59] L. Keviczky and C. Banyasz. Interplay between robustness versus control and identification performances. In *IFAC Robust Control Design (ROCOND'97)*, pages 333–339, Budapest, Hungary, 1997.

[60] R. L. Kosut and B. D. O. Anderson. Uncertainty model unfalsification. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 1, pages 163–168, San Diego, CA, 1997.

[61] W. E. Larimore. Canonical variate analysis in identification, filtering and adaptive control. In *Proceedings of the 29th IEEE Conference on Decision and Control*, pages 596–604, Honolulu, Hawaii, 1990.

[62] W. S. Lee, B. D. O. Anderson, R. L. Kosut, and I. M. Y. Mareels. On adaptive robust control and control relevant system identification. In *Proceedings of the American Control Conference*, pages 2834–2841, Chicago, IL, 1992.

[63] W. S. Lee, B. D. O. Anderson, R. L. Kosut, and I. M. Y. Mareels. On robust performace improvement through the windsurfer approach to adaptive robust control. In *Proceedings of the 32nd Conference on Decision and Control*, pages 2821–2827, San Antonio, TX, 1993.

[64] W. S. Lee, B. D. O. Anderson, I. M. Y. Mareels, and R. L. Kosut. A new approach to adaptive robust control. *International Journal of Adaptive Control and Signal Processing*, 7:183–211, 1993.

[65] W. S. Lee, B. D. O. Anderson, I. M. Y. Mareels, and R. L. Kosut. On some key issues in the windsurfer approach to adaptive robust control. *Automatica*, 31:1619–1636, 1995.

[66] W. S. Lee, B. D. O. Anderson, I. M. Y. Mareels, and R. L. Kosut. On some practical issues in system identification for the windsurfer approach to adaptive robust control. In M. Blanke and T. Söderström, editors, *Postprint Volume from the 10th IFAC Symposium on System Identification*, volume 1, pages 467–472, Copenhagen, 1995.

[67] L. Ljung. Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, 23:770–783, 1978.

[68] L. Ljung. Asymptotic variance expressions for identified black-box transfer function models. *IEEE Transactions on Automatic Control*, 30:834–844, 1985.

[69] L. Ljung. Identification, model validation, and control. Plenary address at the 36th IEEE Conference on Decision and Control, San Diego, USA, 1997.

[70] L. Ljung. Identification for control – What is there to learn? In *Proceedings of Workshop on Learning, Control and Hybrid Systems*, Bangalore, India, January 1998.

[71] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 2nd edition, 1999.

[72] L. Ljung and P. E. Caines. Asymptotic normality of prediction error estimation for approximate system models. *Stochastics*, 3:29–46, 1979.

[73] L. Ljung and U. Forssell. Variance results for closed-loop identification methods. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 3, pages 2435–2440, San Diego, C.A., 1997.

[74] L. Ljung and U. Forssell. Bias, variance and optimal experiment design: Some comments on closed loop identification. In *Proceedings of the Colloqium on Control Problems in Honor of Prof. I.D. Landau*, Paris, June 1998. Springer Verlag.

[75] L. Ljung and U. Forssell. An alternative motivation for the indirect approach to closed-loop identification. *IEEE Transactions on Automatic Control*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-1989, Linköping University, Linköping, Sweden.

[76] L. Ljung and T. Glad. *Modeling of Dynamic Systems*. Information and System Sciences Series. Prentice-Hall, 1994.

[77] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. MIT Press, 1983.

[78] P. M. Mäkilä, J. R. Partington, and T. K. Gustafsson. Worst-case control-relevant identification. *Automatica*, 31:1799–1819, 1995.

[79] D. McFarlane and K. Glover. *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions*, volume 138 of *Lecture Notes in Control and Information Sciences*. Springer Verlag, 1990.

[80] D. McFarlane and K. Glover. A loop shaping design procedure using H_∞ synthesis. *IEEE Transactions on Automatic Control*, pages 759–769, 1992.

[81] M. Milanese. Properties of least squares estimates in set membership identification. *Automatica*, 31:327–332, 1995.

[82] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set meembership uncertainty: An overview. *Automatica*, 27:997–1009, 1991.

[83] M. Milanese and A. Vicino. Information based complexity and nonparametric worst-case system identification. *Journal of Complexity*, 9:427–446, 1993.

[84] M. Morari and E. Zafiriou. *Robust Process Control*. Prentice-Hall, 1989.

[85] B. Ninness. *Stochastic and Deterministic Modelling*. PhD thesis, The University of Newcastle, Newcastle, NSW, Australia, 1993.

[86] B. Ninness and G. Goodwin. Estimation of model quality. *Automatica*, 31:1771–1797, 1995.

[87] B. Ninness, H. Hjalmarsson, and F. Gustafsson. The fundamental role of general orthonormal bases in system identification. *IEEE Transactions on Automatic Control*, to appear.

[88] J. P. Norton. Identification and application of bounded-parameter models. *Automatica*, 23:497–507, 1987.

[89] A. G. Partanen and R. R. Bitmead. Two stage iterative identification/controller design and direct experimental controller refinement. In *Proceedings of 32nd IEEE Conference on Decision and Control*, pages 2833–2838, San Antonio, Texas, December 1993.

[90] K. Peternell, W. Scherrer, and M. Deistler. Statistical analysis of novel subspace identification methods. *Signal Processing*, 52:161–177, 1996.

[91] K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal. A time-domain approach to model validation. *IEEE Transactions on Automatic Control*, 39:951–959, 1994.

[92] D. E. Rivera, J. F. Pollard, and C. E. Garcia. Control-relevant prefiltering: A systematic design approach and case study. *IEEE Transactions on Automatic Control*, 37:964–974, 1992.

[93] J. Schoukens, R. Pintelon, G. Vandersteen, and P. Guillaume. Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 33:1073–1086, 1997.

[94] J. Schoukens, Y. Rolain, F. Gustafsson, and R. Pintelon. Fast calculations of linear and non-linear least-squares estimates for system identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3408–3410, Tampa, Florida, 1998.

[95] R. J. P. Schrama. An open-loop solution to the approximate closed-loop identification problem. In *Proceedings of the 9th IFAC/IFORS Symposium Identification and System Parameter Estimation*, pages 1602–1607, Budapest, Hungary, 1991.

[96] R. J. P. Schrama. Accurate models for control design: the necessity of an iterative scheme. *IEEE Transactions on Automatic Control*, 37:991–994, 1992.

[97] R. J. P. Schrama. *Approximate Identification and Control Design.* PhD thesis, Delft University of Technology, Delft, The Netherlands, 1992.

[98] R. J. P. Schrama and O. H. Bosgra. Adaptive performance enhancement by iterative identification and control design. *International Journal of Adaptive Control & Signal Processing*, 7:475–487, 1993.

[99] R. J. P. Schrama and P. M. J. Van den Hof. An iterative scheme for identification and control desing based on coprime factorizations. In *Proceedings of the American Control Conference*, pages 2842–2846, Chicago, IL, 1992.

[100] M. M. Seron, J. H. Braslavsky, and G. C. Goodwin. *Fundamental Limitations in Filtering and Control.* Springer-Verlag, 1997.

[101] D. S. Shook, C. Mohtadi, and S. L. Shah. A control-relevant identification strategy for GPC. *IEEE Transactions on Automatic Control*, 37:975–980, 1992.

[102] R. E. Skelton. Model error concepts in control design. *International Journal of Control*, 49:1725–1753, 1989.

[103] R. Smith, G. Dullerud, S. Rangan, and K. Poolla. Model validation for dynamically uncertain systems. *Mathematical Modelling of Systems*, 3(1):43–58, 1997.

[104] R. S. Smith and J. C. Doyle. Model validation: A connection between robust control and identification. *IEEE Transactions on Automatic Control*, 37:942–952, 1992.

[105] T. Söderström, I. Gustavsson, and L. Ljung. Identifiability conditions for linear systems operating in closed loop. *International Journal of Control*, 21:234–255, 1975.

[106] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer-Verlag, 1983.

[107] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall International, 1989.

[108] T. Söderström, P. Stoica, and E. Trulsson. Instrumental variable methods for closed loop systems. In *Preprints IFAC 10th World Congress on Automatic Control*, pages 364–369, Munich, Germany, July 1987.

[109] G. Stein and J. C. Doyle. Beyond singular values and loop shapes. *Journal of Guidance, Control & Dynamics*, 14:5–16, 1991.

[110] T. T. Tay, J. B. Moore, and R. Horowitz. Indirect adaptive techniques for fixed controller performance enhancement. *International Journal of Control*, 50:1941–1959, 1989.

[111] P. M. J. Van den Hof and R. J. P. Schrama. An indirect method for transfer function estimation from closed loop data. *Automatica*, 29:1523–1527, 1993.

[112] P. M. J. Van den Hof and R. J. P. Schrama. Identification and control — Closed-loop issues. *Automatica*, 31:1751–1770, 1995.

[113] P. M. J. Van den Hof, R. J. P. Schrama, R. A. de Callafon, and O. H. Bosgra. Identification of normalized coprime factors from closed-loop experimental data. *European Journal of Control*, 1:62–74, 1995.

[114] P. Van Overschee and B. De Moor. *Subspace identification for linear systems*. Kluwer, 1996.

[115] P. Van Overschee and B. De Moor. Closed-loop subspace system identification. *Submitted to Automatica*, 1997.

[116] M. Verhaegen. Application of a subspace model identification technique to identify LTI systems operating in closed-loop. *Automatica*, 29:1027–1040, 1993.

[117] M. Verhaegen. Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica*, 30:61–74, 1994.

[118] M. Viberg. Subspace-based methods for identification of linear time-invariant systems. *Automatica*, 31:1835–1851, 1995.

[119] B. Wahlberg and L. Ljung. Hard frequency-domain model error bounds from least-squares like identification techniques. *IEEE Transactions on Automatic Control*, 37:900–912, 1992.

[120] G. Zames. Feddback and optimal sensitivity: model reference transformations, multiplicative seminorms and approximate inverses. *IEEE Transactions on Automatic Control*, 26:301–320, 1981.

[121] Z. Zang, R. R. Bitmead, and M. Gevers. H_2 iterative model refinement and control robustness enhancement. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 279–284, Brighton, U.K., 1991.

[122] Z. Zang, R. R. Bitmead, and M. Gevers. Disturbance rejection: on-line refinement of controllers by closed-loop modelling. In *Proceedings of the American Control Conference*, pages 2829–2833, Chicago, IL, 1992.

[123] Z. Zang, R. R. Bitmead, and M. Gevers. Iterative weighted least-squares identification and weighted LQG control design. *Automatica*, 31:1577–1594, 1995.

[124] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, 1996.

[125] Y.-C. Zhu. Black-box identification of MIMO transfer functions: Asymptotic properties of prediction error models. *International Journal of Adaptive Control and Signal Processing*, 3:357–373, 1989.

[126] J. G. Ziegler and N. B. Nichols. Optimum settings for automatic controllers. *Transactions of the American Society of Mechanical Engineers*, 64:759–768, 1942.

Part II

Publications

Paper A

Closed-loop Identification Revisited

Authors: Urban Forssell and Lennart Ljung

Edited version of paper to appear in *Automatica*.

Preliminary version published as Technical Report LiTH-ISY-R-2021, Department of Electrical Engineering, Linköping University, Linköping, Sweden.

Closed-loop Identification Revisited

Urban Forssell and Lennart Ljung

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
URL: http://www.control.isy.liu.se/.

Abstract

Identification of systems operating in closed loop has long been of prime interest in industrial applications. The problem offers many possibilities, and also some fallacies, and a wide variety of approaches have been suggested, many quite recently. The purpose of the current contribution is to place most of these approaches in a coherent framework, thereby showing their connections and display similarities and differences in the asymptotic properties of the resulting estimates. The common framework is created by the basic prediction error method, and it is shown that most of the common methods correspond to different parameterizations of the dynamics and noise models. The so called *indirect methods*, e.g., are indeed "direct" methods employing noise models that contain the regulator. The asymptotic properties of the estimates then follow from the general theory and take different forms as they are translated to the particular parameterizations. We also study a new projection approach to closed-loop identification with the advantage of allowing approximation of the open loop dynamics in a given, and user-chosen frequency domain norm, even in the case of an unknown, nonlinear regulator.

Keywords: System identification; Closed-loop identification; Prediction error methods; Modelling; Statistical analysis.

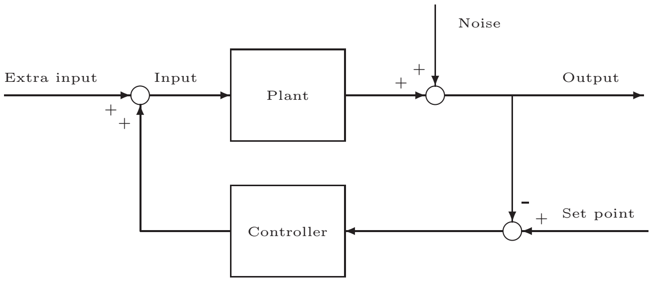

Figure 1 *A closed-loop system*

1 Introduction

1.1 Background and Scope

System identification is a well established field with a number of approaches, that can broadly be classified into the *prediction error family*, e.g,, [24], the *subspace approaches*, e.g., [35], and the nonparametric *correlation and spectral analysis methods*, e.g., [5]. Of special interest is the situation when the data to be used has been collected under closed-loop operation, as in Fig. 1.

The fundamental problem with closed-loop data is the correlation between the unmeasurable noise and the input. It is clear that whenever the feedback controller is not identically zero, the input and the noise will be correlated. This is the reason why several methods that work in open loop fail when applied to closed-loop data. This is for example true for the subspace approach and the nonparametric methods, unless special measures are taken. Despite these problems, performing identification experiments under output feedback (i.e., in closed loop) may be necessary due to safety or economic reasons, or if the system contains inherent feedback mechanisms. Closed-loop experiments may also be advantageous in certain situations:

- In [13] the problem of optimal experiment design is studied. It is shown that if the model is to be used for minimum variance control design the identification experiment should be performed in closed-loop with the optimal minimum variance controller in the loop. In general it can be seen that optimal experiment design with criteria penalizing both the error in the dynamics model and in the noise model leads to closed-loop solutions [24, 18]. The same is true with variance constraints on the output [9].
- In "identification for control" (see, e.g., [37, 21, 7]) the objective is to achieve a model that is suited for robust control design. Thus one has to tailor the experiment and preprocessing of data so that the model is reliable in regions

where the design process does not tolerate significant uncertainties. The use of closed-loop experiments has been a prominent feature in these approaches.

Historically, there has been a substantial interest in both special identification techniques for closed-loop data, and for analysis of existing methods when applied to such data. One of the earliest results was given by Akaike [1] who analyzed the effect of feedback loops in the system on correlation and spectral analysis. In the seventies there was a very active interest in questions concerning closed-loop identification, as summarized in the survey paper [15]. See also [2]. Up to this point much of the attention had been directed towards identifiability and accuracy problems. With the increasing interest "identification for control", the focus has shifted to the ability to shape the bias distribution so that control-relevant model approximations of the system are obtained. The surveys [12] and [32] cover most of the results along this line of research.

It is the purpose of the present paper to "revisit" the area of closed-loop identification, to put some of the new results and methods into perspective, and to give a status report of what can be done and what cannot. In the course of this exposé, new results will also be generated.

We will exclusively deal with methods derived in the prediction error framework and most of the results will be given for the multi-input multi-output (MIMO) case. The leading idea in the paper will be to provide a unified framework for many closed-loop methods by treating them as different parameterizations of *the* prediction error method. This will be a useful starting point for a comparative study of the different methods and it allows a simultaneous analysis of the methods using the basic statements on the asymptotic statistical properties of the general prediction error method:

- Convergence and bias distribution of the limit transfer function estimates.
- Asymptotic variance of the transfer function estimates (as the model orders tend to infinity).
- Asymptotic variance and distribution of the parameter estimates.

(For completeness and ease of reference we have included precise formulations of these theoretical results in Appendix A.)

1.2 Approaches to Closed-loop Identification

In the literature several different types of closed-loop identification methods have been suggested. We will use the set-up depicted in Fig. 1 to explain the ideas. One may distinguish between methods that

(a) Assume no knowledge about the nature of the feedback mechanism and use only measurements of the plant input, ignoring extra inputs or set-point signals even if known.

(b) Assume the feedback to be known and use possible extra inputs or set-point signals.

(c) Assume the regulator to be unknown, but of a certain structure (like in Fig. 1).

In (b) and (c) it is common to assume that the controller is linear, but the same ideas can also be applied if nonlinear and/or time-varying controllers are used. The price is, of course, that the estimation problems then become much more involved. For simplicity, we will in general assume linear, time-invariant feedback when discussing methods that fall under (b) and (c) in this paper. If the plant input is measured it can be noted that knowing the controller and extra inputs or set-point signals brings no additional information about the open-loop system. Conversely, with a known linear, time-invariant, and noise-free controller we can easily reconstruct the plant input using measurements of the output and extra inputs or set-point signals and then these carry no further information about the open-loop system. Thus there is no major difference between (a), (b), and (c) when the controller is linear, time-invariant, and noise-free. The problem in industrial practice is rather that no controller has this simple form: Various delimiters, anti-windup functions and other nonlinearities will have the controller deviate from the linear, time-invariant form, even if the regulator parameters (e.g. PID-coefficients) are known. This strongly disfavors the second approach.

In this paper we will use a classification of the different methods that is similar to the one in [15]. See also [28]. The basis for the classification is the different kinds of possible assumptions on the feedback listed above. The closed-loop identification methods correspondingly fall into the following main groups (cf. Fig. 1):

1. *The Direct Approach*: Ignore the feedback and identify the open-loop system using measurements of the input and the output.
2. *The Indirect Approach*: Identify some closed-loop transfer function and determine the open-loop parameters using the knowledge of the (linear) controller.
3. *The Joint Input-Output Approach*: Regard the input and output jointly as the output from a system driven by some extra input or set-point signal and noise. Use some method to determine the open-loop parameters from an estimate of this augmented system.

These categories are basically the same as those in [15], the only difference is that in the joint input-output approach we allow the augmented system to have a measurable input in addition to the unmeasurable noise. The indirect and joint input-output approaches typically assume linear feedback but can also be applied with nonlinear feedback, although these schemes then tend to be very involved.

In the closed-loop identification literature it has been common to classify the methods primarily based on how the final estimates are computed (e.g. directly or indirectly using multi-step estimation schemes), and then the main groupings have been into “direct” and “indirect” methods. This should not, however, be confused with the classification (1)-(3) which is based on the assumptions made on the feedback.

1.3 Outline

The rest of the paper is organized as follows. Next, in Section 2 the assumptions we make on the data generating mechanism are formalized. This section also introduces some of the notation that will be used in the paper. Section 3 presents the prediction error method and contains a summary of the asymptotic statistical properties of this method. The three basic approaches to closed-loop identification are discussed in Sections 4 and 5. The direct and indirect approaches are covered in Section 4, while Section 5 deals with the joint input-output approach. The asymptotic properties of the different methods are studied in Sections 6-8. In Section 6 convergence results for the limit estimates are given. Section 7 deals with the asymptotic variance of the transfer function estimates as the model orders tend to infinity. The asymptotic distribution of the parameter estimates is studied in Section 8. The last section, Section 9, contains a summarizing discussion on the main results. Two appendices have also been included in the paper. Appendix A contains the basic results for the statistical properties of the prediction error method, here given in full detail. Appendix B contains the deferred proofs from Sections 6-8.

2 Technical Assumptions and Notation

The basis of all identification is the data set

$$Z^N = \{u(1), y(1), \ldots, u(N), y(N)\} \tag{1}$$

consisting of measured input and output signals $u(t)$ and $y(t)$, $t = 1, \ldots, N$. We will make the following assumptions regarding how this data set was generated.

Assumption 1 *The true system $\mathcal{S}$ is causal, linear, and time-invariant with p outputs and m inputs and given by*

$$y(t) = G_0(q)u(t) + v(t), \quad v(t) = H_0(q)e(t) \tag{2}$$

where $\{e(t)\}$ $(p \times 1)$ is a zero-mean white noise process with covariance matrix Λ_0, and bounded moments of order $4+\delta$, some $\delta > 0$, and $H_0(q)$ is an inversely stable, monic filter.

Assumption 2 *The input $\{u(t)\}$ is given by*

$$u(t) = k(t, y^t, u^{t-1}, r(t)) \tag{3}$$

where $y^t = [y(1), \dots, y(t)]$, etc., and where the reference signal $\{r(t)\}$ is a given quasi-stationary signal, independent of $\{v(t)\}$ and k is a given deterministic function such that the closed-loop system (2) and (3) is exponentially stable, which we define as follows: For each $t, s;\ t \geq s$ there exist random variables $\bar{y}_s(t), \bar{u}_s(t)$, independent of r^s and v^s but not independent of r^t and v^t, such that

$$\bar{E}\|y(t) - \bar{y}_s(t)\|^4 < C\lambda^{t-s} \tag{4}$$

$$\bar{E}\|u(t) - \bar{u}_s(t)\|^4 < C\lambda^{t-s} \tag{5}$$

for some $C < \infty,\ \lambda < 1$. In addition, k is such that $G_0(q)u(t)$ is independent of $e(s)$ for $s \geq t$.

In the paper q will denote the time-shift operator (i.e., $q^{-1}u(t) = u(t-1)$). The symbol $\bar{E}$ is defined as

$$\bar{E}f(t) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} Ef(t) \tag{6}$$

Unless otherwise stated, we will assume that all signals are quasi-stationary [24]. This means that we can define spectra and cross-spectra of signals, which will be denoted $\Phi_u(\omega)$, $\Phi_{ue}(\omega)$, etc. We will assume that we can factorize the $\Phi_u(\omega)$ (and all other signal spectra) as

$$\Phi_u(\omega) = M_u(e^{i\omega})M_u^T(e^{-i\omega}) \tag{7}$$

where $M_u(e^{i\omega})$ is causal, stable, and inversely stable.

Introduce the signal $\chi_0(t) = \begin{bmatrix} u^T(t) & e^T(t) \end{bmatrix}^T$. The spectrum is

$$\Phi_{\chi_0} = \begin{bmatrix} \Phi_u & \Phi_{ue} \\ \Phi_{eu} & \Lambda_0 \end{bmatrix} \tag{8}$$

Here we have suppressed the argument ω which also will be done in the sequel whenever there is no risk of confusion. Similarly, we will also frequently suppress the arguments t, q, and $e^{i\omega}$ for notational convenience. The following factorizations of Φ_{χ_0} will be useful (it is assumed that $\Phi_u > 0$ and $\Lambda_0 > 0$):

$$\Phi_{\chi_0} = \begin{bmatrix} I & \Phi_{ue}\Lambda_0^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Phi_u^r & 0 \\ 0 & \Lambda_0 \end{bmatrix} \begin{bmatrix} I & 0 \\ \Lambda_0^{-1}\Phi_{eu} & I \end{bmatrix} \tag{9}$$

$$= \begin{bmatrix} I & 0 \\ \Phi_{eu}\Phi_u^{-1} & I \end{bmatrix} \begin{bmatrix} \Phi_u & 0 \\ 0 & \Phi_e^r \end{bmatrix} \begin{bmatrix} I & \Phi_u^{-1}\Phi_{ue} \\ 0 & I \end{bmatrix} \tag{10}$$

Here

$$\Phi_u^r = \Phi_u - \Phi_{ue}\Lambda_0^{-1}\Phi_{eu} \tag{11}$$

$$\Phi_e^r = \Lambda_0 - \Phi_{eu}\Phi_u^{-1}\Phi_{ue} \tag{12}$$

We will also use

$$\Phi_{\chi_0} = \Phi_{\chi_0}^r + \Phi_{\chi_0}^e, \quad \Phi_{\chi_0}^r = \begin{bmatrix} \Phi_u^r & 0 \\ 0 & 0 \end{bmatrix}, \quad \Phi_{\chi_0}^r = \begin{bmatrix} \Phi_u^e & \Phi_{ue} \\ \Phi_{eu} & \Lambda_0 \end{bmatrix} \tag{13}$$

$$\Phi_u^e = \Phi_u - \Phi_u^r = \Phi_{ue}\Lambda_0^{-1}\Phi_{eu} \tag{14}$$

The matrix Φ_u^r defined in (11) can be seen as that part of the input spectrum that originates from the reference signal r. Similarly, Φ_u^e defined in (14) can be seen as that part of the input spectrum that originates from the noise e.

For some of the analytic treatment we shall assume that the input u is generated as

$$u(t) = r(t) - K(q)y(t) \tag{15}$$

where K is a linear, time-invariant regulator of appropriate dimensions and where r is the reference signal (e.g., an extra input or a set-point signal), which is assumed independent of the noise v. If (15) holds then Assumption 2 simply means that the closed-loop system is asymptotically stable.

Let us now introduce some further notation for the linear feedback case. By combining the equations (2) and (15) we get the closed-loop relations

$$y(t) = S_0(q)G_0(q)r(t) + S_0(q)v(t) \tag{16}$$

$$u(t) = S_0^i(q)r(t) - S_0^i(q)K(q)v(t) \tag{17}$$

where S_0 and S_0^i are the output and input sensitivity functions, respectively,

$$S_0(q) = (I + G_0(q)K(q))^{-1} \tag{18}$$

$$S_0^i(q) = (I + K(q)G_0(q))^{-1} \tag{19}$$

For future use we also introduce

$$G_0^c(q) = S_0(q)G_0(q) \text{ and } H_0^c(q) = S_0(q)H_0(q) \tag{20}$$

so that we can rewrite the closed-loop system (16) as

$$y(t) = G_0^c(q)r(t) + v_c(t), \quad v_c(t) = H_0^c(q)e(t) \tag{21}$$

Using (17) we can easily derive the following expressions for Φ_u^r and Φ_u^e under (15):

$$\Phi_u^r = S_0^i\Phi_r(S_0^i)^* \tag{22}$$

$$\Phi_u^e = S_0^iK\Phi_vK^*(S_0^i)^* = KS_0\Phi_vS_0^*K^* \tag{23}$$

Here Φ_r is the spectrum of the reference signal and $\Phi_v = H_0\Lambda_0 H_0^*$ the noise spectrum. Superscript $*$ denotes complex conjugate transpose. If u is generated as in (15) the cross-spectrum between u and e is

$$\Phi_{ue} = -KS_0H_0\Lambda_0 = -S_0^iKH_0\Lambda_0 \tag{24}$$

Occasionally we shall also consider the case where the regulator is linear as in (15) but contains an unknown additive disturbance d:

$$u(t) = r(t) - K(q)y(t) + d(t) \tag{25}$$

We will assume that the nature of d is such that the closed-loop system (2), (25) is exponentially stable in the sense defined in Assumption 2. The origin of the disturbance d can, e.g., be an imperfect knowledge of the true regulator parameters: Suppose that the true regulator is given by $K + \Delta_K$ where K is known and Δ_K is unknown. We could then use the model (25) to describe how the input was generated. (In this case $d = -\Delta_K y$.) It is also worth noting that if K is known and both u and y are measured we can regard d as part of the external excitation since then we can always construct

$$r'(t) = u(t) + K(q)y(t) \ (= r(t) + d(t)) \tag{26}$$

and replace r by this signal in our model. This would lead us back to the situation (15). See, e.g., [30].

3 Prediction Error Identification

In this section we shall review some basic results on prediction error methods, that will be used in the sequel. See Appendix A and [24] for more details.

3.1 The Method

We will work with a model structure $\mathcal{M}$ of the form

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{27}$$

G will be called the dynamics model and H the noise model. Unless otherwise explicitly stated, we will assume that G is causal and such that $G(q,\theta)u(t)$ is independent of $e(s)$ for $s \geq t$ and that H is monic and causal. The parameter vector θ ranges over a set $D_{\mathcal{M}}$ which is assumed compact and connected. The one-step-ahead predictor for the model structure (27) is, [24]

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (I - H^{-1}(q,\theta))y(t) \tag{28}$$

The prediction error is

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{29}$$

Given the model (28) and measured data Z^N we determine the prediction error estimate through

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N) \tag{30}$$

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_F^T(t,\theta)\Lambda^{-1}\varepsilon_F(t,\theta) \tag{31}$$

$$\varepsilon_F(t,\theta) = L(q,\theta)\varepsilon(t,\theta) \tag{32}$$

Here Λ is a symmetric, positive definite weighting matrix and L a (possibly parameter-dependent) monic prefilter that can be used to enhance certain frequency regions. It is easy to see that

$$\varepsilon_F(t,\theta) = L(q,\theta)H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{33}$$

Thus the effect of the prefilter L can be included in the noise model and $L(q,\theta) = I$ can be assumed without loss of generality. This will be done in the sequel. The resulting transfer function (matrix) estimates will be denoted $\hat{G}_N(q)$, $\hat{G}_N(q) = G(q,\hat{\theta}_N)$, etc.

We say that the true system is contained in the model set if, for some $\theta_0 \in D_{\mathcal{M}}$,

$$G(q,\theta_0) = G_0(q), \;\; H(q,\theta_0) = H_0(q) \tag{34}$$

This will also be written $\mathcal{S} \in \mathcal{M}$. The case when the true noise properties cannot be correctly described within the model set but where there exists a $\theta_0 \in D_{\mathcal{M}}$ such that

$$G(q,\theta_0) = G_0(q) \tag{35}$$

will be denoted $G_0 \in \mathcal{G}$.

3.2 Convergence

Define the average criterion $\bar{V}(\theta)$ as

$$\bar{V}(\theta) = \bar{E}\varepsilon^T(t,\theta)\Lambda^{-1}\varepsilon(t,\theta) \tag{36}$$

Then we have the following result (see, e.g., [22, 24]):

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \bar{V}(\theta) \quad \text{with probability (w. p.) 1 as } N \to \infty \tag{37}$$

Note that we can write $\bar{V}(\theta)$ as

$$\bar{V}(\theta) = \bar{E}\,\mathrm{tr}\big[\Lambda^{-1}\varepsilon(t,\theta)\varepsilon^T(t,\theta)\big] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{tr}\big[\Lambda^{-1}\Phi_\varepsilon\big]d\omega \tag{38}$$

Here Φ_ε is the spectrum of the prediction error $\varepsilon(t,\theta)$. It can thus be seen that D_c defined in (37) depends only on the second order properties of (the spectrum) of the prediction error.

In case the input-output data can be described by (2) we have the following characterization of D_c (G_θ is short for $G(q,\theta)$, etc.):

$$D_c = \arg\min_{\theta\in D_{\mathcal{M}}}\int_{-\pi}^{\pi}\mathrm{tr}\Big[\big[(G_0-G_\theta)\quad(H_0-H_\theta)\big]\Phi_{\chi_0}\begin{bmatrix}(G_0-G_\theta)^*\\(H_0-H_\theta)^*\end{bmatrix}(H_\theta\Lambda H_\theta^*)^{-1}\Big]d\omega \tag{39}$$

where Φ_{χ_0} is given by (8). This is shown in Appendix A.1. Note that the result holds regardless of the nature of the regulator, as long as Assumptions 1 and 2 hold and the signals involved are quasistationary (which is a standing assumption).

From (39) several conclusions regarding the consistency of the method can be drawn. First of all, suppose that the parameterization of G and H is sufficiently flexible so that $\mathcal{S}\in\mathcal{M}$. If this holds, the method will give consistent estimates of G_0 and H_0 if the experiment is *informative* [24], which means that the matrix Φ_{χ_0} is positive definite for all frequencies. (Note that it will always be positive *semi*-definite since it is a spectral matrix.) From the factorization (9) we see that this holds if, and only if, Φ_u^r defined in (11) is positive definite. The general condition for this to be true is that there should *not* be a linear, time-invariant, and noise-free relationship between u and y. With an external reference signal this is automatically satisfied, but it should be clear that informative closed-loop experiments can also be guaranteed if we switch between different linear regulators or use a nonlinear regulator. (See, e.g., [15, 24].) With a linear controller we need a persistently exciting reference signal to ensure that Φ_u^r is positive definite. To see this note that, in the linear case, $\Phi_u^r = S_0^i\Phi_r(S_0^i)^*$ so Φ_u^r is positive definite for all frequencies if, and only if, Φ_r is (which is the same as to say that the reference signal is persistently exciting [24]), since the analytical function S_0^i (cf. (22)) can be zero at at most finitely many points.

It can also be noted that the prediction error method can give consistent estimates even if the experiment is not informative, depending on the parameterization of G and H: If $\mathcal{S}\in\mathcal{M}$ the method still gives consistent estimates of G_0 and H_0 even though Φ_{χ_0} is not positive definite provided that $\big[(G_0-G_\theta)\quad(H_0-H_\theta)\big]$ does not lie in the (left) null space of Φ_{χ_0}. In [24] this is called that the experiment is *informative enough* with respect to the chosen model structure.

3.3 Asymptotic Variance of Black Box Transfer Function Estimates

Consider the model structure (27). Introduce

$$T(q,\theta) = \operatorname{vec}\big[G(q,\theta)\ H(q,\theta)\big] \tag{40}$$

(The vec-operator stacks the columns of its argument on top of each other in a vector. A more formal definition is given in Appendix A.2.) Suppose that the vector θ can be decomposed so that

$$\theta = \big[\theta_1^T,\ \theta_2^T,\ \ldots,\ \theta_n^T\big]^T \quad \dim\theta_k = s \quad \dim\theta = n\cdot s \tag{41}$$

We shall call n the order of the model structure (27) and for some results we will allow n to tend to infinity as N tends to infinity. Suppose also that T in (40) has the following shift structure:

$$\frac{\partial}{\partial\theta_k}T(q,\theta) = q^{-k+1}\frac{\partial}{\partial\theta_1}T(q,\theta) \tag{42}$$

It should be noted that most polynomial-type model structures, including the ones studied in this paper, satisfy this shift structure. Thus (42) is a rather weak assumption.

More background material including further technical assumptions and additional notation can be found in Appendix A.2. For brevity reasons we here go directly to the main result ($\otimes$ denotes the Kronecker product):

$$\operatorname{Cov}\operatorname{vec}\big[\hat{T}_N(e^{i\omega})\big] \sim \frac{n}{N}(\Phi_{\chi_0}(\omega))^{-T}\otimes\Phi_v(\omega) \tag{43}$$

The covariance matrix is thus proportional to the model order n divided by the number of data N. This holds asymptotically as both n and N tend to infinity. In open loop we have $\Phi_{ue}=0$ and

$$\operatorname{Cov}\operatorname{vec}\big[\hat{G}_N(e^{i\omega})\big] \sim \frac{n}{N}(\Phi_u(\omega))^{-T}\otimes\Phi_v(\omega) \tag{44}$$

$$\operatorname{Cov}\operatorname{vec}\big[\hat{H}_N(e^{i\omega})\big] \sim \frac{n}{N}\Lambda_0^{-1}\otimes\Phi_v(\omega) \tag{45}$$

Notice that the result (44) for the dynamics model holds also in case the noise model is fixed (e.g., $H(q,\theta)=I$).

3.4 Asymptotic Distribution of Parameter Vector Estimates

If $\mathcal{S} \in \mathcal{M}$ then $\hat{\theta}_N \to \theta_0$ as $N \to \infty$ under reasonable conditions on the data set [24]. Then, if $\Lambda = \Lambda_0$,

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, P_\theta) \tag{46a}$$

$$P_\theta = \left[\bar{E}\psi(t,\theta_0)\Lambda_0^{-1}\psi^T(t,\theta_0)\right]^{-1} \tag{46b}$$

where ψ is the negative gradient of the prediction error ε with respect to θ.

In this paper we will restrict to the SISO case when discussing the asymptotic distribution of the parameter vector estimates for notational convenience. For ease of reference we have in Appendix A.3 stated a variant of (46) as a theorem.

4 The Direct and Indirect Approaches

4.1 The Direct Approach

The direct approach amounts to applying a prediction error method directly to input-output data, ignoring possible feedback. This implies that this method can be applied to system with arbitrary (unknown) feedback mechanisms. In general one works with model structures of the form (27), that is

$$y(t) = G(q,\theta)u(t) + H(q,\theta))e(t) \tag{47}$$

The direct method can thus be formulated as in (30)-(32) and coincides with the standard (open-loop) prediction error method (e.g., [24, 28]).

It is sometimes claimed that the direct approach cannot be used if the underlying system is unstable. This is not correct: The prediction error method can be applied directly also to unstable systems as long as the *predictor* is stable. This is automatically satisfied with ARX and ARMAX models and can also be guaranteed for the output error and Box-Jenkins model structures if special measures are taken [10]. Let us now turn to the indirect approach.

4.2 The Idea Behind the Indirect Approach

If the controller and some extra input or set-point signal (cf. Fig. 1) is known we may use indirect identification. This method can be applied to systems with nonlinear feedback, but to explain the idea it is easier to restrict to linear feedback. This has also been a standard assumption in the literature. Suppose therefore that the input is generated as in (15). The closed-loop system is (cf. (21)):

$$y(t) = G_0^c(q)r(t) + v_c(t) \tag{48}$$

where $G_0^c = (I + G_0K)^{-1}G_0$. If the regulator K is known and r is measurable, we can identify the closed-loop system, e.g., using the prediction error method with a model

$$y(t) = G^c(q, \theta)r(t) + H_*(q)e(t) \tag{49}$$

and compute an estimate $\hat{G}_N$ of the open-loop system by solving the equation

$$\hat{G}_N^c(q) = (I + \hat{G}_N(q)K(q))^{-1}\hat{G}_N(q) \tag{50}$$

This is the basic idea behind the indirect method and illustrates the some of the features of this method. For instance, note that by identifying the closed-loop system we have transformed the closed-loop identification problem into an "open loop" one, since r and v are uncorrelated. This means that we may use any method that works in open loop to find a model of the closed-loop system, for instance a subspace method or instrumental variables. Here we will focus on prediction error methods.

We have outlined a procedure consisting of two steps: First the closed-loop system is identified, then an open-loop model is computed. It should be realized, however, that there are alternative ways of implementing this method. For instance, instead of identifying the closed-loop system we may also identify the sensitivity function using measurements of u and r, see [6]. Here we will limit the discussion to methods that aim at identifying the closed-loop system. The prediction error approach allows arbitrary parameterizations of the models, so if this method is used it is also possible to construct different indirect "methods" that suit our purposes by considering different parameterizations. Such ideas will be studied next.

4.3 Indirect Identification Using the Prediction Error Method

With the indirect method we can use output error models with fixed noise models/prefilters to shape the bias distribution (cf. Corollary 6 below). The model structure would then be given by (49) where $G^c(q, \theta)$ is a parameterized model of the closed-loop system. This gives us an estimate $\hat{G}_N^c$ of the closed-loop system and the open-loop estimate can then be found using (50). The exact solution is

$$\hat{G}_N(q) = \hat{G}_N^c(q)(I - \hat{G}_N^c(q)K(q))^{-1} \tag{51}$$

This is of course a perfectly valid solution, although an estimate $\hat{G}_N$ computed this way will generally be of high order. (The order will typically be equal to the sum of the orders of $\hat{G}_N^c$ and K.) If we attempt to solve (50) with the additional constraint that $\hat{G}_N$ should be of a certain (low) order we end up with an over-determined system of equations which can be solved in many ways, for instance in a weighted least squares sense. Examples of such indirect "methods" will be given in Section 8.2 below. A particularly simple and attractive approach is to let the parameters θ relate to properties of the open-loop system G, so that in the first step we should parameterize $G^c(q,\theta)$ as

$$G^c(q,\theta) = (I + G(q,\theta)K(q))^{-1}G(q,\theta) \tag{52}$$

(This was apparently first suggested as an exercise in [24]. See also [34].) The advantage with this parameterization is that the (problematic) second step (50) is avoided and an open loop model $\hat{G}_N$ is delivered directly.

The parameterization of $G^c(q,\theta)$ may thus be important for numeric and algebraic reasons, but we emphasize that it does not affect the asymptotic statistical properties of the estimate $\hat{G}_N^c(q)$, as long as it covers the same set of transfer functions.

Another nice and interesting parameterization idea is to use the, so called, dual Youla parameterization that parameterizes all systems that are stabilized by a certain regulator K (see, e.g., [36]). This will be studied next.

The Dual Youla Parameterization

To present the idea, the concept of coprime factorizations of transfer functions is required: A pair of stable transfer functions $N,\ D$ is a right coprime factorization (rcf) of G if $G = ND^{-1}$ and there exist stable transfer functions $X,\ Y$ such that $XN + YD = I$. The dual Youla parameterization now works as follows. Let G_{nom} with $\mathrm{rcf}(N,D)$ be any system that is stabilized by K with $\mathrm{rcf}(X,Y)$. Then, as R ranges over all stable transfer functions, the set

$$\big\{G : G(q,\theta) = (N(q) + Y(q)R(q,\theta))(D(q) - X(q)R(q,\theta))^{-1}\big\} \tag{53}$$

describes all systems that are stabilized by K. The unique value of R that corresponds to the true plant G_0 is given by

$$R_0(q) = Y^{-1}(q)(I + G_0(q)K(q))^{-1}(G_0(q) - G_{nom}(q))D(q) \tag{54}$$

This idea can now be used for identification (see, e.g., [16], [17], and [6]): Given an estimate $\hat{R}_N$ of R_0 we can compute an estimate of G_0 as

$$\hat{G}_N(q) = (N(q) + Y(q)\hat{R}_N(q))(D(q) - X(q)\hat{R}_N(q))^{-1} \tag{55}$$

Using the dual Youla parameterization we can write

$$G^c(q,\theta) = (N(q) + Y(q)R(q,\theta))(D(q) + X(q)Y^{-1}(q)N(q))^{-1} \tag{56}$$
$$\triangleq (N(q) + Y(q)R(q,\theta))M(q) \tag{57}$$

With this parameterization the identification problem (49) becomes

$$z(t) = R(q,\theta)x(t) + \bar{v}_c(t) \tag{58}$$

where

$$z(t) = Y^{-1}(q)(y(t) - N(q)M(q)r(t)) \tag{59}$$
$$x(t) = M(q)r(t) \tag{60}$$
$$\bar{v}_c(t) = Y^{-1}(q)v_c(t) \tag{61}$$

Thus it can be seen that the dual Youla method is a special parameterization of the general indirect method. This means, as mentioned, that the statistical properties of the resulting estimates for the indirect method remain unaffected for the dual Youla method. The main advantage of this method is of course that the obtained estimate $\hat{G}_N$ is guaranteed to be stabilized by K, which clearly is a nice feature. A drawback is that this method typically will give high-order estimates – typically the order will be equal to the sum of the orders of G_{nom} and $\hat{R}_N$.

Before turning to the joint input-output approach, let us pause and study an interesting variant of the parameterization idea used in (52) which will provide useful insights into the connection between the direct and indirect methods.

4.4 A Formal Connection Between Direct and Indirect Methods

The noise model H in a linear dynamics model structure has often turned out to be a key to interpretation of different "methods". The distinction between the models/"methods" ARX, ARMAX, output error, Box-Jenkins, etc., is entirely explained by the choice of the noise model. Also the practically important feature of prefiltering is equivalent to changing the noise model. Even the choice between minimizing one- or k-step prediction errors can be seen as a noise model issue. See, e.g., [24] for all this.

Therefore it should not come as a surprise that also the distinction between the fundamental approaches of direct and indirect identification can be seen as a choice of noise model.

The idea is to use a model parameterized as

$$y(t) = G(q,\theta)u(t) + (I + G(q,\theta)K(q))H_1(q,\theta)e(t) \tag{62}$$

We thus link the noise model to the dynamics model. There is nothing strange with that: So do ARX and ARMAX models. The parameterization (62) is thus perfectly valid, but it should still be pointed out that this is a highly specialized parameterization using the knowledge of K. Also note that this particular parameterization scales H_1 with the inverse model sensitivity function. Similar parameterization ideas have been studied in, e.g., [4, 20, 8, 34].

The predictor for (62) is

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (I - H^{-1}(q,\theta))y(t) \tag{63}$$

Using $u = r - Ky$ and inserting (62) we get

$$\hat{y}(t|\theta) = H_1^{-1}(q,\theta)(I + G(q,\theta)K(q))^{-1}G(q,\theta)r(t) + (I - H_1^{-1}(q,\theta))y(t) \tag{64}$$

But this is exactly the predictor also for the closed-loop model structure

$$y(t) = (I + G(q,\theta)K(q))^{-1}G(q,\theta)r(t) + H_1(q,\theta)e(t) \tag{65}$$

and hence the two approaches are equivalent. We formulate this result as a lemma:

Lemma 3 *Suppose that the input is generated as in* (15) *and that both u and r are measurable and that the linear regulator K is known. Then, applying a prediction error method to* (62) *or to* (65) *gives identical estimates $\hat{\theta}_N$. This holds regardless of the parameterization of G and H_1.*

Among other things, this shows that we can use any theory developed for the direct approach (allowing for feedback) to evaluate properties of the indirect approach, and vice versa. It can also be noted that the particular noise model used in (62) is the answer to the question how H should be parameterized in the direct method in order to avoid the bias in the G-estimate in the case of closed-loop data, even if the true noise characteristics is not correctly modeled. This is further studied in [25].

After having discussed so many variants of the indirect method it is perhaps in place to remark that we will refer to the model structure (49) with G^c is parameterized as in (52) when discussing the indirect method in the sequel. The only exception will be Section 8.2 where we discuss a different implementation of the indirect method.

5 The Joint Input-output Approach

The third main approach to closed-loop identification is the so called joint input-output approach. The basic assumption in this approach is that the input is generated using a regulator of a certain form, e.g., (15). Exact knowledge of the regulator parameters is not required – an advantage over the indirect method where this is a necessity.

5.1 The Idea

Suppose that the feedback is linear and given by (15). Closed-loop expressions for the output and the input were given in (16) and (17), respectively:

$$y(t) = G_0^c(q)r(t) + S_0(q)v(t) \tag{66}$$

$$u(t) = S_0^i(q)r(t) - S_0^i(q)K(q)v(t) \tag{67}$$

The basic idea in the joint input-output approach is to compute estimates of the open-loop system using estimates of the different transfer functions in (66) and (67). A straightforward approach would be to identify the closed-loop system and the input sensitivity function and form the estimate $\hat{G}_N = \hat{G}_N^c(\hat{S}_N^i)^{-1}$. (The rationale behind this is the relation $G_0^c(S_0^i)^{-1} = G_0$.) This idea really goes back to Akaike [1] who showed that spectral analysis of closed-loop data should be performed as follows: Compute the spectral estimates (SISO)

$$\hat{G}_N^{yx} = \frac{\hat{\Phi}_{yx}^N}{\hat{\Phi}_x^N} \quad \text{and} \quad \hat{G}_N^{ux} = \frac{\hat{\Phi}_{ux}^N}{\hat{\Phi}_x^N} \tag{68}$$

where the signal x is correlated with y and u, but uncorrelated with the noise e (a standard choice is $x = r$) and estimate the open-loop system as

$$\hat{G}_N^{yu} = \frac{\hat{G}_N^{yx}}{\hat{G}_N^{ux}} = \frac{\hat{\Phi}_{yx}^N}{\hat{\Phi}_{ux}^N} \tag{69}$$

For a discussion on "classical" joint input-output methods, see [15]. Recently, several new prediction error methods that fall in this framework have been suggested, e.g., the two-stage method [31] and the related projection method [11], and the coprime factor identification scheme [33]. In the following we will discuss how to implement the joint input-output approach utilizing the fact that the prediction error method allows arbitrary parameterizations of the models. This will allow us to do a unified analysis of all joint input-output methods using the general theory and to compare the results with the corresponding ones obtained for the direct and indirect approaches.

5.2 Joint Input-output Identification Using the Prediction Error Method

Suppose that the regulator is linear and of the form (25) and suppose e and d are independent noise sources (d does not necessarily have to be white). The output y

and input u then obey

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \mathsf{G}_0(q) r(t) + \mathsf{H}_0(q) \begin{bmatrix} e(t) \\ d(t) \end{bmatrix} \tag{70}$$

$$\mathsf{G}_0(q) = \begin{bmatrix} G_0^c(q) \\ S_0^i(q) \end{bmatrix}, \quad \mathsf{H}_0(q) = \begin{bmatrix} S_0(q) H_0(q) & G_0(q) S_0^i(q) \\ -K(q) S_0(q) H_0(q) & S_0^i(q) \end{bmatrix} \tag{71}$$

The idea is to identify the augmented system (70) using a model of the form

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \mathsf{G}(q,\theta) r(t) + \mathsf{H}(q,\theta) \begin{bmatrix} e(t) \\ d(t) \end{bmatrix} \tag{72}$$

where the parameterizations of the transfer functions $\mathsf{G}(q,\theta)$ and $\mathsf{H}(q,\theta)$, for the time being, are not further specified, and to compute an open loop estimate $\hat{G}_N$ using $\hat{\mathsf{G}}_N$ and $\hat{\mathsf{H}}_N$. Different parameterization will lead to different methods, as we shall see. Here we have assumed a nonzero reference signal r. There are also methods that assume $r = 0$ and model y and u jointly as time series and retrieves the open-loop system from an estimate of H_0. See, e.g., [15, 28].

If we parameterize $\mathsf{G}(q,\theta)$ in (72) as

$$\mathsf{G}(q,\theta) = \begin{bmatrix} G^{yr}(q,\theta) \\ G^{ur}(q,\theta) \end{bmatrix} \tag{73}$$

and estimate the parameters θ using, e.g., a fixed noise model $\mathsf{H}(q,\theta) = \mathsf{H}_*(q)$ we can use the straightforward approach of simply dividing the estimates $\hat{G}_N^{yr}$ and $\hat{G}_N^{ur}$ (cf. (68)-(69)):

$$\hat{G}_N(q) = \hat{G}_N^{yr}(q)(\hat{G}_N^{ur}(q))^{-1} \tag{74}$$

We can also include a prefilter for r and use the filtered signal $x = F_r r$ (again, cf. (68)-(69)) in the model:

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} G^{yx}(q,\theta) \\ G^{ux}(q,\theta) \end{bmatrix} x(t) + \mathsf{H}_*(q) \begin{bmatrix} e(t) \\ d(t) \end{bmatrix} \tag{75}$$

and use

$$\hat{G}_N(q) = \hat{G}_N^{yx}(q)(\hat{G}_N^{ux}(q))^{-1} \tag{76}$$

It can be shown that the prefilter F_r can be used to shape the bias distribution of the resulting models (in case of undermodeling). See Corollary 7 below.

A drawback with proceeding bluntly as in (76) is that the estimate $\hat{G}_N$ inevitably will be of high order. This problem is similar to the one we are faced with in the indirect method, where we noted that solving for the open-loop estimate in (50) typically gives high-order estimates. A solution is to parameterize $G^{yx}(q,\theta)$ and $G^{ux}(q,\theta)$ as

$$G^{yx}(q,\theta) = G^{yu}(q,\rho)G^{ux}(q,\beta), \quad G^{ux}(q,\theta) = G^{ux}(q,\beta), \quad \theta = \begin{bmatrix}\rho \\ \beta\end{bmatrix} \tag{77}$$

This would give us control over the order of $\hat{G}_N$ through the factor $G^{yu}(q,\rho)$, and circumvent the step (76). If we disregard the correlation between the noise sources affecting y and u, we may also first estimate the β-parameters separately using u and r and then estimate the ρ-parameters using y and r, keeping the β-parameters fixed to their estimated values. We will study such methods further in Section 5.3 below.

Another interesting idea is to use the prefilter F_r to tune the method to give models $\hat{G}_N^{yx}$ and $\hat{G}_N^{ux}$ that are normalized coprime ($\hat{G}_N^{yx}$, $\hat{G}_N^{yx}$ stable, $(\hat{G}_N^{yx})^*\hat{G}_N^{yx} + (\hat{G}_N^{ux})^*\hat{G}_N^{ux} = I$). If we succeed in estimating normalized coprime factors, the open-loop model $\hat{G}_N$ will be of minimal order (the McMillan degree of $\hat{G}_N$ will be equal to that of $\hat{G}_N^{yx}$ and $\hat{G}_N^{ux}$) and we will achieve a complete elimination of redundant dynamics. This is the idea behind the so called coprime factor identification scheme [33]. (Note, though, that in practice $\hat{G}_N^{yx}$ and $\hat{G}_N^{ux}$ will never be exactly normalized coprime due to noise and the fact that the proper filter F_r to achieve normalized coprimeness cannot be calculated unless the system is perfectly known. In [33] it is therefore suggested to use an iterative procedure to achieve approximate normalized coprimeness.) In this method it is also suggested that the models $\hat{G}_N^{yx}$ and $\hat{G}_N^{ux}$ are parameterized on a common denominator form, similar to (77), to simplify the computation of $\hat{G}_N$.

5.3 The Two-stage and Projection Methods

The two-stage method [31] and the projection method [11] can both be explained using the following two steps:

1. Estimate the β-parameters in the model
$$u(t) = S(q,\beta)r(t) + H_1(q)e(t) \tag{78}$$
and construct the signal $\hat{u} = \hat{S}_N r$.
2. Identify the open-loop system using the model
$$y(t) = G(q,\rho)\hat{u}(t) + H_2(q)e(t) \tag{79}$$

This is the same parameterization idea as in (77) and these methods can thus be seen as joint input-output methods where the correlation between the noise sources in (78) and (79) is ignored. This classification will be further motivated in the next section.

The two-stage method was originally suggested as a method for consistently estimating the open-loop system, regardless of the noise model used, in case of linear feedback. Indeed, if the feedback is linear, given by (15), and the first step of the algorithm was successful so that $\hat{S}_N = S_0^i$, then the simulated signal $\hat{u}$ will be exactly equal to the noise-free part of the input and the open-loop system can be consistently estimated using (79). See Corollary 7 below. (Note that in the first step a high-order model of S_0 can be used, since we in the second step can control the open-loop model order independently. This implies that in most cases we can find a practically unbiased estimate of S_0^i in the first step.)

The two-stage method is quite robust and rather simple to use, therefore it is an attractive alternative to many of the existing closed-loop identification methods, including the direct method. However, the two-stage method will fail if the controller is nonlinear and/or contains some unknown disturbance signal d like in (25) that is not accounted for in the model. And, as we mentioned in the introduction, this will typically be the case in industrial applications even when the basic controller structure is linear. With the projection method these problems can be circumvented and, consequently, this method can be applied to systems with arbitrary feedback mechanisms, just as the direct method, and gives consistent estimates of G_0 regardless of the noise model used in the second step (79). The idea is to use a noncausal FIR model

$$u(t) = S(q, \beta)r(t) + e(t) = \sum_{k=-M_1}^{M_2} s_k r(t-k) + e(t) \tag{80}$$

with M_1 and M_2 chosen so large that any correlation between $u(t)$ and $r(s)$ for $t - s \notin [-M_1, \dots, M_2]$ can be ignored, in the first step of the algorithm. This will imply that $\hat{u}$ and $\tilde{u} = u - \hat{u}$ (asymptotically) will be uncorrelated or, in a least squares estimation language, *orthogonal* (hence the name *projection* method). As we shall see, this is the key to obtaining a consistent open-loop estimate.

It can thus be seen that the noncausal taps in the FIR model (80) is what distinguishes the projection method from the two-stage method, which considers only causal models. In [11] (see also below) it is shown that the noncausal taps cannot be excluded in general if the controller is nonlinear, even though the feedback system is causal (which we assume).

The model (80) can be seen as an approximation of the *Wiener smoother* (e.g., [19]) solving the linear least mean square problem (assuming infinite data records):

$$\min_S \bar{E}(u(t) - S(q)r(t))^T (u(t) - S(q)r(t)), \quad S(q) = \sum_{k=-\infty}^{\infty} s_k q^{-k} \tag{81}$$

The Wiener smoother has the property that $\hat{u}$ $(= \hat{S}r)$ and $\tilde{u}$ $(= u - \hat{u})$ are uncorrelated. In the frequency domain the solution can be expressed as $\hat{S} = \Phi_{ur}(\Phi_r)^{-1}$ (cf. (68)). In the linear case this equals S_0^i which of course is causal operator. In the nonlinear case, however, the asymptotically optimal Wiener smoother need not be causal even though the feedback system is causal. This is the reason why a noncausal model can be strictly better than a causal one (or, equivalently, why the projection method outperforms the two-stage method) asymptotically. In practice, with finite data records, the noncausal taps in (80) can be used to decrease the *de facto* correlation between $\hat{u}$ and $\tilde{u}$, which will have a positive effect on the quality of the final estimate. This is illustrated in the following example.

Example 4 *Consider the closed-loop system in Fig. 2. The system G_0 is given by*

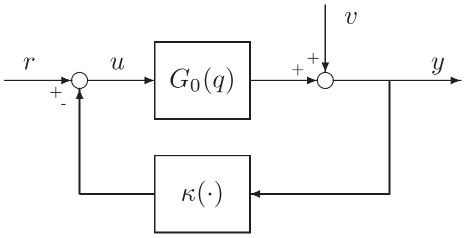

Figure 2 *Closed-loop system with nonlinear feedback.*

$$G_0(q) = q^{-1} - 0.8q^{-2} \tag{82}$$

and the feedback controller is

$$\kappa(x) = \begin{cases} 0.25x + 4, & \text{if } x \geq 0 \\ 0.25x - 4, & \text{if } x < 0 \end{cases} \tag{83}$$

Thus, apart from just a proportional term, the controller consists of a columbic friction element.

In order to identify G_0 we performed a noise-free ($v = 0$) simulation of this system using a Gaussian white noise reference signal. 5000 data samples were collected. Thus the only errors in the models should be bias errors.

To identify the system we used the two-stage and the projection methods. The following models were employed in the first step of the two-stage method:

1. *An FIR model:*

$$u(t) = b_0 r(t) + b_1 r(t-1) + \cdots + b_{10} r(t-10) + e(t) \tag{84}$$

2. An ARX model:

$$u(t) = -a_1u(t-1) - \cdots - a_5u(t-5) + b_0r(t) + \cdots + b_5r(t-5) + e(t) \quad (85)$$

In the first step of the projection method the model was

$$u(t) = s_{-5}r(t+5) + \cdots + s_0r(t) + \cdots + s_5r(t-5) + e(t) \quad (86)$$

That is, a noncausal FIR model. Notice that the same number of parameters are used in all three cases. In the second step of the methods the model was chosen as

$$y(t) = b_1\hat{u}(t-1) + b_2\hat{u}(t-2) + e(t) \quad (87)$$

where $\hat{u}$ was obtained using the first step models.

The identification results are summarized in Table 1:

Table 1. Summary of identification results.

	b_1	b_2
True system	1.0000	−0.8000
Two-stage (FIR)	0.9595	−0.8258
Two-stage (ARX)	0.8857	−0.6906
Projection	0.9985	−0.7965

Here we can note that the two variants of the two-stage method give biased results while the projection method gives a very accurate estimate. Bode plots of the resulting model errors are shown in Fig. 3 together with a plot of G_0. The only practical difference between the two-stage method and the projection method is that the latter uses a noncausal model in the first step of the method. As we saw above, the noncausal taps in the FIR model used in the projection method are important for obtaining asymptotically unbiased results. This example shows that the noncausal taps can also make a difference in case of finite data records. Note especially that we in this example have used a Gaussian white noise reference signal and that the nonlinearity is static, so there should be no noncausal correlation between r and u (asymptotically). Despite this, the projection method still outperforms the two-stage method in this example. The reason is that there is in fact a significant noncausal correlation between r and u, in this finite sample case, that is picked up by the noncausal FIR model.

Finally it can be noted that the direct method, applied with a correct/parameterized noise model, will retrieve the true system perfectly regardless of the controller, since the simulation is noise free. The indirect method is not easy to apply in this kind of situation, even if the controller is perfectly known. The reason is that the nonlinearity will make the estimation problem nontrivial. If we ignore the nonlinearity

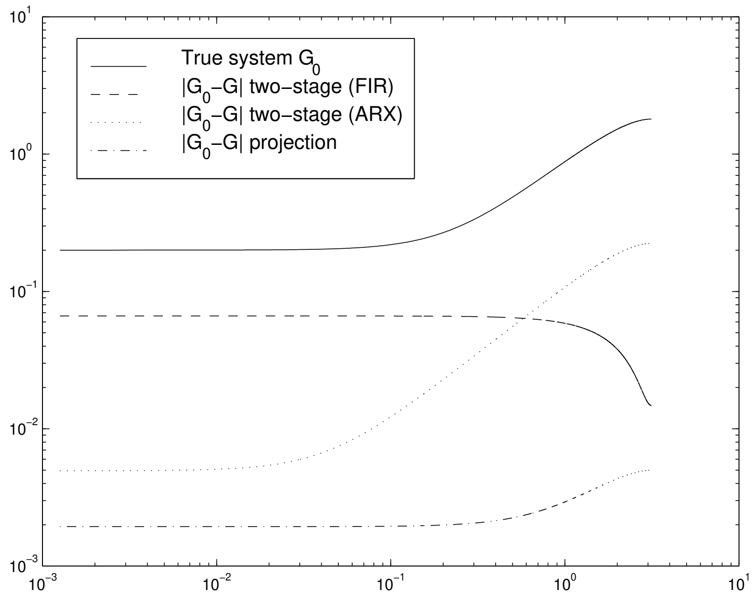

Figure 3 *Bode plot of the true system and the errors in the estimated models.*

when applying the indirect method the resulting estimate will (quite naturally) be very poor.

5.4 Unifying Framework for All Joint Input-Output Methods

It will now be shown that the joint input-output approach can be seen as a "direct" approach where both the system and the controller is identified and that the two-stage and projection methods are indeed joint input-output methods with specially chosen noise models. To achieve this we will exploit the links between the prediction error method (30)-(32) and maximum likelihood estimation.

Consider the joint system (70) and assume that d is white noise with covariance matrix Λ_d, independent of e. The maximum likelihood estimates of G_0 and H_0 are computed as

$$\min_{\theta \in D_{\mathcal{M}}} \frac{1}{N} \sum_{t=1}^{N} \begin{bmatrix} y(t) - \hat{y}(t|\theta) \\ u(t) - \hat{u}(t|\theta) \end{bmatrix}^T \begin{bmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_d \end{bmatrix}^{-1} \begin{bmatrix} y(t) - \hat{y}(t|\theta) \\ u(t) - \hat{u}(t|\theta) \end{bmatrix} \tag{88}$$

where

$$\begin{bmatrix} \hat{y}(t|\theta) \\ \hat{u}(t|\theta) \end{bmatrix} = \mathsf{H}^{-1}(q,\theta)\mathsf{G}(q,\theta)r(t) + (I - \mathsf{H}^{-1}(q,\theta)) \begin{bmatrix} y(t) \\ u(t) \end{bmatrix} \tag{89}$$

The parameterizations of G and H can be arbitrary. The relation (70) was obtained using the assumption that the noise e affect the open-loop system only and the disturbance d affect the regulator only. The natural way to parameterize H in order to reflect these assumptions in the model is

$$\mathsf{H}(q,\theta) = \begin{bmatrix} (I + G(q,\theta)K(q,\theta))^{-1}H(q,\theta) & G(q,\theta)(I + K(q,\theta)G(q,\theta))^{-1} \\ -K(q,\theta)(I + G(q,\theta)K(q,\theta))^{-1}H(q,\theta) & (I + K(q,\theta)G(q,\theta))^{-1} \end{bmatrix} \tag{90}$$

With G parameterized as

$$\mathsf{G}(q,\theta) = \begin{bmatrix} G(q,\theta)(I + K(q,\theta)G(q,\theta))^{-1} \\ (I + K(q,\theta)G(q,\theta))^{-1} \end{bmatrix} \tag{91}$$

we get

$$\begin{bmatrix} \hat{y}(t|\theta) \\ \hat{u}(t|\theta) \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix} r(t) + \begin{bmatrix} I - H^{-1}(q,\theta) & H^{-1}(q,\theta)G(q,\theta) \\ -K(q,\theta) & 0 \end{bmatrix} \begin{bmatrix} y(t) \\ u(t) \end{bmatrix} \tag{92}$$

or

$$\hat{y}(t|\theta) = (I - H^{-1}(q,\theta))y(t) + H^{-1}(q,\theta)G(q,\theta)u(t) \tag{93}$$

$$\hat{u}(t|\theta) = r(t) - K(q,\theta)y(t) \tag{94}$$

The predictor (93) is the same as for the direct method (cf. (28)), while (94) is the natural predictor for estimating the regulator K. The maximum likelihood estimate becomes

$$\min_{\theta \in D_{\mathcal{M}}} \frac{1}{N} \Bigg\{ \sum_{t=1}^{N} (y(t) - \hat{y}(t|\theta))^T \Lambda_0^{-1} (y(t) - \hat{y}(t|\theta)) + \sum_{t=1}^{N} (u(t) - \hat{u}(t|\theta))^T \Lambda_d^{-1} (u(t) - \hat{u}(t|\theta)) \Bigg\} \tag{95}$$

This shows that we may view the joint input-output method as a combination of direct identification of the open-loop system and a direct identification of the regulator. Note that this holds even in the case where $r(t) = 0$.

If the parameterization of the regulator K is independent of the one of the system the two terms in (95) can be minimized separately which decouples the two identification problems. This idea is used in the two-stage and projection methods to be studied next.

Let us return to (70). Suppose that the model (cf. (77) and (78)-(79))

$$\begin{bmatrix} \hat{y}(t|\theta) \\ \hat{u}(t|\theta) \end{bmatrix} = \begin{bmatrix} G(q,\rho) \\ I \end{bmatrix} S^i(q,\beta)r(t), \quad \theta = \begin{bmatrix} \rho \\ \beta \end{bmatrix} \tag{96}$$

is used. The maximum likelihood estimate becomes

$$\min_{\theta \in D_{\mathcal{M}}} \frac{1}{N} \left\{ \sum_{t=1}^{N} (y(t) - G(q,\rho)S^i(q,\beta)r(t))^T \Lambda_0^{-1} (y(t) - G(q,\rho)S^i(q,\beta)r(t)) \right. \\ \left. + \sum_{t=1}^{N} (u(t) - S^i(q,\beta)r(t))^T \Lambda_d^{-1} (u(t) - S^i(q,\beta)r(t)) \right\} \tag{97}$$

It will be assumed that $\Lambda_0 = \lambda_0 \cdot I$ and $\Lambda_d = \lambda_d \cdot I$. (If Λ_0 and Λ_d are known this can always be achieved by suitable prefiltering or, equivalently, by changing the noise models.) As λ_d tends to zero the maximum likelihood estimate (97) will approach the one obtained with the two-stage/projection methods. This is true because for small λ_d the β-parameters will minimize

$$\frac{1}{N} \sum_{t=1}^{N} (u(t) - S^i(q,\beta)r(t))^T (u(t) - S^i(q,\beta)r(t)) \tag{98}$$

regardless of the ρ-parameters, which then will minimize

$$\frac{1}{N} \sum_{t=1}^{N} (y(t) - G(q,\rho)S^i(q,\beta)r(t))^T (y(t) - G(q,\rho)S^i(q,\beta)r(t)) \tag{99}$$

Thus the two-stage and projection methods may be viewed as special cases of the general joint input-output approach corresponding to special choices of the noise models. This means especially that any results that hold for the joint input-output approach without constraints on the noise models, hold for the two-stage and projection methods as well. We will use this fact in Corollary 9 below.

6 Convergence Results for the Closed-loop Identification Methods

Let us now apply the result of Theorem A.1 to the special case of closed-loop identification. We will first study the direct approach. (In the following we will write G_0 as short for $G_0(e^{i\omega})$ and G_θ as short for $G(e^{i\omega},\theta)$, etc. The subscript θ is included to emphasize the parameter dependence.)

6.1 The Direct Approach

Corollary 5 *Consider the situation in Theorem A.1 and suppose that u is persistently exciting. With a model structure*

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{100}$$

we then have that, w. p. 1 as $N \to \infty$,

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\Big[\big\{(G_0 + B_\theta - G_\theta)\Phi_u(G_0 + B_\theta - G_\theta)^* \\ + (H_0 - H_\theta)\Phi_e^r(H_0 - H_\theta)^*\big\}(H_\theta \Lambda H_\theta^*)^{-1}\Big] d\omega \tag{101}$$

where

$$B_\theta = (H_0 - H_\theta)\Phi_{eu}\Phi_u^{-1} \tag{102}$$

The proof is given in Appendix B.1.

Remark: The "bias-pull" term B_θ defined in (102) can be noncausal depending, e.g., on the nature of the feedback. However, since both G_0 and G_θ are causal we can replace B_θ by a causal counterpart in (101). The exact statement is as follows. If we let $[A]_+$ denote the causal part and $[A]_-$ the anti-causal part of A we can reformulate (101) as

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\Big[\big\{(G_0 + B_\theta^+ - G_\theta)\Phi_u(G_0 + B_\theta^+ - G_\theta)^* \\ + (H_0 - H_\theta)\Phi_e^r(H_0 - H_\theta)^* + B_\theta^- \Phi_u (B_\theta^-)^*\big\}(H_\theta \Lambda H_\theta^*)^{-1}\Big] d\omega \tag{103}$$

where

$$B_\theta^+ = H_\theta \left[H_\theta^{-1}(H_0 - H_\theta)\Phi_{eu}M_u^{-*}\right]_+ M_u^{-1} \tag{104}$$

$$B_\theta^- = H_\theta \left[H_\theta^{-1}(H_0 - H_\theta)\Phi_{eu}M_u^{-*}\right]_- M_u^{-1} \tag{105}$$

A proof is included in Appendix B.1. □

If the parameterization of the model G and the noise model H is flexible enough so that $\mathcal{S} \in \mathcal{M}$ then $\bar{V}(\theta_0) = \Lambda_0$, and provided this is a unique minimum we thus have $D_c = \{\theta_0\}$. See, e. g., the discussion following Eq. (39) and Theorem 8.3 in [24].

We can also note that if the system operates in open loop, so that $\Phi_{ue} = 0$, then the bias term $B_\theta = 0$ regardless of the noise model H_θ and the limit model will be a compromise between making G_θ an as good approximation of G_0 as possible and making H_θ an as good approximation of H_0 as possible (with the indicated frequency weightings). A consequence of this is that in open loop we can use fixed noise models and still get consistent estimates of G_0 provided that $G_0 \in \mathcal{G}$ and certain identifiability conditions hold. See, e.g., Theorem 8.4 in [24].

In case of undermodeling we get a characterization of the bias error through the term B_θ in (102). A straightforward upper bound on $\bar{\sigma}(B_\theta)$ (the maximal singular value of B_θ) is

$$\bar{\sigma}(B_\theta) \leq \bar{\sigma}(H_0 - H_\theta) \cdot \bar{\sigma}(\Phi_{eu}) \cdot \frac{1}{\underline{\sigma}(\Phi_u)} \tag{106}$$

($\underline{\sigma}(\Phi_u)$ denotes the minimal singular values of Φ_u.) If the feedback is linear and given by (15) we get the following alternative upper bound:

$$\bar{\sigma}(B_\theta) \leq \bar{\sigma}(H_0 - H_\theta) \cdot \sqrt{\frac{\bar{\sigma}(\Lambda_0)}{\underline{\sigma}(\Phi_u)}} \cdot \sqrt{\frac{\bar{\sigma}(\Phi_u^e)}{\underline{\sigma}(\Phi_u)}} \tag{107}$$

It follows that the bias inclination in the G-estimate will be small in frequency ranges where either (or all) of the following holds:

- The noise model is good ($\bar{\sigma}(H_0 - H_\theta)$ is small).
- The feedback noise contribution to the input spectrum ($\bar{\sigma}(\Phi_u^e)/\underline{\sigma}(\Phi_u)$) is small.
- The signal-to-noise ratio is high ($\bar{\sigma}(\Lambda_0)/\underline{\sigma}(\Phi_u)$ is small).

In particular, if a reasonable flexible, independently parameterized noise model is used the bias-inclination of the G-estimate can be small.

6.2 The Indirect Approach

We will now give the corresponding convergence result for the indirect method. Here we assume that this method is implemented using a model structure of the form (49) where G^c is parameterized as in (52).

Corollary 6 *Consider the situation in Theorem A.1 and suppose that r is persistently exciting and that the input u is given by* (25). *Then, if the model structure is*

$$y(t) = G^c(q, \theta) r(t) + H_*(q) e(t) \tag{108}$$

we have that, w. p. 1 as $N \to \infty$,

$$\hat{\theta}_N \to D_c = \arg \min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\left[(G_0^c + \bar{B} - G_\theta^c)\Phi_r (G_0^c + \bar{B} - G_\theta^c)^* (H_* \Lambda H_*^*)^{-1}\right] d\omega \tag{109}$$

where

$$\bar{B} = G_0^c \Phi_{dr} \Phi_r^{-1} \tag{110}$$

The proof is given in Appendix B.2.

Remark: Since we consider causal systems and models we can alternatively write (109) as

$$\hat{\theta}_N \to D_c = \arg \min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\left[(G_0^c + \bar{B}^+ - G_\theta^c)\Phi_r (G_0^c + \bar{B}^+ - G_\theta^c)^* (H_* \Lambda H_*^*)^{-1}\right] d\omega \tag{111}$$

where

$$\bar{B}^+ = H_* \left[H_*^{-1} G_0^c \Phi_{dr} M_r^{-*}\right]_+ M_r^{-1} \tag{112}$$

This is shown in Appendix B.2. □

From (109) we can draw a number of conclusions regarding the limit estimates for the indirect method. We remember that the closed-loop model G^c is parameterized as $G^c(q,\theta) = (I + G(q,\theta)K(q))^{-1}G(q,\theta)$.

If d is uncorrelated with r then $\Phi_{dr} = 0$ and $\bar{B} = 0$ (cf. (110)), and the G-estimate will try to minimize the norm of

$$G_0^c - G_\theta^c = G_0^c - S_\theta G_\theta = S_0(G_0 - G_\theta)S_\theta^i \tag{113}$$

with the indicated frequency weighting. This shows that the indirect method can give consistent estimates of G_0 if the parameterization of G is flexible enough, that is, if $G_0 \in \mathcal{G}$, even with a fixed noise model. This is the main advantage of the indirect method. In case of under-modeling the resulting G-estimate will try to minimize the mismatch between G_0 and G and at the same time try to minimize the model sensitivity function S^i. There will thus be a "bias-pull" towards transfer functions that give a small sensitivity for the given regulator, which can be advantageous, e.g., if the model is to be used for control [12, 32]. Unlike (101) it is not easy to quantify this bias error in the open-loop model, though.

If d and r are correlated then $\bar{B} \neq 0$ and the G-estimate will try to minimize the (frequency weighted) norm of

$$G_0^c + \bar{B} - (I + G_\theta K)^{-1} G_\theta \tag{114}$$

which can lead to an arbitrarily bad open-loop estimate depending on $\bar{B}$. This situation will, e.g., occur if the assumed controller parameters deviate from the true ones or if the true controller is nonlinear. See the discussion in Section 2. (In case we know u in addition to y and r there is a simple remedy for this problem, since then we can construct $r' = u + Ky$ $(= r + d)$ and use this signal instead of r which would eliminate the bias.)

6.3 The Joint Input-output Approach

Finally, we will treat the joint input-output approach. Here we will study two cases, first a "classical" joint input-output method, where the final G estimate is found as in (72)-(74), and, second, the two-stage/projection method.

Corollary 7 *Consider the situation in Theorem A.1 and suppose that r is persistently exciting and that the input u is given by* (25)*. Then*

1. *if the model structure is*

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} G^c(q,\theta) \\ S^i(q,\theta) \end{bmatrix} r(t) + \mathsf{H}_*(q) \begin{bmatrix} e(t) \\ d(t) \end{bmatrix} \tag{115}$$

we have that, w. p. 1 as $N \to \infty$,

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \mathrm{tr}\Big[\begin{bmatrix} G_0^c \cdot D - G^c(q,\theta) \\ S_0^i \cdot D - S^i(q,\theta) \end{bmatrix} \Phi_r \begin{bmatrix} G_0^c \cdot D - G^c(q,\theta) \\ S_0^i \cdot D - S^i(q,\theta) \end{bmatrix}^* (\mathsf{H}_* \Lambda \mathsf{H}_*^*)^{-1} \Big] d\omega \tag{116}$$

where

$$D = I + \Phi_{dr}\Phi_r^{-1} \tag{117}$$

2. *if the model structure is*

$$y(t) = G(q,\theta)\hat{u}(t) + H_*(q)e(t), \quad \hat{u}(t) = \hat{S}_N(q)r(t), \;\; \tilde{u}(t) = u(t) - \hat{u}(t) \tag{118}$$

we have that, w. p. 1 as $N \to \infty$,

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \mathrm{tr}\big[(G_0 + \tilde{B} - G_\theta)\Phi_{\hat{u}}(G_0 + \tilde{B} - G_\theta)^* (H_* \Lambda H_*^*)^{-1} \big] d\omega \tag{119}$$

where

$$\tilde{B} = G_0 \Phi_{\tilde{u}\hat{u}} \Phi_{\hat{u}}^{-1} \tag{120}$$

The proof is given in Appendix B.3.

Remark: We could also rewrite (116) and (119) in similar forms as (111), but for brevity we do not give the details here. □

Let us first comment on the result (116). We assume that the open-loop estimate is computed as $\hat{G}_N = \hat{G}_N^c(\hat{S}_N^i)^{-1}$. If $D = I$ and the parameterizations of G^c

and S^i are flexible enough, then this method will give consistent estimates of G_0, since $\hat{G}_N = \hat{G}^c_N(\hat{S}^i_N)^{-1} \to G^c_0(S^i_0)^{-1} = G_0$ as N tends to infinity, regardless of the noise models used, just as for the indirect method. An advantage compared to the indirect method is that the controller need not be known. We can also note that if the parameterizations of G^c and S^i are such that $\hat{G}^c_N \to G^c_0 D$ and $\hat{S}^i_N \to S^i_0 D$ as N tends to infinity, then this joint input-output method gives consistent G-estimates regardless of the disturbance d. This follows since $G^c_0 D(S^i_0 D)^{-1} = G_0$ regardless of D.

The result (119) holds for both the two-stage and the projection method. It is clear that any correlation between $\hat{u}$ and $\tilde{u}$ will deteriorate the G-estimate in these methods. The error is quantified by the term $\tilde{B}$ defined in (120). With the projection method $\Phi_{\tilde{u}\hat{u}} = 0$ is (asymptotically) achieved, by construction, which implies that $\tilde{B} = 0$. Hence, with this method it is possible to obtain consistent estimates of G_0 with fixed noise models (prefilters), regardless of the feedback just as with the direct method. An important advantage compared to the direct method is that with the projection method we can guarantee consistency even if a fixed noise model is used. This also means that in case of undermodeling the model can be fit to the data with arbitrary frequency weighting with this method.

7 Asymptotic Variance of Black Box Transfer Function Estimates

Consider first the direct approach. The following result follows directly from Theorem A.2 after noting that

$$\Phi_{\chi_0}^{-1} = \begin{bmatrix} (\Phi_u^r)^{-1} & -(\Phi_u^r)^{-1}\Phi_{ue}\Lambda_0^{-1} \\ -\Lambda_0^{-1}\Phi_{eu}(\Phi_u^r)^{-1} & (\Phi_e^r)^{-1} \end{bmatrix} \tag{121}$$

Corollary 8 *Suppose that the conditions in Theorem A.2 hold. Then*

$$\sqrt{N}\,\mathrm{vec}\big[\hat{G}_N(e^{i\omega}, n, \delta) - G^*_n(e^{i\omega})\big] \in AsN(0, P(\omega, n, \delta)) \quad \text{as } N \to \infty \text{ for fixed } n,\ \delta \tag{122}$$

where

$$\lim_{\delta\to 0}\lim_{n\to\infty} \frac{1}{n} P(\omega, n, \delta) = (\Phi_u^r(\omega))^{-T} \otimes \Phi_v(\omega) \tag{123}$$

(The notation is explained in Appendix A.2.)

An intuitive, but not formally correct, interpretation of (122), (123) is

$$\mathrm{Cov}\,\mathrm{vec}\big[\hat{G}_N(e^{i\omega})\big] \sim \frac{n}{N}(\Phi_u^r(\omega))^{-T} \otimes \Phi_v(\omega) \tag{124}$$

as the model order n as well as the number of data N tends to infinity.

The difference between (124) (in the sequel we will use (124) when referring to the result in Corollary 8) and the corresponding open-loop expression (44) is thus that Φ_u^r replaces Φ_u, or, in other words, that the part of the input spectrum that originates from the reference signal r replaces the total input spectrum.

The expression (124) – which also is the asymptotic Cramèr-Rao lower limit – tells us precisely "the value of information" of closed-loop experiments. It is the noise-to-signal ratio (where "signal" is what derives from the injected reference) that determines how well the open-loop transfer function can be estimated. From this perspective, that part of the input that originates from the feedback noise has no information value when estimating G.

If we in the direct method use a fixed noise model equal to the true one, $H(q,\theta) = H_*(q) = H_0(q)$, we have that

$$\operatorname{Cov}\operatorname{vec}\left[\hat{G}_N(e^{i\omega}, n)\right] \sim \frac{n}{N}(\Phi_u(\omega))^{-T} \otimes \Phi_v(\omega) \tag{125}$$

as the model order tends to infinity, just as in the open-loop case. This is true since the covariance matrix is determined entirely by the second order properties (the spectrum) of the input, and it is immaterial whether this spectrum is a result of open-loop or closed-loop operation. We emphasize that this alternative result holds only if we use a *fixed* and correct noise model. The importance of this for the improved accuracy will be further discussed below.

It should be noted that (124) holds for the direct method if $\mathcal{S} \in \mathcal{M}$. If we restrict to linear feedback as in (15) we have $\Phi_u^r = S_0^i \Phi_r (S_0^i)^*$. In this case (124) also holds for the indirect and joint input-output approaches if $G_0 \in \mathcal{G}$, regardless of the parameterizations and the noise models used. This follows from

Corollary 9 *Suppose that the input is generated as in* (15). *If the estimate* $\hat{G}_N$ *is found using either the indirect or the joint input-output approach and the conditions in Corollary A.3 hold. Then* (122), (123) *holds also for these methods.*

The proof is given in Appendix B.4.

We stress that in Corollary 9 we restricted to linear feedback. Hence, Φ_u^r in (123) should really be replaced by $\Phi_u^r = S_0^i \Phi_r (S_0^i)^*$ in this case.

It can also be noted that the results in, for example, (43), (44), and (124) are independent of the parameterization and an important assumption is that the model order tends to infinity. It is then quite natural that the differences between the closed-loop identification methods disappear, since, as we have seen, they may be viewed as different parameterizations of the same general prediction error method.

8 Asymptotic Distribution of Parameter Vector Estimates in the Case of Closed-loop Data

8.1 The Direct and Indirect Approaches

We will first consider the case when the dynamics and noise models are independently parameterized. The two following corollaries to Theorem A.4 in Appendix A.3 states what the asymptotic distribution of the parameter estimates will be for the direct and indirect approaches, respectively, in their standard formulations. The results will be given for the SISO case only and in the following G'_ρ will be used as short for $(d/d\rho)G(e^{i\omega},\rho)|_{\rho=\rho_0}$, etc. λ_0 denotes the variance of e in the SISO case.

Corollary 10 *Assume that the conditions in Theorem A.4 hold and that the model structure $\mathcal{M}$ is*

$$y(t) = G(q,\rho)u(t) + H(q,\beta)e(t), \quad \theta = \begin{bmatrix}\rho\\ \beta\end{bmatrix} \tag{126}$$

Suppose that $\mathcal{S} \in \mathcal{M}$ and that $\mathcal{M}$ is globally identifiable at θ_0. Suppose also that the regulator is given by (15) and that the data set Z^∞ is informative enough with respect to $\mathcal{M}$. Then

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P^D_\rho) \tag{127a}$$

where

$$P^D_\rho = \lambda_0(R^r_\rho + \Delta)^{-1} \tag{127b}$$

$$R^r_\rho = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\Phi^r_u}{|H_0|^2} G'_\rho G'^*_\rho \, d\omega \tag{127c}$$

$$\Delta = R^e_\rho - R^e_{\rho\beta}(R^e_\beta)^{-1}R^e_{\beta\rho} \tag{127d}$$

$$R^e_\rho = \frac{1}{2\pi}\int_{-\pi}^{\pi} \lambda_0|K|^2|S_0|^2 G'_\rho G'^*_\rho \, d\omega \tag{127e}$$

$$R^e_{\rho\beta} = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\lambda_0}{|H_0|^2} K S_0 H_0 G'_\rho H'^*_\beta \, d\omega \tag{127f}$$

$$R^e_\beta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\lambda_0}{|H_0|^2} H'_\beta H'^*_\beta \, d\omega \tag{127g}$$

Furthermore,

$$0 \le \Delta \le R^e_\rho \tag{128}$$

and Δ approaches zero as the order of the noise model tends to infinity, while $\Delta = R^e_\rho$ if the noise model is fixed and equal to the true one, that is, if $H(q,\beta) = H_*(q) = H_0(q)$.

The proof is given in Appendix B.5.

Corollary 11 *Consider the same situation as in Corollary 10 but assume that the model structure is*

$$y(t) = \frac{G(q,\rho)}{1+G(q,\rho)K(q)} r(t) + H^c(q,\beta)e(t), \quad \theta = \begin{bmatrix} \rho \\ \beta \end{bmatrix} \tag{129}$$

Then

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P_\rho^I) \tag{130a}$$

where

$$P_\rho^I = \lambda_0 (R_\rho^r)^{-1} \tag{130b}$$

with R_ρ^r *given by* (127c)

If the noise model in (129) *is fixed* $H^c(q,\beta) = H_*(q)$ *and the model structure now satisfies* $G_0 \in \mathcal{G}$ *and is globally identifiable at* ρ_0, *we have that*

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, \bar{P}_\rho^I) \tag{131a}$$

where

$$\bar{P}_\rho^I = R^{-1} Q R^{-1} \tag{131b}$$

$$R = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|S_0|^2 \Phi_u^r}{|H_*|^2} G'_\rho G'^*_\rho \, d\omega \tag{131c}$$

$$Q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|S_0|^4 \Phi_v \Phi_u^r}{|H_*|^4} G'_\rho G'^*_\rho \, d\omega \tag{131d}$$

Proof *The main result is a consequence of Theorem A.4. The expressions for the covariance matrices follow from the corresponding open-loop expressions (see, e.g., Equation (9.55) in [24]) after noting that* $(G_\rho^c)' = S_0^2 G'_\rho$ *and* $H_0^c = S_0 H_0$. □ □

An important observation regarding the result (127) is that the term Δ is entirely due to the noise part of the input spectrum and since $\Delta \geq 0$ this contribution has a positive effect on the accuracy, contrarily to what one might have guessed. We conclude that in the direct method the noise in the loop is utilized in reducing the variance. For the indirect methods this contribution is zero as can be seen from (130). The reason for this is that in the indirect method only that part of the input that originates from the reference signal is used in the identification.

From (127) it is also clear that – from the accuracy point of view – the worst-case experimental conditions with respect to the reference signal is when there is *no* external reference signal present, that is, when $\Phi_r = 0$. In this case

$$Cov\ \hat{\rho}_N \sim \frac{\lambda_0}{N}\Delta^{-1} \tag{132}$$

Thus Δ characterizes the lower limit of achievable accuracy for the direct method. Now, if Δ is nonsingular, (127) says that we can consistently estimate the system parameters using the direct method even though no reference signal is present. The exact conditions for this to happen are given in [26] for some common special cases. However, even if Δ is singular it will have a beneficial effect on the variance of the estimates, according to (127). Only when $\Delta = 0$ there is no positive effect from the noise source on the accuracy of the estimates. According to Corollary 10 this would be the case when the order of the noise model tends to infinity. At the other extreme we have that a fixed (and correct) noise model will make Δ as "large" as it can be.

This puts the finger on the value of information in the noise source e for estimating the dynamics:

> *It is the knowledge/assumption of a constrained noise model that improves the estimate of G_0.*

This also explains the difference between (124) (which assumes the noise model order to tend to infinity) and (125) (which assumes a fixed and correct noise model).

For the indirect approach, we can further note that for all fixed H_*,

$$P_\rho^{\mathrm{I}} \leq \bar{P}_\rho^{\mathrm{I}} \tag{133}$$

with equality for $H_* = H_0^c = S_0 H_0$. By comparing (130) with (127) we also see that the covariance matrix for the indirect method will always be larger than the one for the direct. The difference is the term Δ that is missing in (130). We thus have the following ranking:

$$P_\rho^{\mathrm{D}} \leq P_\rho^{\mathrm{I}} \leq \bar{P}_\rho^{\mathrm{I}} \tag{134}$$

8.2 Further Results for the Indirect Approach

In Corollary 11 we considered two particular variants of the indirect approach. It is important to realize that the conclusions drawn in the previous section do not hold for all possible variants and parameterizations of the indirect approach. In this section we will study two other variants where first the closed-loop system is identified and then the open-loop system is estimated by solving an over-determined system of equations, relating the open-loop parameters to the closed-loop parameters, in a least squares sense using the knowledge of the controller (cf. the discussion in Section 4.2). We will then see that it is possible to get the same level of accuracy with the indirect method as with the direct, even with finite model orders. This was first shown in [27] and in this section we will review and extend the results there. See also [15] and [29].

Indirect Identification Using ARMAX Models

Let us now outline the idea. Suppose that the regulator is linear and of the form (15), that is

$$u(t) = r(t) - K(q)y(t) \tag{135}$$

Factorize K as

$$K(q) = \frac{X(q)}{Y(q)} \tag{136}$$

where the polynomials X and Y are coprime. Let the closed-loop and open-loop model structures be

$$A^c(q)y(t) = B^c(q)r(t) + C^c(q)e(t) \tag{137}$$

and

$$A(q)y(t) = B(q)u(t) + C(q)e(t) \tag{138}$$

respectively. From

$$G^c(q) = \frac{G(q)}{1 + G(q)K(q)} \text{ and } H^c(q) = \frac{H(q)}{1 + G(q)K(q)} \tag{139}$$

(cf. (21)) it follows that we may solve for A, B and C in

$$\begin{cases} A^c(q) = A(q)Y(q) + B(q)X(q) \\ B^c(q) = B(q)Y(q) \\ C^c(q) = C(q)Y(q) \end{cases} \tag{140}$$

Let η and θ denote the closed-loop and open-loop parameter vectors, respectively. Then, for given η, (140) is a system of linear equations in θ from which θ can be estimated. If the Markov estimate is used, that is, if (140) is solved in a least squares sense, using the estimated covariances of $\hat{\eta}_N$ as weights, Corollary 12 below states that this indirect method is consistent under reasonable assumptions and gives the same accuracy as the direct method with an ARMAX model structure.

Corollary 12 *Consider the situation in Theorem A.4 and let the model structure $\mathcal{M}$ be given by (137). Assume that $\mathcal{M}$ is identifiable at η_0 and that the data set $Z^\infty = \{y(1), r(1), \ldots, y(\infty), r(\infty)\}$ is informative enough with respect to $\mathcal{M}$. Assume also that $\mathcal{S} \in \mathcal{M}$ holds for the model structure (138). Then, if the open-loop parameters θ are estimated from (140) using the Markov estimate, we have that*

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, P_\theta^{AX}) \tag{141a}$$

where

$$P_\theta^{AX} = \lambda_0 \left[\bar{E}\psi(t,\theta_0)\psi^T(t,\theta_0)\right]^{-1} \tag{141b}$$

$$\psi(t,\theta) = -\frac{d}{d\theta}\frac{A(q)}{C(q)}\left(y(t) - \frac{B(q)}{A(q)}u(t)\right) \tag{141c}$$

The proof is given in Appendix B.6.

We emphasize that (141) is exactly what we would get with the direct method applied using an ARMAX model. Through this corollary we have therefore established yet another link between the direct and indirect approaches. As will become more apparent below, the fact that the dynamics *model* and the noise *model* share the same poles is crucial for obtaining the same accuracy as with the direct approach.

Indirect Identification Using Box-Jenkins Models

Suppose now that we in the first step in the indirect method use the following model structure with an independently parameterized noise model:

$$\begin{aligned} y(t) &= G^c(q,\rho_\eta)r(t) + H^c(q,\beta_\eta)e(t) \\ &= \frac{B^c(q)}{F^c(q)}r(t) + \frac{C^c(q)}{D^c(q)}e(t) \end{aligned} \tag{142}$$

This is also known as a Box-Jenkins model structure. Let the open-loop model be

$$\begin{aligned} y(t) &= G(q,\rho)u(t) + H(q,\beta)e(t) \\ &= \frac{B(q)}{F(q)}r(t) + \frac{C(q)}{D(q)}e(t) \end{aligned} \tag{143}$$

To find the indirect estimate of G_0 we may solve for B and F in (cf. (140))

$$\begin{cases} B^c(q) = B(q)Y(q) \\ F^c(q) = F(q)Y(q) + B(q)X(q) \end{cases} \tag{144}$$

The final estimate will also this time be consistent (under mild conditions) but not of optimal accuracy. Instead we get

Corollary 13 *Consider the situation in Theorem A.4 and let the model structure $\mathcal{M}$ be given by (142). Assume that $\mathcal{M}$ is identifiable at η_0 and that the data set $Z^\infty = \{y(1), r(1), \dots, y(\infty), r(\infty)\}$ is informative enough with respect to $\mathcal{M}$. Assume also that $G_0 \in \mathcal{G}$ holds for the model structure (143). Then, if the open-loop parameters θ are estimated from (144) using the Markov estimate, we have that*

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P_\rho^{BJ}) \tag{145a}$$

where

$$P_\rho^{BJ} = \lambda_0 \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Phi_u^r}{|H_0|^2} G'_\rho G'^*_\rho d\omega \right]^{-1} \tag{145b}$$

The proof is given in Appendix B.7.

Thus

$$P_\rho^{\mathrm{BJ}} = P_\rho^{\mathrm{I}} \; (\geq P_\rho^{\mathrm{D}}) \tag{146}$$

and once again we conclude that with an independently parameterized noise model the indirect method gives worse accuracy than the direct method. The difference is quantified by the term Δ (cf. (127)) which is missing in (145).

8.3 Variance Results for the Projection Method

In this section some results for the efficiency of the projection method will be presented. Exact analysis of the asymptotic variance of the parameter estimates is not easy in the general case and we will not go into any further details on this here. However, if we consider the model (cf. (79))

$$y(t) = G(q, \rho)\hat{u}(t) + H_*(q)e(t) \tag{147}$$

and rewrite the true system (2) as

$$y(t) = G_0(q)\hat{u}(t) + w(t) \tag{148}$$

where

$$w(t) = v(t) + G_0(q)(u(t) - \hat{u}(t)) \tag{149}$$

we can nevertheless gain some insight into the variance properties of the projection method.

With the projection method, $\hat{u}$ and w in (148) are uncorrelated. We may thus regard the second step of the projection method as an open-loop identification of

G_0 using measurements of y and $\hat{u}$. The difference the second step of the projection method and direct identification of G_0 using measurements of y and u is that $\hat{u}$ replaces u and w replaces v. Typically we will have

$$\Phi_{\hat{u}} < \Phi_u \quad \text{and} \quad \Phi_w > \Phi_v \tag{150}$$

implying that

$$\frac{\Phi_{\hat{u}}}{\Phi_w} < \frac{\Phi_u}{\Phi_v} \tag{151}$$

This shows that the signal-to-noise ratio will be worse than for direct identification using y and u, and we can thus expect that the accuracy of the projection method will be suboptimal. See also [11].

If u is generated as in (15) and $\hat{u} = S_0 r$ (SISO) we have that $w = S_0 v$ and we are back in an open-loop situation and all standard open-loop results hold (since r and v are independent). As an example we have the following consequence of Theorem A.4 (G'_ρ is short for $d/d\rho G(e^{i\omega}, \rho)|_{\rho=\rho_0}$):

Corollary 14 *Consider the estimate $\hat{\rho}_N$ defined by* (30)-(32) *where ε is determined from the model structure* (147) *which we assume is uniformly stable and satisfies $G_0 \in \mathcal{G}$. Suppose that Assumptions 1 and 2 hold and that the input is generated as in* (15) *and that $\hat{u} = S_0 r$. Suppose also that for a unique value ρ_0 interior to $D_{\mathcal{M}}$ we have*

$$\hat{\rho}_N \to \rho_0, \quad \text{w.p. } 1 \text{ as } N \to \infty \tag{152}$$

$$R = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\Phi_{\hat{u}}}{|H_*|^2} G'_\rho G'^*_\rho \, d\omega > 0 \tag{153}$$

$$\sqrt{N} E\left[\frac{1}{N}\sum_{t=1}^{N}[\psi(t,\rho_0)\varepsilon(t,\rho_0) - \bar{E}\psi(t,\rho_0)\varepsilon(t,\rho_0)]\right] \to 0, \text{ as } N \to \infty \tag{154}$$

where ψ denotes the negative gradient of ε. Then

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P_\rho) \tag{155a}$$

$$P_\rho = R^{-1} Q R^{-1} \tag{155b}$$

$$Q = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\Phi_w \Phi_{\hat{u}}}{|H_*|^4} G'_\rho G'^*_\rho \, d\omega \tag{155c}$$

9 Summarizing Discussion

We have attempted to give a status report on identification of closed-loop systems. Several methods have been studied and we have shown that most of the common

methods can be viewed as special parameterizations of the general prediction error method. This allows simultaneous analysis of the statistical properties of the methods using the general prediction error theory. The main results can be summarized as follows.
The direct approach

- gives consistency and optimal accuracy, regardless of the feedback, if the noise and dynamic models contain a true description of the system.
- requires a correct noise model (possibly estimated) to avoid bias in the dynamics transfer function for closed-loop data (Corollary 5). For this reason, user chosen weightings for custom-made bias distribution cannot be used.

The indirect approach

- requires perfect knowledge of the regulator but can be applied with fixed noise models and still guarantee consistency (Corollary 6).
- is equivalent to the direct approach applied with a noise model containing the regulator (Lemma 3).

The joint input-output approach

- gives consistent estimates of the open-loop system regardless of the noise models used, as long as the feedback has a certain (linear) structure.
- can be seen as a combination of direct identification of the open-loop system and of the regulator.

The projection method

- allows the true system to be approximated in an arbitrary, user chosen frequency domain norm regardless of the feedback (Corollary 7).

Further comments:

1. As the model orders tend to infinity, the asymptotic variance expressions are the same for all methods in case of linear feedback (Corollaries 8 and 9).
2. For the finite model order case, the direct method meets the Cramèr-Rao bound and is thus optimal. The indirect method generally gives worse accuracy. The improved accuracy for the direct method can be traced to the fact that a constrained noise model (that includes the true noise description) allows some leverage from the noise part of the input to estimate the dynamics (Corollaries 10 and 11, and Corollaries 12 and 13).

The optimal statistical properties, the simplicity, and the general applicability of the direct method implies that this should be seen as the first choice of methods for closed-loop identification. When this method fails no other method will succeed. In the literature several other methods, e.g., the indirect method, have been advocated mainly due to their ability to shape the bias distribution by use of fixed noise models (prefilters). However, a drawback with most of these methods is that they cannot be applied to systems with arbitrary/unknown feedback mechanisms (the indirect method, e.g., requires a perfectly known feedback law to be applicable). An exception is the projection method that works regardless of the nature of the feedback. Another drawback with the indirect and joint input-output methods is that they generally give suboptimal accuracy.

If a low order model is sought that should approximate the system dynamics in a pre-specified frequency norm there are two ways to proceed:

1. First estimate a high order model, with small bias, using the direct approach and then reduce this model to lower order using the proper frequency weighting.
2. Use the projection method.

Compared to the former approach, the projection method has the advantage that a low order model can be fitted directly to the data with arbitrary frequency weighting. However, if the high order model used in the first approach is (essentially) bias free this will also give us explicit information about the error in the low order model, which favors this approach.

Acknowledgment

The authors wish to thank the anonymous reviewers for their many helpful comments and suggestions for improvements of the manuscript.

References

[1] H. Akaike. Some problems in the application of the cross-spectral method. In B. Harris, editor, *Spectral Analysis of Time Series*, pages 81–107. John Wiley & Sons, 1967.

[2] B. D. O. Anderson and M. Gevers. Identifiability of linear stochastic systems operating under linear feedback. *Automatica*, 18:195–213, 1982.

[3] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Information and System Sciences Series. Prentice-Hall, 1979.

[4] K. J. Åström. Matching criteria for control and identification. In *Proceedings of the 2nd European Control Conference*, pages 248–251, Groningen, The Netherlands, 1993.

[5] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Holden-Day, 1981.

[6] R. de Callafon and P. Van den Hof. Multivariable closed-loop identification: From indirect identification to Dual-Youla parametrization. In *Proceedings of the 35th Conference on Decision and Control*, pages 1397–1402, Kobe, Japan, 1996.

[7] R. de Callafon, P. Van den Hof, and M. Steinbuch. Control relevant identification of a compact disc pick-up mechanism. In *Proceedings of the 32nd Conference on Decision and Control*, volume 3, pages 2050–2055, San Antonio,TX, 1993.

[8] B. Egardt. On the role of noise models for approximate closed loop identification. In *Proceedings of the 4th European Control Conference*, Brussels, Belgium, 1997.

[9] U. Forssell and L. Ljung. Identification for control: Some results on optimal experiment design. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3384–3389, Tampa, FL, 1998. Submitted to Automatica.

[10] U. Forssell and L. Ljung. Identification of unstable systems using output error and Box-Jenkins model structures. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3932–3937, Tampa, FL, 1998. To appear in IEEE Transactions on Automatic Control.

[11] U. Forssell and L. Ljung. A projection method for closed-loop identification. *IEEE Transactions on Automatic Control*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-1984, Linköping University, Linköping, Sweden.

[12] M. Gevers. Towards a joint design of identification and control. In H. L. Trentelman and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 111–151. Birkhäuser, 1993.

[13] M. Gevers and L. Ljung. Optimal experiment design with respect to the intended model application. *Automatica*, 22:543–554, 1986.

[14] A. Graham. *Kronecker Products and Matrix Calculus: with Applications*. Ellis Horwood Limited, 1981.

[15] I. Gustavsson, L. Ljung, and T. Söderström. Identification of processes in closed loop — Identifiability and accuracy aspects. *Automatica*, 13:59–75, 1977.

[16] F. R. Hansen. *A Fractional Representation to Closed-loop System Identification and Experiment Design*. PhD thesis, Stanford University, Stanford, CA, USA, 1989.

[17] F. R. Hansen, G. F. Franklin, and R. Kosut. Closed-loop identification via the fractional representation: experiment design. In *Proceedings of the American Control Conference*, pages 1422–1427, Pittsburg, PA, 1989.

[18] H. Hjalmarsson, M. Gevers, and F. De Bruyne. For model-based control design, closed loop identification gives better performance. *Automatica*, 32:1659–1673, 1996.

[19] T. Kailath. *Lectures on Wiener and Kalman Filtering*. Springer-Verlag, 1981.

[20] I. D. Landau and K. Boumaiza. An output error recursive algorithm for identification in closed loop. In *Proceedings of the 13th IFAC World Congress*, volume I, pages 215–220, San Francisco, CA, 1996.

[21] W. S. Lee, B. D. O. Anderson, I. M. Y. Mareels, and R. L. Kosut. On some key issues in the windsurfer approach to adaptive robust control. *Automatica*, 31:1619–1636, 1995.

[22] L. Ljung. Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, 23:770–783, 1978.

[23] L. Ljung. Asymptotic variance expressions for identified black-box transfer function models. *IEEE Transactions on Automatic Control*, 30:834–844, 1985.

[24] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[25] L. Ljung and U. Forssell. An alternative motivation for the indirect approach to closed-loop identification. *IEEE Transactions on Automatic Control*, to appear. Preliminary version available as Technical Report LiTH-ISY-R-1989, Linköping University, Linköping, Sweden.

[26] T. Söderström, I. Gustavsson, and L. Ljung. Identifiability conditions for linear systems operating in closed loop. *International Journal of Control*, 21:234–255, 1975.

[27] T. Söderström, L. Ljung, and I. Gustavsson. On the accuracy of identification and the design of identification experiments. Technical Report 7428, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, 1974.

[28] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall International, 1989.

[29] T. Söderström, P. Stoica, and B. Friedlander. An indirect prediction error method for system identification. *Automatica*, 27:183–188, 1991.

[30] P. Van den Hof. Closed-loop issues in system identification. In *Procedings of the 11th IFAC Symposium on System Identification*, volume 4, pages 1651–1664, Fukuoka, Japan, 1997.

[31] P. M. J. Van den Hof and R. J. P. Schrama. An indirect method for transfer function estimation from closed loop data. *Automatica*, 29:1523–1527, 1993.

[32] P. M. J. Van den Hof and R. J. P. Schrama. Identification and control — Closed-loop issues. *Automatica*, 31:1751–1770, 1995.

[33] P. M. J. Van den Hof, R. J. P. Schrama, R. A. de Callafon, and O. H. Bosgra. Identification of normalized coprime factors from closed-loop experimental data. *European Journal of Control*, 1:62–74, 1995.

[34] E. T. van Donkelaar and P. M. J. Van den Hof. Analysis of closed-loop identification with a taylor-made parametrization. In *Proceedings of the 4th European Control Conference*, Brussels, Belgium, 1997.

[35] P. Van Overschee and B. De Moor. *Subspace identification for linear systems*. Kluwer, 1996.

[36] M. Vidyasagar. *Control System Synthesis: A Factorization Approach*. MIT Press, 1985.

[37] Z. Zang, R. R. Bitmead, and M. Gevers. Iterative weighted least-squares identification and weighted LQG control design. *Automatica*, 31:1577–1594, 1995.

[38] Y.-C. Zhu. Black-box identification of MIMO transfer functions: Asymptotic properties of prediction error models. *International Journal of Adaptive Control and Signal Processing*, 3:357–373, 1989.

A Theoretical Results for the Prediction Error Method

A.1 Complement to Section 3.2

Here we will prove (39). Let us formalize the result in a theorem:

Theorem A.1 *Let $\hat{\theta}_N$ be defined by* (30)-(32) *where $L(q,\theta) = I$. Suppose that the model structure* (27) *is uniformly stable and that Assumptions 1 and 2 hold. Then, w. p.* 1 *as $N \to \infty$,*

$$\hat{\theta}_N \to D_c =$$
$$\arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \operatorname{tr}\Big[\big[(G_0 - G_\theta) \quad (H_0 - H_\theta)\big] \Phi_{\chi_0} \begin{bmatrix} (G_0 - G_\theta)^* \\ (H_0 - H_\theta)^* \end{bmatrix} (H_\theta \Lambda H_\theta^*)^{-1}\Big] d\omega \tag{A.1}$$

Proof *From [22] we have that $\hat{\theta}_N$ will tend to the minimizing argument of $\bar{V}(\theta)$, w. p.* 1 *as $N \to \infty$. This is* (37). *To derive the expression* (A.1) *we need to derive an expression for the spectrum of the prediction error ε, cf.* (38). *Write H_θ as short for $H(e^{i\omega}, \theta)$, etc. Then, using* (2) *we get*

$$\varepsilon = H_\theta^{-1}(y - G_\theta u) = H_\theta^{-1}\big[(G_0 - G_\theta)u + (H_0 - H_\theta)e\big] + e \tag{A.2}$$

The last term is independent of the rest since $(G_0(q) - G(q,\theta))u(t)$ depends only on $e(s)$ for $s < t$ (by assumption) and $H_0(q) - H(q,\theta)$ contains a delay (since both H_0 and H are monic). Using (8) *the spectrum of ε thus becomes*

$$\Phi_\varepsilon = H_\theta^{-1}\big[(G_0 - G_\theta) \;\; (H_0 - H_\theta)\big]\Phi_{\chi_0}\begin{bmatrix}(G_0 - G_\theta)^* \\ (H_0 - H_\theta)^*\end{bmatrix} H_\theta^{-*} + \Lambda_0 \tag{A.3}$$

The result now follows since the last term is independent of θ and $\mathrm{tr}[AB] = \mathrm{tr}[BA]$ for any matrices A and B of appropriate dimensions. □

A.2 Complement to Section 3.3

In this appendix we will give the exact statements of the results (43)-(45), including the necessary technical assumptions. The main results are given in Theorem A.2 and Corollary A.3 below.

First of all the vec-operator introduced in (40) is defined as ($A_{i,j}$ is the (i,j) element of the $(n \times m)$ matrix A)

$$\mathrm{vec}\big[A\big] = \big[A_{1,1}, A_{2,1}, \ldots, A_{n,1}, A_{1,2}, A_{2,2}, \ldots, A_{n,m}\big]^T \tag{A.4}$$

The model structure considered is

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{A.5}$$

Let

$$\theta^*(n) = \arg\min_{\theta \in D_\mathcal{M}} \bar{E}\varepsilon^T(t,\theta)\varepsilon(t,\theta) \tag{A.6}$$

(If the minimum is not unique, let $\theta^*(n)$ denote *any* minimizing element.) Now define the estimate $\hat{\theta}_N(n,\delta)$ by

$$\hat{\theta}_N(n,\delta) = \arg\min_{\theta \in D_\mathcal{M}} V_N(\theta, n, \delta, Z^N) \tag{A.7}$$

$$V_N(\theta, n, \delta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon^T(t,\theta)\varepsilon(t,\theta) + \delta\|\theta - \theta^*(n)\|^2 \tag{A.8}$$

Here δ is a regularization parameter helping us to select a unique minimizing element in (A.7) in case $\delta = 0$ leads to nonunique minima. Let

$$T_n^*(e^{i\omega}) = T(e^{i\omega}, \theta^*(n)) \tag{A.9}$$

$$\hat{T}_N(e^{i\omega}, n, \delta) = T(e^{i\omega}, \hat{\theta}_N(n,\delta)) \tag{A.10}$$

$$T_0(e^{i\omega}) = \mathrm{vec}\big[G_0(e^{i\omega}) \;\; H_0(e^{i\omega})\big] \tag{A.11}$$

It will be assumed that

$$\lim_{n\to\infty} n^2 E\|\varepsilon(t,\theta^*(n)) - e(t)\|^2 = 0 \tag{A.12}$$

which implies that $T_n^*(e^{i\omega})$ tends to $T_0(e^{i\omega})$ as n tends to infinity. Define $Z(q,\theta)$ through (cf. (42))

$$\frac{\partial}{\partial\theta_k}T(q,\theta) = q^{-k+1}\frac{\partial}{\partial\theta_1}T(q,\theta) \triangleq q^{-k+1}Z(q,\theta) \tag{A.13}$$

Let the matrix function $Z(e^{i\omega},\theta)$ be denoted by $Z_0(e^{i\omega})$ when evaluated for $T_0(e^{i\omega})$. It will also be assumed that

$$Z_0(e^{i\omega})Z_0^T(e^{i\omega}) \tag{A.14}$$

is invertible and that

$$\lim_{N\to\infty}\frac{1}{\sqrt{N}}\sum_{t=1}^{N} E\Big[\frac{d}{d\theta}\|\varepsilon(t,\theta^*(n))\|^2\Big] = 0 \quad (n \text{ fixed}) \tag{A.15}$$

Finally it will be assumed that all signal spectra involved exist and are well defined and that

$$\bar{\sigma}(\Phi_u(\omega)) \le C, \quad \underline{\sigma}(\Phi_u(\omega)) \ge \lambda > 0 \quad \forall\omega \tag{A.16}$$

The main variance result is now

Theorem A.2 *Consider the estimate* $\hat{T}_N(e^{i\omega},n,\delta)$ *under the assumptions* (42) *and* (A.5)-(A.16). *Suppose also that Assumptions 1 and 2 hold. Then*

$$\sqrt{N}\,\mathrm{vec}\big[\hat{T}_N(e^{i\omega},n,\delta) - T_n^*(e^{i\omega})\big] \in AsN(0,P(\omega,n,\delta)) \quad \text{as } N\to\infty \text{ for fixed } n,\ \delta \tag{A.17}$$

where

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}P(\omega,n,\delta) = (\Phi_{\chi_0}(\omega))^{-T}\otimes\Phi_v(\omega) \tag{A.18}$$

(Here $\otimes$ denotes the Kronecker product.)

This was proven for the SISO case in [23]. The extention to the MIMO case was established in [38].

The expression (43) can be seen as an intuitive, but not formally correct, interpretation of (A.17), (A.18).

In open loop we have [23, 38]

Corollary A.3 *Consider the same situation as in Theorem A.2 but assume that $H(q,\theta)$ is independent of θ. Assume that $\Phi_{ue} = 0$ and that* (A.12) *is relaxed to*

$$n^2 \| G_n^*(e^{i\omega}) - G_0(e^{i\omega}) \| \to 0 \quad \text{as } n \to \infty \tag{A.19}$$

Then

$$\sqrt{N}\,\mathrm{vec}\big[\hat{G}_N(e^{i\omega}, n, \delta) - G_n^*(e^{i\omega})\big] \in AsN(0, P(\omega, n, \delta)) \qquad \text{as } N \to \infty \text{ for fixed } n,\ \delta \tag{A.20}$$

where

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{n} P(\omega, n, \delta) = (\Phi_u(\omega))^{-T} \otimes \Phi_v(\omega) \tag{A.21}$$

A.3 Complement to Section 3.4

We will now turn to the question of asymptotic distribution of the parameter vector estimates. For notational convenience the discussion will be limited to the SISO case. Let λ_0 denote the variance of the noise e in (2) in the SISO case. The following result is a variant of Theorem 9.1 in [24]:

Theorem A.4 *Consider the estimate $\hat{\theta}_N$ determined by* (30)-(32) *with $L(q,\theta) =$ 1. Assume that the model structure is linear and uniformly stable and that Assumptions 1 and 2 hold. Suppose that for a unique value θ^* interior to $D_{\mathcal{M}}$ we have*

$$\hat{\theta}_N \to \theta^*, \quad \text{w. p. 1 as } N \to \infty \tag{A.22}$$

$$R = \bar{V}''(\theta^*) > 0 \tag{A.23}$$

$$\sqrt{N} E\Big[\frac{1}{N}\sum_{t=1}^{N}[\psi(t,\theta^*)\varepsilon(t,\theta^*) - \bar{E}\psi(t,\theta^*)\varepsilon(t,\theta^*)]\Big] \to 0, \quad \text{as } N \to \infty \tag{A.24}$$

Then

$$\sqrt{N}(\hat{\theta}_N - \theta^*) \in AsN(0, P_\theta) \tag{A.25a}$$

$$P_\theta = R^{-1} Q R^{-1} \tag{A.25b}$$

$$Q = \lim_{N \to \infty} N \cdot E\big\{[V_N'(\theta^*, Z^N)][V_N'(\theta^*, Z^N)]^T\big\} \tag{A.25c}$$

Remark: Frequency domain expressions for R and Q can easily be derived using Parseval's relationship (cf. [24], Section 9.4). □

B Additional Proofs

B.1 Proof of Corollary 5

The result is direct consequence of Theorem A.1. To derive the expression (101) we simply have to insert the factorization (10) of Φ_{χ_0} into (A.1). □

We now turn to the proof of (103). Consider the first term in (101). If we define $\tilde{G}_\theta = G_0 - G_\theta$ we have (neglecting the trace operator and the scaling Λ^{-1} which are irrelevant here)

$$\begin{aligned}
&\int_{-\pi}^{\pi} H_\theta^{-1}(\tilde{G}_\theta + B_\theta)\Phi_u(\tilde{G}_\theta + B_\theta)^* H_\theta^{-*} d\omega \\
&= \int_{-\pi}^{\pi} (H_\theta^{-1}\tilde{G}_\theta M_u + H_\theta^{-1}B_\theta M_u)(H_\theta^{-1}\tilde{G}_\theta M_u + H_\theta^{-1}B_\theta M_u)^* d\omega \\
&= \int_{-\pi}^{\pi} \Big(H_\theta^{-1}\tilde{G}_\theta M_u + \big[H_\theta^{-1}B_\theta M_u\big]_+\Big)\Big(H_\theta^{-1}\tilde{G}_\theta M_u + \big[H_\theta^{-1}B_\theta M_u\big]_+\Big)^* \\
&\qquad + \big[H_\theta^{-1}B_\theta M_u\big]_- \big[H_\theta^{-1}B_\theta M_u\big]_-^* \, d\omega \\
&= \int_{-\pi}^{\pi} H_\theta^{-1}\big\{(\tilde{G}_\theta + B_\theta^+)\Phi_u(\tilde{G}_\theta + B_\theta^+)^* + B_\theta^- \Phi_u (B_\theta^-)^*\big\} H_\theta^{-*} d\omega
\end{aligned} \tag{B.26}$$

where we in the second step used that $H_\theta^{-1}\tilde{G}_\theta M_u$ is causal. (103) follows. □

B.2 Proof of Corollary 6

The main convergence result was given in Theorem A.1. To derive the expression (109) we need to compute Φ_ε. The prediction error is (using (25))

$$\varepsilon = H_*^{-1}(y - G_\theta^c r) = H_*^{-1}\big[(G_0^c - G_\theta^c)r + S_0 v + G_0^c d\big] \tag{B.27}$$

Thus

$$\begin{aligned}
\Phi_\varepsilon = H_*^{-1}\big[(G_0^c - G_\theta^c)\Phi_r(G_0^c - G_\theta^c)^* + (G_0^c - G_\theta^c)\Phi_{rd}(G_0^c)^* + G_0^c\Phi_{dr}(G_0^c - G_\theta^c)^*\big] H_*^{-*} \\
+ \theta\text{-indep. terms}
\end{aligned} \tag{B.28}$$

With $\bar{B}$ given by (110) this can be rewritten as

$$\Phi_\varepsilon \;=\; H_*^{-1}\big[(G_0^c + \bar{B} - G_\theta^c)\Phi_r(G_0^c + \bar{B} - G_\theta^c)^*\big] H_*^{-*} + \theta\text{-indep. terms} \tag{B.29}$$

(109) follows if we consider only the θ-dependent terms. □

The alternative expression (111) will now be derived. Define $\tilde{G}_\theta^c = G_0^c - G_\theta^c$. Then

$$\begin{aligned}
&\int_{-\pi}^{\pi} H_*^{-1}(\tilde{G}_\theta^c + \bar{B})\Phi_r(\tilde{G}_\theta^c + \bar{B})^* H_*^{-*} d\omega \\
&= \int_{-\pi}^{\pi} (H_*^{-1}\tilde{G}_\theta^c M_r + H_*^{-1}\bar{B}M_r)(H_*^{-1}\tilde{G}_\theta^c M_r + H_*^{-1}\bar{B}M_r)^* d\omega \\
&= \int_{-\pi}^{\pi} \left(H_*^{-1}\tilde{G}_\theta^c M_r + \left[H_*^{-1}\bar{B}M_r\right]_+\right)\left(H_*^{-1}\tilde{G}_\theta M_r + \left[H_*^{-1}\bar{B}M_r\right]_+\right)^* \\
&\quad + \left[H_*^{-1}\bar{B}M_r\right]_- \left[H_*^{-1}\bar{B}M_r\right]_-^* d\omega \\
&= \int_{-\pi}^{\pi} H_*^{-1}(\tilde{G}_\theta^c + \bar{B}^+)\Phi_r(\tilde{G}_\theta^c + \bar{B}^+)^* H_*^{-*} d\omega + C
\end{aligned} \tag{B.30}$$

where $\bar{B}^+$ is given by (112) and C is a constant matrix, independent of θ. In the second step used that $H_*^{-1}\tilde{G}_\theta^c M_r$ is causal. It follows that (111) is equivalent to (109), as claimed. □

B.3 Proof of Corollary 7

The proof of (116) is analogous to the corresponding proof for the indirect approach and hence omitted. For the proof of (119) we first note that the main convergence result follows from Theorem A.1. We need to compute Φ_ε. The prediction error is

$$\varepsilon = H_*^{-1}(y - G_\theta \hat{u}) = H_*^{-1}\left[(G_0 - G_\theta)\hat{u} + w(t)\right] \tag{B.31}$$

where $w = G_0\tilde{u} + v$. Hence

$$\begin{aligned}
\Phi_\varepsilon = H_*^{-1}\left[(G_0 - G_\theta)\Phi_{\hat{u}}(G_0 - G_\theta)^* + (G_0 - G_\theta)\Phi_{w\hat{u}} + \Phi_{\hat{u}w}(G_0 - G_\theta)^*\right]H_*^{-*} \\
+ (H_*)^{-1}\Phi_w H_*^{-*}
\end{aligned} \tag{B.32}$$

where we can note that $\Phi_{w\hat{u}} = G_0\Phi_{\tilde{u}\hat{u}}$. If we use $\tilde{B}$ defined in (120) we can thus alternatively write Φ_ε as

$$\Phi_\varepsilon = H_*^{-1}\left[(G_0 + \tilde{B} - G_\theta)\Phi_{\hat{u}}(G_0 + \tilde{B} - G_\theta)^*\right]H_*^{-*} + \theta\text{-indep. terms} \tag{B.33}$$

from which (119) follows. □

B.4 Proof of Corollary 9

For the proof of the result for the indirect approach we will assume that the model is given by (49). (The parameterization as well as the noise model can be arbitrary.

Thus the result applies to all variants of the indirect method studied in this paper.) According to Corollary A.3 the random variable

$$\sqrt{N}\,\mathrm{vec}\big[\hat{G}_N^c(e^{i\omega},n,\delta)-G_n^{c,*}(e^{i\omega})\big] \tag{B.34}$$

will have an asymptotic normal distribution with covariance matrix $\bar{P}(\omega,n,\delta)$ satisfying

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\bar{P}(\omega,n,\delta)=(\Phi_r(\omega))^{-T}\otimes\Phi_{v_c}(\omega) \tag{B.35}$$

where $\Phi_{v_c}=S_0\Phi_v S_0^*$. From the Taylor expansion

$$\begin{aligned}\mathrm{vec}\big[\hat{G}_N(e^{i\omega},n,\delta)-G_n^*(e^{i\omega})\big]&=\\ =\left(\frac{d\,\mathrm{vec}[G]}{d\,\mathrm{vec}[G^c]}(e^{i\omega},\theta^*(n))\right)^T&\Big\{\mathrm{vec}\big[\hat{G}_N^c(e^{i\omega},n,\delta)-G_n^{c,*}(e^{i\omega})\big]\Big\}\\ &+O(\|\mathrm{vec}\big[\hat{G}_N^c(e^{i\omega},n,\delta)-G_n^{c,*}(e^{i\omega})\big]\|)\end{aligned} \tag{B.36}$$

we see that the covariance matrix of

$$\sqrt{N}\,\mathrm{vec}\big[\hat{G}_N(e^{i\omega},n,\delta)-G_n^*(e^{i\omega})\big] \tag{B.37}$$

will be

$$P(\omega,n,\delta)=\left(\frac{d\,\mathrm{vec}[G]}{d\,\mathrm{vec}[G^c]}(e^{i\omega},\theta^*(n))\right)^{-T}\bar{P}(\omega,n,\delta)\left(\frac{d\,\mathrm{vec}[G]}{d\,\mathrm{vec}[G^c]}(e^{-i\omega},\theta^*(n))\right)^{-1} \tag{B.38}$$

Note that

$$G^c=SG=(I+GK)^{-1}G=G-GKG+GKGKG-\ldots \tag{B.39}$$

Repeated use of the formula [14]

$$\frac{d\,\mathrm{vec}[AXB]}{d\,\mathrm{vec}[X]}=(B\otimes A^T) \tag{B.40}$$

and the chain rule now gives

$$\begin{aligned}\frac{d\,\mathrm{vec}[G^c]}{d\,\mathrm{vec}[G]}&=I-(KG\otimes I)-(I\otimes K^TG^T)+(KG\otimes K^TG^T)+\\ &+(KGKG\otimes I)+(I\otimes K^TG^TK^TG^T)-\ldots\end{aligned} \tag{B.41}$$

As can be readily verified, with $S^i=(I+KG)^{-1}$ and $S=(I+GK)^{-1}$, we thus have that

$$\frac{d\,\mathrm{vec}[G^c]}{d\,\mathrm{vec}[G]}=\big(S^i\otimes S^T\big)=\big((S^i)^T\otimes S\big)^T \tag{B.42}$$

Consequently

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}P(\omega,n,\delta) &= \left((S_0^i)^T\otimes S_0\right)^{-1}\left((\Phi_r)^{-T}\otimes\Phi_{v_c}\right)\left((S_0^i)^T\otimes S_0\right)^{-*}\\ &= \left((S_0^i)^{-T}\otimes S_0^{-1}\right)\left((\Phi_r)^{-T}\otimes\Phi_{v_c}\right)\left(((S_0^i)^T)^{-*}\otimes S_0^{-*}\right)\\ &= \left((S_0^i)^{-T}(\Phi_r)^{-T}((S_0^i)^*)^{-T}\right)\otimes\left(S_0^{-1}\Phi_{v_c}S_0^{-*}\right)\\ &= (\Phi_u^r)^{-T}\otimes\Phi_v \qquad\qquad (B.43)\end{aligned}$$

which concludes the proof for the indirect approach.

For the joint input-output approach we will in this proof assume that the model is

$$\begin{bmatrix}\hat{y}(t,\theta)\\ \hat{u}(t,\theta)\end{bmatrix} = \begin{bmatrix}G^c(q,\theta)\\ S^i(q,\theta)\end{bmatrix} r(t) \qquad (B.44)$$

and that the open-loop estimate is computed as $\hat{G}_N = \hat{G}_N^c(\hat{S}_N^i)^{-1}$. (The result holds regardless of the parameterization and regardless of the noise model used, hence it holds for all joint input-output method studied in this paper, including the two-stage/projection method.) Write Q as short for $\left[(G^c)^T\ (S^i)^T\right]^T$. According to Corollary A.3 the random variable

$$\sqrt{N}\,\mathrm{vec}\left[\hat{Q}_N(e^{i\omega},n,\delta)-Q_n^*(e^{i\omega})\right] \qquad (B.45)$$

will have an asymptotic normal distribution with covariance matrix $\bar{P}(\omega,n,\delta)$ satisfying

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\bar{P}(\omega,n,\delta) = (\Phi_r(\omega))^{-T}\otimes\begin{bmatrix}I\\ -K(e^{i\omega})\end{bmatrix}\Phi_{v_c}(\omega)\begin{bmatrix}I\\ -K(e^{i\omega})\end{bmatrix}^* \qquad (B.46)$$

where $\Phi_{v_c} = S_0\Phi_v S_0^*$. From the Taylor expansion

$$\begin{aligned}&\mathrm{vec}\left[\hat{G}_N(e^{i\omega},n,\delta)-G_n^*(e^{i\omega})\right] =\\ &\quad= \left(\frac{d\,\mathrm{vec}[G]}{d\,\mathrm{vec}[Q]}(e^{i\omega},\theta^*(n))\right)^T\left\{\mathrm{vec}\left[\hat{Q}_N(e^{i\omega},n,\delta)-Q_n^*(e^{i\omega})\right]\right\}\\ &\qquad + O(\|\mathrm{vec}\left[\hat{Q}_N(e^{i\omega},n,\delta)-Q_n^*(e^{i\omega})\right]\|) \qquad (B.47)\end{aligned}$$

we see that the covariance matrix of

$$\sqrt{N}\,\mathrm{vec}\left[\hat{G}_N(e^{i\omega},n,\delta)-G_n^*(e^{i\omega})\right] \qquad (B.48)$$

will be

$$P(\omega,n,\delta) = \left(\frac{d\,\mathrm{vec}[G]}{d\,\mathrm{vec}[Q]}(e^{i\omega},\theta^*(n))\right)^{-T}\bar{P}(\omega,n,\delta)\left(\frac{d\,\mathrm{vec}[G]}{d\,\mathrm{vec}[Q]}(e^{-i\omega},\theta^*(n))\right)^{-1} \qquad (B.49)$$

By use of the formula (B.40) and the chain rule it can be shown that

$$\begin{aligned}\frac{d\,\text{vec}[G]}{d\,\text{vec}[Q]} &= \Big((S^i)^{-1} \otimes \big[I \ 0\big]^T\Big) + \Big((S^i)^{-1} \otimes \big[0 \ -G^c(S^i)^{-1}\big]^T\Big) \\ &= (S^i)^{-1} \otimes \big[I \ -G^c(S^i)^{-1}\big]^T \end{aligned} \tag{B.50}$$

Note that $G^c(S^i)^{-1} \to G_0$ as $n \to \infty$. Thus

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} P(\omega, n, \delta) &= \big((S_0^i)^{-T} \otimes \big[I \ -G_0\big]\big) \cdot \\ &\cdot \left((\Phi_r)^{-T} \otimes \begin{bmatrix} I \\ -K \end{bmatrix} \Phi_{v_c} \begin{bmatrix} I \\ -K \end{bmatrix}^*\right) \big((S_0^i)^{-T} \otimes \big[I \ -G_0\big]\big)^* \\ &= \big((S_0^i)^{-T} (\Phi_r)^{-T} ((S_0^i)^*)^{-T}\big) \otimes \big(S_0^{-1} \Phi_{v_c} S_0^{-*}\big) \\ &= (\Phi_u^r)^{-T} \otimes \Phi_v \end{aligned} \tag{B.51}$$

which concludes the proof for the joint input-output approach and the also the proof of Corollary 9. □

B.5 Proof of Corollary 10

Under the conditions in the corollary we have that $\hat{\theta}_N \to \theta_0$ w. p. 1 as $N \to \infty$. See, e.g., Theorem 8.3 in [24]. It follows from Theorem A.4 that

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, P_\theta) \tag{B.52}$$

where

$$P_\theta = \lambda_0 R_\theta \tag{B.53}$$

$$R_\theta = \big[\bar{E}\psi(t, \theta_0)\psi^T(t, \theta_0)\big] \tag{B.54}$$

Using Parseval's relationship R_θ can be written as (see, e.g., [24])

$$R_\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|H_0|^2} \bar{T}'_\theta \Phi_{\chi_0} \bar{T}'^*_\theta \, d\omega \tag{B.55}$$

where $\bar{T} = [G \ H]$ and where $\bar{T}'_\theta = \frac{d}{d\theta}\bar{T}|_{\theta=\theta_0}$. From (13) it follows that

$$\bar{T}'_\theta \Phi_{\chi_0} \bar{T}'^*_\theta = \bar{T}'_\theta \Phi^r_{\chi_0} \bar{T}'^*_\theta + \bar{T}'_\theta \Phi^e_{\chi_0} \bar{T}'^*_\theta \tag{B.56}$$

We may thus write

$$R_\theta = R_\theta^r + R_\theta^e \tag{B.57}$$

where, due to the chosen parameterization,

$$R_\theta^r = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{|H_0|^2}\bar{T}_\theta' \Phi_{\chi_0}^e \bar{T}_\theta'^* d\omega = \begin{bmatrix} R_\rho^r & 0 \\ 0 & 0 \end{bmatrix} \tag{B.58}$$

and

$$R_\theta^e = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{|H_0|^2}\bar{T}_\theta' \Phi_{\chi_0}^r \bar{T}_\theta'^* d\omega = \begin{bmatrix} R_\rho^e & R_{\rho\beta}^e \\ R_{\beta\rho}^e & R_\beta^e \end{bmatrix} \tag{B.59}$$

Straight-forward calculations will also show that

$$R_\rho^r = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\Phi_u^r}{|H_0|^2} G_\rho' G_\rho'^* d\omega \tag{B.60}$$

and

$$R_\theta^e = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{|H_0|^2}\bar{T}_\theta' \Phi_{\chi_0}^e \bar{T}_\theta'^* d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\lambda_0}{|H_0|^2} \begin{bmatrix} KS_0H_0G_\rho' \\ -H_\beta' \end{bmatrix} \begin{bmatrix} KS_0H_0G_\rho' \\ -H_\beta' \end{bmatrix}^* d\omega \tag{B.61}$$

P_ρ, the covariance matrix for $\hat{\rho}_N$ can be found as the top left block of P_θ. Combining (B.53), (B.54), (B.58), and (B.59) we thus have that

$$P_\rho = \lambda_0 (R_\rho^r + \Delta)^{-1} \tag{B.62}$$

$$\Delta = R_\rho^e - R_{\rho\beta}^e (R_\beta^e)^{-1} R_{\beta\rho}^e \tag{B.63}$$

where R_ρ^r is given by (B.60) and expressions for R_ρ^e, $R_{\rho\beta}^e$, and R_β^e can be extracted from (B.61).

To prove the lower bound in (128) note that Δ is the Schur complement of R_β^e in the matrix R_θ^e given by (B.59). Since $R_\theta^e \geq 0$ we thus have $\Delta \geq 0$. Next, define

$$\psi_e(t,\theta_0) = \bar{T}_\theta'(q,\theta_0) \begin{bmatrix} K(q)S_0(q)H_0(q) \\ -1 \end{bmatrix} e(t) \tag{B.64}$$

Then we can alternatively write R_θ^e as

$$R_\theta^e = \bar{E}\psi_e(t,\theta_0)\psi_e^T(t,\theta_0) \tag{B.65}$$

Now introduce the notation

$$w(t) = H_0^{-1}(q)e(t) \tag{B.66}$$

$$\bar{G}_\rho'(q) = K(q)S_0(q)H_0(q)G_\rho'(q) \tag{B.67}$$

so that

$$\psi_e(t,\theta_0) = \begin{bmatrix} \bar{G}_\rho'(q)w(t) \\ -H_\beta'(q)w(t) \end{bmatrix} \tag{B.68}$$

Here the number of rows in $\bar{G}'_\rho$ and H'_β are consistent with the partitioning (B.58), (B.59). From well known least squares projections (see, e.g., [3]), we recognize Δ as *the error covariance matrix when estimating* $\bar{G}'_\rho w$ *from* $H'_\beta w$. Suppose now that the order of the noise model tends to infinity. Without loss of generality we can assume that

$$H(q,\beta) = 1 + \lim_{M\to\infty} \sum_{k=1}^{M} h_k q^{-k} \tag{B.69}$$

Then knowing $H'_\beta w$ is equivalent to knowing all past w. Then $\bar{G}'_\rho w$ can be determined exactly from $H'_\beta w$, and $\Delta = 0$. At the other extreme, a fixed (and correct) noise model will make $\Delta = R^e_\rho = \bar{E}\bar{G}'_\rho w(\bar{G}'_\rho w)^T$, which is the largest value Δ may have. This proves the upper bound in (128) and the remaining statements in the corollary. □

B.6 Proof of Corollary 12

Let

$$\begin{aligned} A(q) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \\ B(q) &= b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \\ C(q) &= 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c} \end{aligned} \tag{B.70}$$

and similarly for the closed-loop polynomials. Define θ as

$$\theta = \begin{bmatrix} a_1, & \ldots, & a_{n_a}, & b_1, & \ldots, & b_{n_b}, & c_1, & \ldots, & c_{n_c} \end{bmatrix}^T \tag{B.71}$$

and η equivalently. Furthermore let the regulator polynomials be

$$\begin{aligned} X(q) &= x_0 + x_1 q^{-1} + \cdots + x_{n_x} q^{-n_x} \\ Y(q) &= 1 + y_1 q^{-1} + \cdots + y_{n_y} q^{-n_y} \end{aligned} \tag{B.72}$$

With these definitions the system of equations (140) can be written as

$$\Gamma\theta = \bar{\eta} \tag{B.73}$$

where

$$\Gamma = \begin{bmatrix} \Gamma_Y & \Gamma_X & 0 \\ 0 & \Gamma_Y & 0 \\ 0 & 0 & \Gamma_Y \end{bmatrix}, \quad \Gamma_Y = \begin{bmatrix} 1 & & \\ y_1 & \ddots & \\ \vdots & \ddots & 1 \\ & & y_1 \\ & & \vdots \end{bmatrix}, \quad \Gamma_X = \begin{bmatrix} x_0 & & \\ x_1 & \ddots & \\ \vdots & \ddots & x_0 \\ & & x_1 \\ & & \vdots \end{bmatrix} \tag{B.74}$$

while

$$\bar{\eta} = \big[a_{c,1} - y_1,\ a_{c,2} - y_2,\ \dots,\ b_{c,1},\ b_{c,2},\ \dots,\ c_{c,1} - y_1,\ c_{c,2} - y_2,\ \dots\big]^T \quad \text{(B.75)}$$

Now, if the closed-loop parameters η are estimated from a set of N data we have from Theorem A.4 that

$$\sqrt{N}(\hat{\eta}_N - \eta_0) \in AsN(0, P_\eta) \quad \text{(B.76)}$$

$$P_\eta = \lambda_0 \big[\bar{E}\psi_c(t,\eta_0)\psi_c^T(t,\eta_0)\big]^{-1} \quad \text{(B.77)}$$

$$\psi_c(t,\eta) = -\frac{d}{d\eta}\frac{A^c(q)}{C^c(q)}\left(y(t) - \frac{B^c(q)}{A^c(q)}r(t)\right) \quad \text{(B.78)}$$

Consider equation (B.73). The Markov estimate of θ is

$$\hat{\theta}_N = \big[\Gamma^T(\operatorname{Cov}\hat{\bar{\eta}}_N)^{-1}\Gamma\big]^{-1}\Gamma^T(\operatorname{Cov}\hat{\bar{\eta}}_N)^{-1}\hat{\bar{\eta}}_N \quad \text{(B.79)}$$

from which it follows that

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, P_\theta) \quad \text{(B.80)}$$

$$P_\theta = N\big[\Gamma^T(\operatorname{Cov}\hat{\bar{\eta}}_N)^{-1}\Gamma\big]^{-1} \quad \text{(B.81)}$$

We also have that

$$\operatorname{Cov}\hat{\bar{\eta}}_N = \frac{1}{N}P_\eta = \frac{\lambda_0}{N}\big[\bar{E}\psi_c(t,\eta_0)\psi_c^T(t,\eta_0)\big]^{-1} \quad \text{(B.82)}$$

Hence

$$P_\theta = \lambda_0\big[\bar{E}\Gamma^T\psi_c(t,\eta_0)\psi_c^T(t,\eta_0)\Gamma\big]^{-1} \quad \text{(B.83)}$$

The gradient vector $\psi_c(t,\eta_0)$ is

$$\begin{aligned}\psi_c(t,\eta_0) = \frac{1}{C_0^c(q)}\big[&-y(t-1),\ \dots,\ -y(t-n_{a_c}),\\ &r(t-1),\ \dots,\ r(t-n_{b_c}),\ e(t-1),\ \dots,\ e(t-n_{c_c})\big]^T\end{aligned} \quad \text{(B.84)}$$

and since

$$\frac{Y(q)}{C_0^c(q)} = \frac{1}{C_0(q)} \quad \text{and} \quad \frac{X(q)y(t)}{C_0^c(q)} - \frac{Y(q)r(t)}{C_0^c(q)} = -\frac{1}{C_0(q)}u(t) \quad \text{(B.85)}$$

we have

$$\begin{aligned}\Gamma^T\psi_c(t,\eta_0) = \frac{1}{C_0(q)}\big[&-y(t-1),\ \dots,\ -y(t-n_a),\\ &u(t-1),\ \dots,\ u(t-n_b),\ e(t-1),\ \dots,\ e(t-n_c)\big]^T\end{aligned} \quad \text{(B.86)}$$

It follows that

$$\Gamma^T \psi_c(t, \eta_0) = -\frac{d}{d\theta} \frac{A(q)}{C(q)} \Big(y(t) - \frac{B(q)}{A(q)} u(t) \Big) \tag{B.87}$$

where the right-hand side should be evaluated at $\theta = \theta_0$. But the expression on the right is equal to the negative gradient of the open-loop prediction error, that is $\psi(t, \theta)$, evaluated at $\theta = \theta_0$. Thus $\Gamma^T \psi_c(t, \eta_0) = \psi(t, \theta_0)$ and

$$P_\theta = \lambda_0 \big[\bar{E} \psi(t, \theta_0) \psi^T(t, \theta_0) \big]^{-1} \tag{B.88}$$

which together with (B.80) proves the corollary. □

B.7 Proof of Corollary 13

Let

$$\begin{aligned} B(q) &= b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \\ F(q) &= 1 + f_1 q^{-1} + \cdots + f_{n_f} q^{-n_f} \\ C(q) &= 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c} \\ D(q) &= 1 + d_1 q^{-1} + \cdots + d_{n_d} q^{-n_d} \end{aligned} \tag{B.89}$$

and similarly for the closed-loop polynomials and let X, Y, Γ_X, and Γ_Y be as in the proof of Corollary 12. Then (144) can be written

$$\tilde{\Gamma} \rho = \bar{\rho}_\eta \tag{B.90}$$

where

$$\tilde{\Gamma} = \begin{bmatrix} \Gamma_Y & 0 \\ \Gamma_X & \Gamma_Y \end{bmatrix}, \quad \bar{\rho}_\eta = \big[b_{c,1},\ b_{c,2},\ \ldots,\ f_{c,1} - y_1,\ f_{c,2} - y_2,\ \ldots \big]^T \tag{B.91}$$

Using the same arguments as in the proof of Corollary 12 we get

$$\sqrt{N}(\hat{\rho}_N - \rho_0) \in AsN(0, P_\rho) \tag{B.92}$$

$$P_\rho = N \big[\tilde{\Gamma}^T (\operatorname{Cov} \hat{\bar{\rho}}_{\eta,N})^{-1} \tilde{\Gamma} \big]^{-1} \tag{B.93}$$

We also have that

$$\operatorname{Cov} \hat{\bar{\rho}}_{\eta,N} = \operatorname{Cov} \hat{\rho}_{\eta,N} = \frac{\lambda_0}{N} \big[\bar{E} \psi_{c,\rho_\eta}(t, \eta_0) \psi^T_{c,\rho_\eta}(t, \eta_0) \big]^{-1} \tag{B.94}$$

where ψ_{c,ρ_η} is the negative gradient of the closed-loop predictions error,

$$\varepsilon_c(t, \eta) = \frac{1}{H^c(q, \beta_\eta)} \Big(y(t) - G^c(q, \rho_\eta) r(t) \Big) \tag{B.95}$$

$$= \frac{1}{H^c(q, \beta_\eta)} \Big(y(t) - \frac{B^c(q)}{F^c(q)} r(t) \Big) \tag{B.96}$$

taken w.r.t. ρ_η, giving

$$\psi_{c,\rho_\eta}(t,\eta_0) = \frac{1}{H_0^c(q)F_0^c(q)}\big[r(t-1),\ \ldots,\ r(t-n_{b_c}),$$
$$-\frac{B_0^c(q)}{F_0^c(q)}(r(t-1),\ \ldots,\ r(t-n_{f_c}))\big] \quad \text{(B.97)}$$

Returning to P_ρ we thus have that

$$P_\rho = \lambda_0\big[\bar{E}\tilde{\Gamma}^T\psi_c(t,\eta_0)\psi_c^T(t,\eta_0)\tilde{\Gamma}\big]^{-1} \quad \text{(B.98)}$$

However, since

$$H_0^c = S_0H_0,\ \frac{F_0^cY - B_0^cX}{(F_0^c)^2} = \frac{S_0^2}{F_0}, \quad \frac{B_0^cY}{(F_0^c)^2} = \frac{S_0^2}{F_0}\cdot\frac{B_0}{F_0} \quad \text{(B.99)}$$

we get (after some calculations)

$$\tilde{\Gamma}^T\psi_{c,\rho_\eta}(t,\eta_0) = \frac{S_0(q)}{H_0(q)F_0(q)}\big[r(t-1),\ \ldots,\ r(t-n_b),$$
$$-\frac{B_0(q)}{F_0(q)}(r(t-1),\ \ldots,\ r(t-n_f))\big]^T \quad \text{(B.100)}$$

so

$$\tilde{\Gamma}^T\psi_{c,\rho_\eta}(t,\eta_0) = -\frac{S_0(q)}{H_0(q)}\cdot\frac{d}{d\rho_\theta}\Big(y(t) - \frac{B(q)}{F(q)}r(t)\Big) \quad \text{(B.101)}$$
$$= \frac{S_0(q)}{H_0(q)}\cdot\frac{d}{d\rho_\theta}G(q,\rho_\theta)r(t) \quad \text{(B.102)}$$

where the derivatives on the right should be evaluated at $\rho_\theta = \rho_{\theta,0}$. Using Parseval's formula we thus have that

$$P_\rho = \lambda_0\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\Phi_u^r}{|H_0|^2}G'_\rho G'^*_\rho\,d\omega\right]^{-1} \quad \text{(B.103)}$$

and the corollary follows. □

Paper B

A Projection Method for Closed-loop Identification

Authors: Urban Forssell and Lennart Ljung

Edited version of paper to appear in *IEEE Transactions on Automatic Control*.

Preliminary version published as Technical Report LiTH-ISY-R-2080, Department of Electrical Engineering, Linköping University, Linköping, Sweden.

A Projection Method for Closed-loop Identification

Urban Forssell and Lennart Ljung

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: ufo@isy.liu.se, ljung@isy.liu.se.

Abstract

A new method for closed-loop identification that allows fitting the model to the data with arbitrary frequency weighting is described and analyzed. Just as the direct method this new method is applicable to systems with arbitrary feedback mechanisms. This is in contrast to other methods, such as the indirect method and the two-stage method, that assume linear feedback. The finite sample behavior of the proposed method is illustrated in a simulation study.

Keywords: Closed-loop identification, Prediction error methods.

1 Introduction

In "Identification for Control" the goal is to construct models that are suitable for control design. It is widely appreciated that small model uncertainty around the cross-over frequency is essential for successful control design. Consequently there has been a substantial interest in identification methods that provide a tunable optimality criterion so that the model can be fit to the data with a suitable frequency weighting. With open-loop experiments this is no problem: It is well known that arbitrary frequency weighting can be obtained by applying a prediction error method to an output error model structure with a suitable fixed noise model/prefilter (recall (cf. [4]) that the effect of any prefilter may be included in the noise model). However, open-loop experiments are not always possible since the system might be unstable or has to be controlled for safety or production reasons. In such cases closed-loop experiments have to be used. The problem is now that the simple approach of using an output error model with a fixed noise model/prefilter will give biased results when applied directly to closed-loop data, unless the fixed noise model correctly models the true noise color (see, e.g., Theorem 8.3 in [4]). A way around this would be to use a flexible, parameterized noise model. This would eliminate the bias but the frequency weighting would then not be fixed.

In this contribution we describe and analyze a closed-loop identification method that is consistent and, in the case of under-modeling, allows fitting the model to the data with arbitrary frequency weighting. This method will be referred to as *the projection method*. The projection method is in form similar to the two-stage method [5] but the ideas underlying the projection method are also related to the cross-spectral method described in [1] and to Wiener filtering, e.g., [2].

2 Preliminaries

In this section we will specify the assumptions we make on the system and the feedback and present the identification method considered. Some notation will also be introduced.

Assumption 1 *The true system is scalar, linear, time-invariant, and causal and given by*

$$y(t) = G_0(q)u(t) + v(t), \quad v(t) = H_0(q)e(t) \tag{1}$$

where $\{e(t)\}$ is a zero-mean white noise process with variance λ_0, and bounded moments of order $4+\delta$, some $\delta > 0$, and $H_0(q)$ is an inversely stable, monic filter. (The symbol q denotes the shift operator, i.e., $q^{-1}u(t) = u(t-1)$.)

Assumption 2 *The input is generated as*

$$u(t) = k(t, y^t, u^{t-1}, r(t)) \tag{2}$$

where $y^t = [y(1), \ldots, y(t)]$, etc., and where the reference signal $\{r(t)\}$ is a given quasi-stationary signal, independent of $\{v(t)\}$ and k is a given deterministic function such that the closed-loop system (1) and (2) is exponentially stable, which we define as follows: For each $t, s;\ t \geq s$ there exist random variables $\bar{y}_s(t), \bar{u}_s(t)$, independent of r^s and v^s but not independent of r^t and v^t, such that

$$\bar{E}|y(t) - \bar{y}_s(t)|^4 < C\lambda^{t-s} \tag{3}$$

$$\bar{E}|u(t) - \bar{u}_s(t)|^4 < C\lambda^{t-s} \tag{4}$$

for some $C < \infty,\ \lambda < 1$.

Here we have used the notation

$$\bar{E}f(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Ef(t) \tag{5}$$

The concept of quasi-stationarity is defined in, e.g., [4]. For a quasi-stationary signal $\{s(t)\}$ we define the covariance function and the spectrum as

$$R_s(\tau) = \bar{E}s(t)s(t-\tau) \quad \text{and} \quad \Phi_s(\omega) = \sum_{\tau=-\infty}^{\infty} R_s(\tau)e^{-i\omega\tau} \tag{6}$$

respectively. Similarly, for two jointly quasi-stationary signals $\{s(t)\}$ and $\{w(t)\}$ we define the cross-covariance function and cross-spectrum as

$$R_{sw}(\tau) = \bar{E}s(t)w(t-\tau) \quad \text{and} \quad \Phi_{sw}(\omega) = \sum_{\tau=-\infty}^{\infty} R_{sw}(\tau)e^{-i\omega\tau} \tag{7}$$

respectively. We will assume that we can factorize the spectrum $\Phi_s(\omega)$ as $\Phi_s(\omega) = M_s(e^{i\omega})M_s(e^{-i\omega})$ where $M_s(e^{i\omega})$ is causal, stable, and inversely stable. (To reduce the notational burden we will in the sequel suppress the arguments t, q, ω, and $e^{i\omega}$ whenever there is no risk of confusion.)

In this paper we will consider the standard prediction error method. This can be understood as follows. Given measured data $Z^N = \{y(1), u(1), \dots, y(N), u(N)\}$ and a parameterized predictor model $\hat{y}(t|\theta)$, where the parameters θ range over a set $D_{\mathcal{M}}$ which is assumed compact and connected, the prediction error estimate is found as [4]:

$$\hat{\theta}_N = \arg\min_{\theta\in D_{\mathcal{M}}} V_N(\theta), \quad V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon^2(t,\theta), \quad \varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) \tag{8}$$

Here $\varepsilon(t,\theta)$ is called the prediction error. Define the average criterion $\bar{V}(\theta) = \bar{E}\varepsilon^2(t,\theta)$. We then have the following convergence result [3]:

Theorem 3 *Let $\hat{\theta}_N$ be defined as in* (8)*, where $\varepsilon(t,\theta)$ is determined from a uniformly stable linear model structure, and suppose that Assumptions 1 and 2 hold. Then*

$$\hat{\theta}_N \to \theta_{opt} = \arg\min_{\theta\in D_{\mathcal{M}}} \bar{V}(\theta) \quad \textit{w.p. } 1 \textit{ as } N\to\infty \tag{9}$$

Remark: Using Parseval's relationship we may rewrite $\bar{V}(\theta)$ as

$$\bar{V}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_\varepsilon d\omega \tag{10}$$

where Φ_ε denotes the spectrum of the prediction error. It can thus be seen that $\bar{V}(\theta)$ (and consequently θ_{opt}) depends only on the second order properties (the spectrum) of the prediction error. □

3 The Projection Method

The projection method consists of the following two steps:

1. Estimate the parameters s_k in the noncausal FIR model

$$u(t) = S(q)r(t) + e(t) = \sum_{k=-M_1}^{M_2} s_k r(t-k) + e(t) \tag{11}$$

where M_1 and M_2 are chosen so large that any correlation between $u(t)$ and $r(s)$ for $t - s \notin [-M_1, \dots, M_2]$ can be ignored, and simulate the signal $\hat{u}(t) = \hat{S}(q)r(t)$.

2. Identify the open-loop system using a model of the kind

$$y(t) = G(q, \theta)\hat{u}(t) + H_*(q)e(t) \tag{12}$$

The two-stage method [5] can be explained using essentially the same steps. What distinguishes the projection method from the two-stage method is that a noncausal model is used in the first step instead of a causal one. (In the two-stage method it is common to use (causal) high-order FIR or ARX models in the first step.) This difference might seem subtle, but as we shall see this has important consequences for the applicability of the method and the quality of the estimates.

The first step in the projection method can be viewed as a least squares projection of u onto r – hence the name *projection* method – and this method (asymptotically) achieves that the least squares estimate $\hat{u}$ and the residual $\tilde{u} = u - \hat{u}$ are orthogonal, or uncorrelated ($\Phi_{\hat{u}\tilde{u}} = 0$). This will be further discussed in the next section.

4 Convergence Analysis

In this section we will characterize the limit models as N tends to infinity using the result in Theorem 3. The analysis will apply to both the two-stage method and the projection method. The main result is

Corollary 4 *Consider the situation in Theorem 3 and assume that the prediction error is obtained from the model* (12) *where $\hat{u} = \hat{S}r$. Suppose that the model* (12) *is uniformly stable and that r is persistently exciting. Then*

$$\hat{\theta}_N \to \theta_{opt} = \arg\min_{\theta \in D_{\mathcal{M}}} \int_{-\pi}^{\pi} \left| G_0 + \bar{B} - G_\theta \right|^2 \frac{\Phi_{\hat{u}}}{|H_*|^2} d\omega \ \text{w.p. } 1 \text{ as } N \to \infty, \ \bar{B} = \frac{G_0 \Phi_{\tilde{u}\hat{u}}}{\Phi_{\hat{u}}} \tag{13}$$

If $G(q,\theta)$ and $H_*(q)$ in (12) are causal the term $\bar{B}$ in (13) can be replaced by

$$\bar{B}^+ = \left[\frac{G_0\Phi_{\tilde{u}\hat{u}}}{H_* M_{\hat{u}}^*}\right]_+ \frac{H_*}{M_{\hat{u}}} \tag{14}$$

where superscript $*$ denotes complex conjugate and $[\cdot]_+$ denotes the causal part.

Proof *The main convergence result*

$$\hat{\theta}_N \to \theta_{opt} = \arg\min_{\theta \in D_{\mathcal{M}}} \bar{V}(\theta) \quad w.\ p.\ 1 \text{ as } N \to \infty \tag{15}$$

follows from Theorem 3. To derive the expression (13) we need to compute Φ_ε. The model is given by (12). The prediction error can thus be written

$$\varepsilon(t,\theta) = H_*^{-1}(q)[y(t) - G(q,\theta)\hat{u}(t)] = H_*^{-1}(q)[(G_0(q) - G(q,\theta))\hat{u}(t) + w(t)] \tag{16}$$

where

$$w(t) = G_0(q)(u(t) - \hat{u}(t)) + v(t) \tag{17}$$

We have

$$\Phi_\varepsilon = \frac{1}{|H_*|^2}[|G_0 - G_\theta|^2\Phi_{\hat{u}} + 2\Re\{(G_0 - G_\theta)\Phi_{\hat{u}w}\} + \Phi_w] \tag{18}$$

Since $\Phi_{\hat{u}} > 0$ (this follows since r is persistently exciting [4]) this can also be written

$$\Phi_\varepsilon = \frac{1}{|H_*|^2}\left[\left|G_0 - G_\theta + \frac{\Phi_{w\hat{u}}}{\Phi_{\hat{u}}}\right|^2\Phi_{\hat{u}} - \frac{|\Phi_{w\hat{u}}|^2}{\Phi_{\hat{u}}} + \Phi_w\right] \tag{19}$$

The result (13) now follows from Theorem 3 after noting that $\Phi_{w\hat{u}} = G_0\Phi_{\tilde{u}\hat{u}}$. If $G(q,\theta)$ and $H_*(q)$ are causal we have that

$$\int_{-\pi}^{\pi}\left|G_0 + \bar{B} - G_\theta\right|^2 \frac{\Phi_{\hat{u}}}{|H_*|^2} d\omega = \int_{-\pi}^{\pi}\left|G_0 + \bar{B}^+ - G_\theta\right|^2 \frac{\Phi_{\hat{u}}}{|H_*|^2} d\omega + C \tag{20}$$

where C is a constant. This ends the proof. □

The expression (13) shows that $\Phi_{\hat{u}\tilde{u}} = 0$ (i.e., that $\hat{u}$ and $\tilde{u}$ are uncorrelated) is a necessary condition for consistency. This applies both to the projection method and to the two-stage method. By construction the projection method (asymptotically) achieves this regardless of the nature of the feedback. If u is generated as

$$u(t) = r(t) - K(q)y(t) \tag{21}$$

where K is a linear regulator we have that

$$u(t) = S_0(q)r(t) - K(q)S_0(q)v(t) \tag{22}$$

where S_0 is the sensitivity function $S_0 = (1 + G_0K)^{-1}$. In this case the condition $\Phi_{\hat{u}\tilde{u}} = 0$ is equivalent to $\hat{S} = S_0$. This can be achieved (asymptotically) also with the two-stage method and hence this method can give consistent estimates of G_0 if the feedback is linear. The more general applicability of the projection method is due to the noncausal model used in the first step of the algorithm. This will be further discussed below.

From (13) it is also clear that, in case of undermodeling, the model can be fit to the data with arbitrary frequency weighting – if we have succeeded in achieving $\Phi_{\hat{u}\tilde{u}} = 0$.

Let us return to the question of how to obtain an estimate $\hat{u} = \hat{S}r$ of u such that $\Phi_{\hat{u}\tilde{u}} = 0$ or, equivalently, such that $\Phi_{r\tilde{u}} = 0$. Consider the linear least mean squares (l.l.m.s.) estimation problem (cf. (11))

$$\min_{s_k} \bar{E}(u(t) - \sum_{k=-M_1}^{M_2} s_k r(t-k))^2 \tag{23}$$

The optimal solution can be found by solving the normal equations

$$R_{ur}(\tau) = \sum_{k=-M_1}^{M_2} s_k R_r(k-\tau),\ \tau = -M_1, \ldots, M_2 \tag{24}$$

Note that this implies that, for the optimal solution, $\tilde{u}(t)$ will be uncorrelated with $r(t-\tau)$ for $\tau = -M_1, \ldots, M_2$. To achieve that $\tilde{u}(t)$ is uncorrelated with $r(t-\tau)$ for all τ ($\Phi_{\tilde{u}r} = 0$) we should let both M_1 and M_2 tend to infinity in (24), which then becomes an infinite set of linear equations. However, the optimal solution can still readily be found using Fourier techniques and it is given by

$$S_{opt}(e^{i\omega}) = \frac{\Phi_{ur}(\omega)}{\Phi_r(\omega)} \tag{25}$$

This l.l.m.s. estimator is often called the *Wiener smoother* [2]. In the linear feedback case the optimal solution is $S_{opt}(e^{i\omega}) = S_0(e^{i\omega})$ which of course is a causal operator. With linear feedback it is therefore no restriction to consider only causal models as in the two-stage method. However, in the nonlinear case the Wiener smoother can be noncausal even though the mapping from r to u is causal. In such cases it would thus be suboptimal to restrict to causal models.

Example 5 *One might be tempted to think that the Wiener smoother cannot be noncausal if the mapping from r to u is causal (be it linear or nonlinear).*

However, in general no such guarantees can be given. A simple counterexample can be constructed as follows. Let $r(t) = \epsilon(t) + \epsilon^2(t-1) - \lambda$, *where* $\epsilon(t)$ *is a sequence of independent Gaussian random variables with zero mean and variance* λ, *and let* $u(t) = r^2(t) - Er^2(t)$. *Then* $Er(t)r(t-\tau) = (\lambda + 2\lambda^2)\delta_\tau$ *and* $Eu(t)r(t-\tau) = 8\lambda^3\delta_\tau + 2\lambda^2\delta_{\tau+1}$ *(*$\delta_{(\cdot)}$ *denotes the Kronecker delta operator). It follows that the Wiener smoother is* $S_{opt}(q) = \frac{2\lambda}{1+2\lambda}q + \frac{8\lambda^2}{1+2\lambda}$ *which is noncausal although the mapping from* r *to* u *is causal.* □

To find the optimal solution when Φ_{ur} and Φ_r are unknown we can, for instance, use spectral analysis as in [1]. In the projection method a parametric approach is used and the optimal, doubly infinite ($M_1 \to \infty$, $M_2 \to \infty$) FIR filter is replaced by a finite approximation. If M_1 and M_2 are chosen large enough we will achieve $\Phi_{r\tilde{u}} \approx 0$, which implies that the resulting model will be essentially bias free, regardless of the nature of the feedback.

So far we have mainly discussed the asymptotic properties of the projection method. In the simulation study below we will illustrate the finite sample behavior of this method and it will then be clear that the noncausal FIR model used in the first step of the projection method can be used to decrease the *de facto* correlation between r and u and which will lead to improved results, compared to the two-stage method, even though the asymptotically optimal Wiener smoother is causal.

5 Asymptotic Variance Properties

This section will contain a brief discussion on the asymptotic variance properties of the projection (two-stage) method. Exact analysis of the asymptotic variance of the parameter estimates is not easy in the general case and we will not go into any further details on this here. However, to get some insight into the variance properties we rewrite the true system (1) as

$$y(t) = G_0(q)\hat{u}(t) + w(t), \quad w(t) = G_0(q)(u(t) - \hat{u}(t)) + v(t) \tag{26}$$

As can be seen, the difference compared to (1) is that we have replaced u by $\hat{u}$ and v by w. It follows that we may regard the second step of the projection method as a direct identification of the open-loop system using measurements of y and $\hat{u}$. Typically we will have

$$\Phi_{\hat{u}} < \Phi_u \quad \text{and} \quad \Phi_w > \Phi_v \tag{27}$$

implying that

$$\frac{\Phi_{\hat{u}}}{\Phi_w} < \frac{\Phi_u}{\Phi_v} \tag{28}$$

This shows that the signal to noise ratio will be worse than for direct identification using y and u, and we can thus expect that the accuracy of the projection method will be suboptimal. The Monte Carlo simulation presented in the next section illustrates this quite clearly.

If u is generated as in (21) and $\hat{u} = S_0 r$ we have that $w = S_0 v$ and we are back in an open-loop situation (since r and v are independent) and all standard open-loop results hold. As an example we have [4] (G'_θ is short for $d/d\theta G(e^{i\omega}, \theta)|_{\theta=\theta_0}$):

Theorem 6 *Consider the estimate $\hat{\theta}_N$ defined by (8) where $\hat{y}(t|\theta)$ is determined from the model structure (12) which we assume is uniformly stable. Suppose that Assumptions 1 and 2 hold and that the input is generated as in (21) and that $\hat{u} = S_0 r$. Suppose also that for a unique value θ_0 interior to $D_\mathcal{M}$ such that $G(q, \theta_0) = G_0(q)$ we have*

$$\hat{\theta}_N \to \theta_0, \quad \text{w.p. } 1 \text{ as } N \to \infty \tag{29}$$

$$R = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Phi_{\hat{u}}}{|H_*|^2} G'_\theta G'^*_\theta d\omega > 0 \tag{30}$$

$$\sqrt{N} E \left[\frac{1}{N} \sum_{t=1}^{N} [\psi(t, \theta_0)\varepsilon(t, \theta_0) - \bar{E}\psi(t, \theta_0)\varepsilon(t, \theta_0)] \right] \to 0, \text{ as } N \to \infty \tag{31}$$

where ψ denotes the negative gradient of ε. Then

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, P_\theta) \tag{32a}$$

$$P_\theta = R^{-1} Q R^{-1} \tag{32b}$$

$$Q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Phi_w \Phi_{\hat{u}}}{|H_*|^4} G'_\theta G'^*_\theta d\omega \tag{32c}$$

Remark: In the situation considered in Theorem 6 we can also write the matrices R and Q as

$$R = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|S_0|^2 \Phi_r}{|H_*|^2} G'_\theta G'^*_\theta d\omega \tag{33}$$

$$Q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|S_0|^4 \Phi_v \Phi_r}{|H_*|^4} G'_\theta G'^*_\theta d\omega \tag{34}$$

□

6 Simulation Study

In this section we illustrate the performance of the closed-loop methods presented earlier when applied to a system with a nonlinearity in the loop. The set-up is

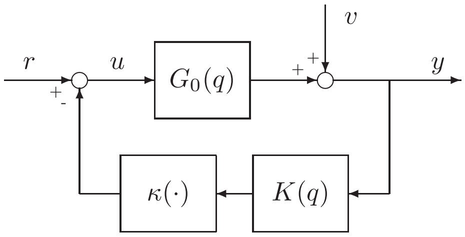

Figure 1 *Closed-loop system with nonlinear feedback.*

illustrated in Figure 1. The system G_0 and the noise model H_0 (cf. (1)) are given by

$$G_0(q) = \frac{b_1^0 q^{-1}}{1 + a_1^0 q^{-1} + a_2^0 q^{-2}}, \quad H_0(q) = 1 \tag{35}$$

where $b_1^0 = 1$, $a_1^0 = -1.2$, and $a_2^0 = 0.61$. The feedback controller is a proportional controller $K(q) = 0.25$ and at the output of the controller there is a static nonlinear element κ given by

$$\kappa(x) = \begin{cases} x + 2.5, & \text{if } x \geq 0 \\ x - 2.5, & \text{if } x < 0 \end{cases} \tag{36}$$

The goal is to identify the open-loop system G_0 using the simulated closed-loop data. Four different methods will be considered: The direct, the indirect, the two-stage, and the projection methods. In the direct method the model was

$$y(t) = G(q,\theta)u(t) + e(t) = \frac{b_1 q^{-1}}{1 + a_1 q^{-1} + a_2 q^{-2}} u(t) + e(t) \tag{37}$$

In the indirect method we used

$$y(t) = \frac{G(q,\theta)}{1 + G(q,\theta)K(q)} r(t) + e(t) \tag{38}$$

where $G(q,\theta)$ as in (37). (This model would give a consistent estimate of G_0 if the nonlinearity was not present.) In the first step of the two-stage method we used a 20th order ARX model and for the projection method the noncausal FIR filter had 41 taps ($M_1 = M_2 = 20$ in Eq. (11)). In the second step of these method we used a model of the form (37) except that u was replaced by $\hat{u}$.

In order to high-light the bias errors in the resulting estimates we first performed a noise-free simulation of the system. The data was generated using a unit variance, Gaussian, white noise reference signal. $N = 512$ data points were collected. The results of the identification are illustrated in Figure 2. With the direct method the true system is retrieved exactly, hence this estimate is not plotted separately. The indirect method performs very badly, which could be expected since there is a considerable error in the assumed feedback law. It is also clear that the projection method is able to model the true system quite accurately while the two-stage method performs less well. The identification results are summarized in Table B.1.

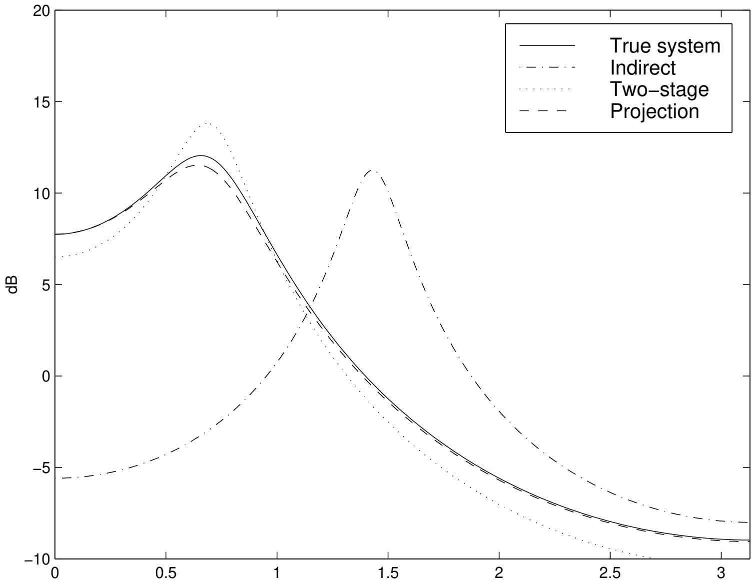

Figure 2 *Bode plot of true and estimated transfer functions.*

Table B.1 *True and estimated parameter values. (D = Direct method; I= Indirect method; T = Two-stage method; P = Projection method)*

	True value	D	I	T	P
a_1	−1.2000	−1.2000	−0.2456	−1.2906	−1.1887
a_2	0.6100	0.6100	0.7777	0.7163	0.5899
b_1	1.0000	1.0000	0.8050	0.9003	0.9797

Note that in this example r is a white noise signal and asymptotically there is no noncausal correlation between r and u so the Wiener smoother is causal. The reason why the projection method outperforms the two-stage method in this situation is that for a finite data set, like this one, the noncausal model used in the projection method reduces the actual correlation between $\hat{u}$ and $\tilde{u}$ and, consequently, improves the quality of the resulting G-estimate. To further illustrate this we studied the impulse response coefficients of the first-step models in the two methods. The

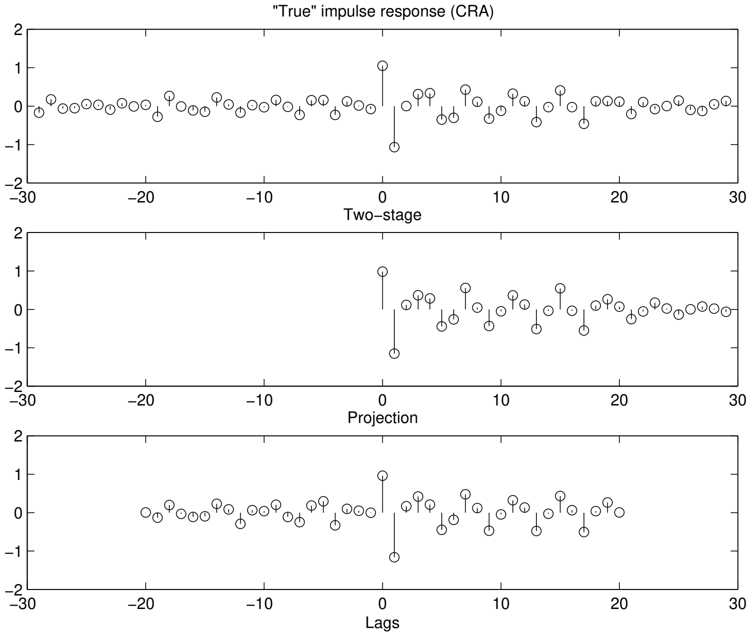

Figure 3 *Impulse response estimated using correlation analysis (top) and impulse responses for the first-step models in the two-stage (middle) and the projection (bottom) methods.*

top plot in Figure 3 shows the estimated impulse response of the map from the reference r to the input u obtained using correlation analysis. This illustrates the actual correlation between $u(t)$ and $r(t-k)$ for different k. The middle and bottom plots show the impulse responses for the first-step models in the two-stage and projection methods, respectively. It is clear that the causal part of the impulse response is quite accurately modeled in the two-stage method while the noncausal part is zero. However, the noncausal FIR filter used in the projection method, which approximates the impulse response well for lags between the chosen limits -20 and 20, apparently picks up more of the significant correlation between r and u and hence gives better results.

To illustrate the suboptimal accuracy of the projection method we performed a simulation study consisting of 128 Monte Carlo simulations. In each run we collected 4096 samples. The noise e was chosen as a unit variance, Gaussian, white noise sequence, independent of the reference signal r. The reference signal was also chosen as a unit variance, Gaussian, white noise sequence, just as before. In this simulation only the optimal direct method and the projection method were compared. In the projection method we used a noncausal FIR filter with 41 taps, also as before. The results of the Monte Carlo simulation are presented in Table B.2. The values shown are the mean values of the parameter estimates together with their estimated standard deviations. Clearly the accuracy of the projection

Table B.2 *Summary of results of Monte Carlo simulation.*

	True value	D	P
a_1	-1.2000	-1.2002	-1.1659
		± 0.0042	± 0.0196
a_2	0.6100	0.6100	0.5894
		± 0.0046	± 0.0146
b_1	1.0000	1.0009	0.9929
		± 0.0055	± 0.0181

method is worse than that of the direct method, as expected. There is also a slight bias in the parameter estimates obtained with the projection method which is due to a nonvanishing correlation between $\hat{u}$ and $\tilde{u}$. One reason for this is that we used rather small M_1 and M_2, another is that the estimated coefficients in the noncausal FIR filter are not perfect due to the noise in u. The latter effect tends to be worse as model order increases, while the former decreases with increasing model order. In a practical situation it is therefore recommended to try different values of M_1 and M_2 to optimize the results.

7 Conclusions

The projection method consists of two steps: First the mapping from the reference signal r to the input u is modeled using a noncausal FIR filter. This gives an estimate $\hat{u}$ of u that is asymptotically uncorrelated with $u - \hat{u}$. In the second step the open-loop system is identified using measurements of y and $\hat{u}$. The method gives consistent estimates regardless of the nature of the feedback but the accuracy is suboptimal. In case of undermodeling, the model can be fit to the data with arbitrary frequency weighting which is a clear advantage compared to the direct method.

References

[1] H. Akaike. Some problems in the application of the cross-spectral method. In B. Harris, editor, *Spectral Analysis of Time Series*, pages 81–107. John Wiley & Sons, 1967.

[2] T. Kailath. *Lectures on Wiener and Kalman Filtering*. Springer-Verlag, 1981.

[3] L. Ljung. Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, 23:770–783, 1978.

[4] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[5] P. M. J. Van den Hof and R. J. P. Schrama. An indirect method for transfer function estimation from closed loop data. *Automatica*, 29:1523–1527, 1993.

Paper C

Efficiency of Prediction Error and Instrumental Variable Methods for Closed-loop Identification

Authors: Urban Forssell and Chun Tung Chou

Edited version of paper originally published in *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 1287-1288, Tampa, FL, 1998.

Efficiency of Prediction Error and Instrumental Variable Methods for Closed-loop Identification

Urban Forssell[1]

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: ufo@isy.liu.se.

Chun Tung Chou[2]

Department of Electrical Engineering,
Delft University of Technology,
P.O. Box 5031, 2600 GA Delft,
The Netherlands.
E-mail: chou@harding.et.tudelft.nl.

Abstract

We study the efficiency of a number of closed-loop identification methods. Results will be given for methods based on the prediction error approach as well as those based on the instrumental variable approach. Moreover, interesting insights in the properties of a recently suggested subspace method for closed-loop identification are obtained by exploring the links between this method and the instrumental variable method.

1 Introduction

The accuracy of different prediction error (PE) methods for closed-loop identification has been studied in, e.g., [3]. Here we will review these results and link them to corresponding results for instrumental variable (IV) methods for closed-loop identification, e.g., [5]. By studying the special case of single-input single-output

[1]U. Forssell is supported by the Swedish Research Council for Engineering Sciences.

[2]C.T. Chou is supported by the Dutch Technology Foundation (STW) under project number DEL55.3891.

ARMAX models it is possible to compare these methods with a modified, but yet representative, form of a recently proposed errors-in-variables (EIV) subspace method [1], that can be applied to closed-loop data.

2 Preliminaries

We assume that the true system is given by the ARMAX model

$$A_0(q)y(t) = B_0(q)u(t) + C_0(q)e(t),\ Ee(s)e(t) = \lambda_0\delta_{s-t} \tag{1}$$

and that there is a stabilizing LTI feedback between $u(t)$ and $y(t)$ given by

$$R(q)u(t) = R(q)r(t) - S(q)y(t) \tag{2}$$

where the reference signal $r(t)$ is independent of $e(t)$.

To identify this system we will work with models of the form

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \tag{3}$$

where $\varphi(t)$ is contains the regressors and θ the parameters of the A- and B-polynomials. We will frequently use the notation

$$\varphi(t) = \varphi_r(t) + \varphi_e(t) \tag{4}$$

where $\varphi_r(t)$ and $\varphi_e(t)$ are the parts of $\varphi(t)$ that are due to $r(t)$ and $e(t)$, respectively. In some cases the model (3) will be applied together with a (possibly parameter-dependent) monic prefilter L that can be used to emphasize certain frequency regions. The resulting parameter vector estimate will be denoted $\hat{\theta}$.

The main focus in the paper will be to characterize the asymptotic covariance matrix P_θ,

$$N \operatorname{Cov} \hat{\theta} \to P_\theta,\ \text{as } N \to \infty \tag{5}$$

for a number of identification methods that guarantee that the parameter estimates converge to the true values also when applied to closed-loop data.

3 Prediction Error Methods

The PE estimate is found as the straightforward fit:

$$\hat{\theta} = \arg\min_{\theta,\eta} \frac{1}{N}\sum_{t=1}^{N}[L(q,\eta)(y(t) - \hat{y}(t|\theta))]^2 \tag{6}$$

This approach can also be applied to closed-loop identification. Depending on how $L(q,\eta)$ is parameterized, different methods are obtained.

If we think of $L(q,\eta)$ as an inverse noise model and parameterize in a straightforward way the direct method is obtained. If the chosen parameterization is sufficiently flexible, the direct method is consistent and gives optimal accuracy [4, 3]. It can be shown that (see [2]) for the direct method we have ('D' stands for *direct*)

$$P_\theta^{\mathrm{D}} = \lambda_0 \Big[\bar{E}\, \frac{1}{C_0(q)} \varphi_r(t) \frac{1}{C_0(q)} \varphi_r^T(t) + \bar{R}_e \Big]^{-1} \tag{7}$$

where $\bar{R}_e \geq 0$ depends on $e(t)$ and not on $r(t)$. (The symbol $\bar{E}$ is defined in, e.g., [4].) This means that the noise actually helps reducing the variance in the direct method. In case the optimal, fixed prefilter $L(q,\eta) = L(q) = 1/C_0(q)$ is used we can write (7) as

$$P_\theta^{\mathrm{D}} = \lambda_0 \Big[\bar{E}\, \frac{1}{C_0(q)} \varphi(t) \frac{1}{C_0(q)} \varphi^T(t) \Big]^{-1} \triangleq P_\theta^{\mathrm{OD}} \tag{8}$$

If we parameterize the prefilter as

$$L(q,\eta) = L(q,\theta) = \frac{L_1(q)R(q)}{A(q)R(q) + B(q)S(q)} \tag{9}$$

we obtain the indirect method (see, e.g., [2]). For this method the asymptotic covariance matrix satisfies

$$P_\theta^{\mathrm{ID}} \geq \lambda_0 \Big[\bar{E}\, \frac{1}{C_0(q)} \varphi_r(t) \frac{1}{C_0(q)} \varphi_r^T(t) \Big]^{-1} \triangleq P_\theta^{\mathrm{OID}} \tag{10}$$

for all choices of the filter $L_1(q)$. Equality holds if
$L_1(q) = (A_0(q)R(q) + B_0(q)S(q))/(C_0(q)R(q))$. Note that $P_\theta^{\mathrm{OID}} \geq P_\theta^{\mathrm{D}}$, due to the term $\bar{R}_e \geq 0$ in (7).

4 Instrumental Variable Methods

Consider the linear regression (3). Let $\zeta(t)$ denote an IV vector (of the same size as $\varphi(t)$) and $L(q)$ be a prefilter, then with the standard IV method the parameter estimate is computed as

$$\hat{\theta} = \Big[\frac{1}{N} \sum_{t=1}^{N} \zeta(t) L(q) \varphi^T(t) \Big]^{-1} \frac{1}{N} \sum_{t=1}^{N} \zeta(t) L(q) y(t) \tag{11}$$

given that the indicated inverse exists. Consistency requires the instruments be chosen such that

$$\bar{E}[\zeta(t)L(q)C_0(q)e(t)] = 0 \tag{12}$$

Depending on the choices of instruments $\zeta(t)$ and prefilter $L(q)$ different instrumental variable "methods" result. For all choices of $\zeta(t)$ and $L(q)$,

$$P_\theta^{\mathrm{IV}} \geq P_\theta^{\mathrm{OD}} \tag{13}$$

Equality holds if $\zeta(t) = 1/C_0(q)\varphi(t)$ and $L(q) = 1/C_0(q)$ (given that these choices satisfy (12)).

If the instruments are chosen as filtered versions of $r(t)$ then (12) will be automatically satisfied since $r(t)$ and $e(t)$ are independent. The resulting method will here be denoted RIV ('R' for reference signal). For this method it can be shown that (see [5]) for any prefilter $L(q)$ and any $\zeta(t)$ constructed from $r(t)$ through filtering we have that

$$P_\theta^{\mathrm{RIV}} \geq P_\theta^{\mathrm{OID}} \tag{14}$$

Equality holds if $\zeta(t) = \frac{1}{C_0(q)}\varphi_r(t)$ and $L(q) = \frac{1}{C_0(q)}$.

Another possibility is to use delayed versions of the regression vector, $\zeta(t) = \varphi(t-k)$. To satisfy (12), the delay k has to be larger than the maximal delay in $L(q)C_0(q)$. This method will be referred to as PIV ('P' for past regressors). The covariance matrix becomes

$$P_\theta^{\mathrm{PIV}} = \lambda_0[\bar{E}\,\varphi(t-k)L(q)\varphi^T(t)]^{-1}[\bar{E}\,\tilde{\varphi}(t-k)\tilde{\varphi}^T(t-k)][\bar{E}\,\varphi(t-k)L(q)\varphi^T(t)]^{-T} \tag{15}$$

where

$$\tilde{\varphi}(\tau) = L(q^{-1})C_0(q^{-1})\varphi(\tau) \tag{16}$$

This can be better or worse than the RIV method depending, e.g., on how $L(q)$ and the delay k are chosen. When the feedback loop is noisy it is likely that PIV gives better accuracy than RIV [2]. Under all circumstances $P_\theta^{\mathrm{PIV}} \geq P_\theta^{\mathrm{OD}}$.

5 Discussion

The direct method gives optimal accuracy, while the other PE and IV methods studied in this paper all give worse accuracy. A key point in understanding this difference is that in the direct method the whole input spectrum is used to reduce the variance, whereas in the indirect and RIV methods only the noise-free part is used. In the PIV method the whole input spectrum is used, but, like all IV methods, this method suffers from the fact that cross-correlation, rather than the auto-correlation, is used to reduce the variance. The special case is when the noise model is known exactly, then one can use the auto-correlation to achieve minimum variance on par with the direct method. The use of a periodic reference signal can also increase the correlation between the instruments and the regressors, and hence improve the accuracy. To balance the picture somewhat we also point out that with an increasing SNR the differences between the methods will disappear and all will perform similarly.

The EIV method [1] can to a certain extent be interpreted as PIV with the prefilter chosen to be unity [2]. This gives new insights into the statistical properties of EIV. We conjecture that with a low SNR this method is likely to be better than RIV and the indirect method, although the efficiency may still be sub-optimal. This erratic type of statistical behavior of EIV is also shared by some other subspace algorithms, as pointed out in [6], and can be corrected by using say subspace fitting [6].

References

[1] C. T. Chou and M. Verhaegen. Subspace algorithms for the identification of multivariable errors-in-variables models. *Automatica*, 33:1857–1869, 1997.

[2] U. Forssell and C.T. Chou. Efficiency of prediction error and instrumental variable methods for closed-loop identification. Technical Report LiTH-ISY-R-2015, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998.

[3] U. Forssell and L. Ljung. Closed-loop identification revisited – Updated version. Technical Report LiTH-ISY-R-2021, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998.

[4] L. Ljung. *System Identification: Theory for the User.* Prentice-Hall, 1987.

[5] T. Söderström, P. Stoica, and E. Trulsson. Instrumental variable methods for closed loop systems. In *Preprints IFAC 10th World Congress on Automatic Control*, pages 364–369, Munich, Germany, July 1987.

[6] M. Viberg, B. Wahlberg, and B. Ottersten. Analysis of state space system identification methods based on instrumental variables and subspace fitting. *Automatica*, 33:1603–1616, 1997.

Paper D

Identification of Unstable Systems Using Output Error and Box-Jenkins Model Structures

Authors: Urban Forssell and Lennart Ljung

Edited version of paper originally published in *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3932-3937, Tampa, FL, 1998.

To appear in *IEEE Transactions on Automatic Control*.

Identification of Unstable Systems Using Output Error and Box-Jenkins Model Structures

Urban Forssell and Lennart Ljung

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: ufo@isy.liu.se, ljung@isy.liu.se.

Abstract

It is well known that the output error and Box-Jenkins model structures cannot be used for prediction error identification of unstable systems. The reason for this is that the predictors in this case generically will be unstable. Typically this problem is handled by projecting the parameter vector into the region of stability which gives erroneous results when the underlying system is unstable. The main contribution of this work is that we derive modified, but asymptotically equivalent, versions of these model structures that can be applied also in the case of unstable systems.

1 Introduction

In this note we will discuss prediction error identification of unstable systems using output error and Box-Jenkins model structures. As is well known from text books on system identification in the prediction error framework (e.g., [5, 9]), it is required that the predictors are stable. In case the parameters are estimated using some numerical search algorithm that requires gradients of the predictor to be computed, these must also be stable. With the ARX and ARMAX model structures this is no problem since the dynamics model and the noise model share denominator polynomials, which cancel when the predictors are formed. For the output error and Box-Jenkins model structures this is not the case and if the underlying system is unstable the predictors will generically be unstable, which seemingly makes the model structures inapplicable in these cases. Also, when identifying stable systems that are "close" to being unstable it might be that the predictors become unstable in one or more of the steps in the search algorithm due to numerical problems, so that the search has to be terminated before the global optimum is reached.

Traditionally the problem of unstable predictors has been handled by projecting the parameter vector into the region of stability. This will, of course, lead to completely useless results if the underlying system is unstable. If the system is stable but the predictors become unstable in some intermediate step in the numerical search, projecting the parameter vector into the region of stability can lead to convergence problems. However, as we shall see it is possible to re-parameterize the output error and Box-Jenkins model structures to guarantee stability of the predictors, and this *without increasing the total number of parameters to be estimated*. With the new versions of the output error and Box-Jenkins model structures we thus gain two things: these model structures can be applied to unstable systems and the numerical properties of the search algorithm are improved. The price we pay is an increase in complexity of the search algorithm in case the predictors are unstable. (If they are stable we can use the standard algorithms.)

If the system is unstable, we will assume that the experimental data are generated under stabilizing feedback. It is clear that such feedback will involve some knowledge of the system. However this knowledge may by quite rudimentary, and there may still be an obvious need for an improved system model. One example – among many – is flight testing unstable aircraft to build accurate dynamic models.

With a stabilizing controller in the loop we are faced with a *closed-loop* identification problem. Closed-loop identification is often used in connection to so called control-relevant identification where the goal is to estimate models that are suitable for (robust) control design, see, e.g., the surveys [3, 11, 2]. It is then often only interesting to model the dynamics of the plant, the noise properties are less interesting, so that it would be natural to use an output error model structure. However, since unstable plants cannot be handled using output error models the conclusion has been that this approach cannot be used when the plant is unstable. Alternative solutions have been suggested in, e.g., [4, 10, 12]. Unfortunately these methods are considerably more involved than a direct application of an output error or a Box-Jenkins model to the closed-loop data.

A problem when identifying systems in closed-loop directly is that the results will be biased unless the noise model accurately describes the true noise characteristics [5, 6, 7]. This has traditionally been a main issue in the closed-loop identification literature that has further motivated the search for alternative closed-loop identification methods. The bias problems will of course also be present when using the new model structures suggested in this paper. However, in many cases the bias errors due to poor modeling of the true noise characteristics will be small, especially if the signal-to-noise ratio is high [7, 2]. It must also be realized that in practice the model errors will be due to *both* bias and variance errors [5] and if a reasonably flexible noise model is used, the bias error due to the feedback will typically be small compared to the variance error. The variance error also increases with the number of parameters which favors model structures with few parameters. We would also like to point out that most other closed-loop identification methods, that are designed to give unbiased estimates, give higher variance errors than the direct method [1, 8]. This is an issue that in our opinion has received too little attention in the closed-loop identification literature.

The organization of the paper is as follows. Next, in Section 2, we study some basic facts on prediction error methods and in Section 3 we discuss some standard choices of model structures. This section also contains an illustration of the problems faced when trying to identify unstable systems using output error models. Section 4 contains the main result of the paper: How the standard output error model structure should be modified to cope also with unstable systems. After introducing some additional notation we present the basic idea and go through the derivation of the required gradient filters in some detail. This is then followed by a simulation study that illustrates the feasibility of the idea. Before concluding we, in Section 5, briefly mention the corresponding changes of the Box-Jenkins model structure that are necessary to make it applicable to unstable systems.

2 Some Basics in Prediction Error Identification

In prediction error identification one typically considers linear model structures parameterized in terms of a parameter vector θ:

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \tag{1}$$

Here $G(q, \theta)$ and $H(q, \theta)$ are rational functions of q^{-1}, the unit delay operator ($q^{-1}u(t) = u(t-1)$, etc.) parameterized in terms of θ; $y(t)$ is the output; $u(t)$ is the input; $e(t)$ is white noise. In this paper we will limit the study to single input single output models for ease of exposition.

Typically θ ranges over some open subset $D_{\mathcal{M}}$ of $\mathbb{R}^d$ ($d = \dim \theta$):

$$\theta \in D_{\mathcal{M}} \subset \mathbb{R}^d \tag{2}$$

Note the distinction between a model and a model structure: When viewed as a function of θ, (1) is a *model structure*, while for a fixed $\theta = \theta^*$, (1) is a *model*.

Furthermore, we also have that (1) together with (2) defines a *model set* and it is our task to find the best model in the set, typically by performing a numerical search over all possible models. Refer to [5] for a comprehensive treatment, including exact definitions, of the concepts model, model structure and model set.

The one-step-ahead *predictor* for (1) is

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t) \tag{3}$$

Here it is required that the filters $H^{-1}(q,\theta)G(q,\theta)$ and $(1 - H^{-1}(q,\theta))$ are stable for the predictor to be well defined. It is easy to see that this calls for an inversely stable noise model $H(q,\theta)$ and that the unstable poles of $G(q,\theta)$ are also poles of $H(q,\theta)$. These issues will play important roles in this paper.

The prediction errors $\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta)$ corresponding to the predictor (3) are

$$\varepsilon(t,\theta) = H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{4}$$

We will also use the following notation for the gradient of $\hat{y}(t|\theta)$:

$$\psi(t,\theta) = \frac{d}{d\theta}\hat{y}(t|\theta) \ (= -\frac{d}{d\theta}\varepsilon(t,\theta)) \tag{5}$$

In the standard case of least-squares prediction error identification one calculates the parameter estimate as the minimizing argument of the criterion function

$$V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N}\frac{1}{2}\varepsilon^2(t,\theta) \tag{6}$$

Typically one finds the estimate through some numerical search routine of the form

$$\hat{\theta}_N^{(i+1)} = \hat{\theta}_N^{(i)} - \mu_N^{(i)}[R_N^{(i)}]^{-1}V_N'(\hat{\theta}_N^{(i)}) \tag{7}$$

where V_N' denotes the gradient of the criterion function, R_N is a matrix that modifies the search direction and μ_N a scaling factor that determines the step length. From (5) we see that

$$V_N'(\theta) = -\frac{1}{N}\sum_{t=1}^{N}\psi(t,\theta)\varepsilon(t,\theta) \tag{8}$$

and typically R_N is chosen approximately equal to the Hessian V_N'' (which would make (7) a Newton algorithm); a standard choice is

$$R_N = \frac{1}{N}\sum_{t=1}^{N}\psi(t,\theta)\psi^T(t,\theta) + \delta I \tag{9}$$

where $\delta \geq 0$ is chosen so that R_N becomes positive definite. This is also called the Levenberg-Marquardt regularization procedure.

Clearly it is required that both the predictor (3) and the gradient (5), which have to be computed and used in the search algorithm (7), are stable. When dealing with unstable systems this introduces constraints on the possible model structures.

3 Commonly Used Model Structures

With these stability requirements in mind, let us now discuss some standard choices of model structures.

A quite general model structure is the following [5]:

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{10}$$

where

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \tag{11}$$

and similarly for the C, D, and F polynomials, while

$$B(q) = q^{-n_k}(b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}) \tag{12}$$

This model structure includes some common special cases:

1. $C(q) = 1, D(q) = 1, F(q) = 1$, an ARX model structure.
2. $D(q) = 1, F(q) = 1$, an ARMAX model structure.
3. $A(q) = 1, C(q) = 1, D(q) = 1$, an output error model structure.
4. $A(q) = 1$, a Box-Jenkins model structure.

In the sequel we will assume $n_k = 0$ (which always can be achieved by replacing $u(t)$ by $u(t - n_k)$).

A sufficient condition for the predictor and gradient filters to be stable is that $C(q) \cdot F(q)$ is stable for all $\theta \in D_{\mathcal{M}}$ (cf. Lemma 4.1 in [5]). Note that this condition is automatically satisfied for ARX models, and for ARMAX models it is sufficient that the C-polynomial is stable, which does not impose any stability constraints on the dynamics model. For identification of unstable system these model structures thus are natural choices.

The output error model structure has a fixed noise model ($H(q, \theta) = 1$) and is a natural choice if only a model of the system dynamics is required. In case one wants to model also the noise characteristics (e.g., to improve the efficiency) but do not want the noise and dynamics models to be dependent as in the ARX and ARMAX cases, then the Box-Jenkins model structure would the one to choose. However, if the underlying system is unstable these model structures can not be used without modifications, e.g., the ones we propose in this paper. To see where the problem lies, let us study the output error case.

Suppose that we want to identify an unstable system, stabilized by some controller, and that we are only interested in modeling the dynamics with no modeling effort spent on the noise characteristics. Then the natural choice would be to use an output error model structure:

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t) \tag{13}$$

Now, since the system is unstable the predictor

$$\hat{y}(t|\theta) = \frac{B(q)}{F(q)}u(t) \tag{14}$$

as well as the gradient filters

$$\frac{\partial}{\partial b_k}\hat{y}(t|\theta) = \frac{1}{F(q)}u(t-k) \tag{15a}$$

$$\frac{\partial}{\partial f_k}\hat{y}(t|\theta) = -\frac{B(q)}{F^2(q)}u(t-k) \tag{15b}$$

will generically be unstable. When implementing a parameter estimation algorithm for the output error case one typically secures stability in every iteration of the algorithm (7) by projecting the parameter vector into the region of stability. For unstable systems this of course leads to erroneous results.

These problems are also present when using the Box-Jenkins model structure:

$$y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{16}$$

Also in this case one has to resort to projections into the region of stability to ensure stability of the predictors and gradient filters, which makes this model structure in its standard form useless for identification of unstable systems.

In the following sections we will describe how to modify these model structures to avoid these problems.

4 An Alternative Output Error Model Structure

4.1 Some Additional Notation

Let $F_s(q)$ ($F_a(q)$) be the stable (anti-stable), monic part of $F(q)$:

$$F(q) = F_s(q)F_a(q) \tag{17}$$

and let the polynomials $F_s(q)$ and $F_a(q)$ be parameterized as

$$F_s(q) = 1 + f_{s,1}q^{-1} + \cdots + f_{s,n_{f_s}}q^{-n_{f_s}} \tag{18}$$

$$F_a(q) = 1 + f_{a,1}q^{-1} + \cdots + f_{a,n_{f_a}}q^{-n_{f_a}} \tag{19}$$

With the notation

$$\bar{f}_{a,k} = \begin{cases} 1 & k = 0 \\ f_{a,k} & 1 \leq k \leq n_{f_a} \\ 0 & \text{else} \end{cases} \tag{20}$$

and

$$\bar{f}_{s,k} = \begin{cases} 1 & k = 0 \\ f_{s,k} & 1 \leq k \leq n_{f_s} \\ 0 & \text{else} \end{cases} \tag{21}$$

we have

$$f_k = \sum_{j=0}^{n_f} \bar{f}_{s,j}\bar{f}_{a,k-j}, \ \ k = 1, 2, \ldots, n_f \tag{22}$$

Furthermore, let $F_a^*(q)$ denote the monic, stabilized F_a-polynomial, i.e., $F_a^*(q)$ is the monic polynomial whose zeros are equal to the zeros of $F_a(q)$ reflected into the unit disc. In terms of $f_{a,i}$, the coefficients of $F_a(q)$, we can write $F_a^*(q)$ as

$$F_a^*(q) = 1 + \frac{f_{a,n_{f_a}-1}}{f_{a,n_{f_a}}}q^{-1} + \cdots + \frac{1}{f_{a,n_{f_a}}}q^{-n_{f_a}} \tag{23}$$

(Here we have used the implicit assumption that $f_{a,n_{f_a}} \neq 0$.)

4.2 The Proposed Model Structure

Now consider the following modified output error model structure:

$$y(t) = \frac{B(q)}{F(q)}u(t) + \frac{F_a^*(q)}{F_a(q)}e(t) \tag{24}$$

with the predictor

$$\begin{aligned}\hat{y}(t|\theta) &= \frac{F_a(q)B(q)}{F_a^*(q)F(q)}u(t) + (1 - \frac{F_a(q)}{F_a^*(q)})y(t) \\ &= \frac{B(q)}{F_a^*(q)F_s(q)}u(t) + (1 - \frac{F_a(q)}{F_a^*(q)})y(t)\end{aligned} \tag{25}$$

The difference between this model structure and the basic output error model structure (13) is thus that we have included a noise model $F_a^*(q)/F_a(q)$ and obtained a different "dummy" noise term

$$\bar{e}(t) = \frac{F_a^*(q)}{F_a(q)}e(t) \tag{26}$$

instead of just $e(t)$. At first glance it may thus seem as the model structures (13) and (24) will give different results, but in fact they are (asymptotically) equivalent as can be seen from the following result:

Proposition 1 *When applying a prediction error method to the model structures* (13) *and* (24) *the resulting estimates will asymptotically, as $N \to \infty$, be the same.*

Proof *From classical prediction error theory we know that under mild conditions the limiting models will minimize the integral of the spectrum of the prediction errors (see, e.g., Theorem 8.2 in [5]). Now, if we let $\varepsilon_F(t,\theta)$ denote the prediction errors obtained with the model* (25) *and let $\varepsilon(t,\theta)$ denote the prediction errors corresponding to the model* (14)*, we have that the spectrum of $\varepsilon_F(t,\theta)$ is given by*

$$\Phi_{\varepsilon_F}(\omega) = \left|\frac{F_a(e^{i\omega})}{F_a^*(e^{i\omega})}\right|^2 \Phi_\varepsilon(\omega) \tag{27}$$

$$= |f_{a,n_{f_a}}|^2 \Phi_\varepsilon(\omega) \tag{28}$$

where $\Phi_\varepsilon(\omega)$ denotes the spectrum of $\varepsilon(t,\theta)$. Thus the spectra differ by only a constant scaling and hence the corresponding limiting models will be the same. □

As we have seen, the results will asymptotically be the same with both model structures; the difference is of course that the predictor (25) will always be stable along with all its derivatives even if $F(q)$ is unstable (as opposed to standard output error case which require a stable $F(q)$ for the predictor to be stable). Note that in (24) the noise model is monic and inversely stable and the unstable poles of the dynamics model are also poles of the noise model (cf. the discussion in Section 2).

The basic idea behind the equivalence result in Proposition 1 is really that a constant spectrum may be factorized in infinitely many ways using all-pass functions. Here we chose convienient pole locations for these all-pass functions to have stable predictors. This is actually very closely related to classical Kalman filter theory, as will be illustrated next.

4.3 Connections to the Kalman Filter

There is a very straightforward interpretation of the second order equivalent (25) of (14). Suppose we realize (13) in state-space form:

$$\begin{aligned} x(t+1) &= \mathcal{A}x(t) + \mathcal{B}u(t) \\ y(t) &= \mathcal{C}x(t) + e(t) \end{aligned} \tag{29}$$

The steady state Kalman filter predictor is

$$\begin{aligned} \hat{x}(t+1) &= (\mathcal{A} - \mathcal{K}\mathcal{C})\hat{x}(t) + \mathcal{B}u(t) + \mathcal{K}y(t) \\ \hat{y}(t|\theta) &= \mathcal{C}\hat{x}(t) \end{aligned} \tag{30}$$

where $\mathcal{K}$ is determined from an algebraic Riccati equation in the usual way: $\mathcal{K} = \mathcal{A}\Pi\mathcal{C}^T/(R+\mathcal{C}\Pi\mathcal{C}^T)$, $\Pi = \mathcal{A}\Pi\mathcal{A}^T - \mathcal{K}(R+\mathcal{C}\Pi\mathcal{C}^T)\mathcal{K}^T$. Here R is the variance of the innovations $e(t)$: $R = Ee^2(t) > 0$. (E denotes mathematical expectation.)

Now, if $\mathcal{A}$ is stable the solution is $\mathcal{K} = 0$ and (30) corresponds to (14). However if $\mathcal{A}$ is not stable, the solution to the Kalman predictor problem is a nonzero $\mathcal{K}$ which makes (30) exactly equal to (25). Note that this holds regardless of R.

We thus have two alternative ways of computing the gradient $\hat{y}(t|\theta)$ in the output error case: (25) and (30). As we have seen these two variants give identical solutions $\hat{y}(t|\theta)$ although (25) is computationally less demanding. In contrast with the Kalman filter solution (30), it is also straightforward to generalize (25) to other model structures, like the Box-Jenkins model structure. This is briefly discussed in Section 5 below. Another important issue is how to compute the gradient (5). Using the explicit predictor formula (25) this is relatively straightforward, albeit tedious, as shown in the next section.

4.4 Computation of the Gradient

As mentioned above, the gradient $\psi(t,\theta)$ is needed for the implementation of the search scheme (7). With the predictor (25) the expression for the gradient will be much more involved than (15) but for completeness we will go through these calculations in some detail (after all, the gradient is needed for the implementation of the estimation algorithm).

Given the predictor model (25) we have that

$$\frac{\partial}{\partial b_k}\hat{y}(t|\theta) = \frac{1}{F_s(q)F_a^*(q)}u(t-k) \tag{31}$$

while

$$\frac{\partial}{\partial f_k}\hat{y}(t|\theta) = \frac{\partial}{\partial f_k}\frac{B(q)}{F_s(q)F_a^*(q)}u(t) - \frac{\partial}{\partial f_k}\frac{F_a(q)}{F_a^*(q)}y(t) \tag{32}$$

Introducing

$$W_1^k(q) = \frac{\partial}{\partial f_k}F_s(q) \tag{33}$$

$$W_2^k(q) = \frac{\partial}{\partial f_k}F_a(q) \tag{34}$$

$$W_3^k(q) = \frac{\partial}{\partial f_k}F_a^*(q) \tag{35}$$

and

$$x_1(t) = -\frac{B(q)}{F_s^2(q)F_a^*(q)}u(t) \tag{36}$$

$$x_2(t) = -\frac{1}{F_a^*(q)}y(t) \tag{37}$$

$$x_3(t) = -\frac{B(q)}{F_s(q)(F_a^*(q))^2}u(t) + \frac{F_a(q)}{(F_a^*(q))^2}y(t) \tag{38}$$

we may write

$$\frac{\partial}{\partial f_k}\hat{y}(t|\theta) = W_1^k(q)x_1(t) + W_2^k(q)x_2(t) + W_3^k(q)x_3(t) \tag{39}$$

What we then finally need in order to be able to compute the gradient $\frac{\partial}{\partial f_k}\hat{y}(t|\theta)$ are explicit expressions for the filters $W_i^k,\ i = 1, 2, 3$. Using (22) we have that

$$W_1^k(q) = \sum_{i=1}^{n_{f_s}} w_{1,i}^k q^{-i},\ w_{1,i}^k = \begin{cases} \bar{f}_{a,k-i}^{-1} & k - n_{f_a} \leq i \leq k \\ 0 & \text{else} \end{cases} \tag{40}$$

$$W_2^k(q) = \sum_{i=1}^{n_{f_a}} w_{2,i}^k q^{-i},\ w_{2,i}^k = \begin{cases} \bar{f}_{s,k-i}^{-1} & k - n_{f_s} \leq i \leq k \\ 0 & \text{else} \end{cases} \tag{41}$$

while

$$W_3^k(q) = \frac{1}{f_{a,n_{f_a}}} \cdot \left[\sum_{i=1}^{n_{f_a}-1} w_{3,i}^k q^{-i} + w_{3,0}^k(1 - F_a^*(q))\right],$$

$$w_{3,i}^k = \begin{cases} \bar{f}_{s,k-n_{f_a}+i}^{-1} & n_{f_a} - k \leq i \leq n_{f_s} + n_{f_a} - k \\ 0 & \text{else} \end{cases} \tag{42}$$

The equations (31)-(42) together constitute a complete and explicit description of the gradient $\psi(t,\theta) = \frac{d}{d\theta}\hat{y}(t|\theta)$ which may be used in an implementation of the search algorithm (7).

4.5 Simulation Example

To illustrate the applicability of the proposed model structure (24) to identification problems involving unstable systems we will in this section present a small simulation study.

The "true" system – to be identified – is given by

$$y(t) = \frac{b_0}{1 + f_1 q^{-1} + f_2 q^{-2}} u(t) + e(t) \tag{43}$$

with $b_0 = 1$, $f_1 = -1.5$, and $f_2 = 1.5$. This system is unstable with poles in $0.75 \pm 0.9682i$.

To generate identification data we simulated this system using the feedback law

$$u(t) = r(t) - (-0.95q^{-2})y(t) = r(t) + 0.95y(t-2) \tag{44}$$

which places the closed-loop poles in 0.8618 and 0.6382. In the simulation we used independent, zero mean, Gaussian white noise reference and noise signals $\{r(t)\}$ and $\{e(t)\}$ with variances 1 and 0.01, respectively. $N = 200$ data samples were used.

In Table D.1 we have summarized the results of the identification, the numbers shown are the estimated parameter values together with their standard deviations. For comparison we have, apart from the model structure (24), used a standard output error model model structure and a second-order ARMAX model structure. As can be seen the standard output error model structure gives completely useless estimates while the modified output error and the ARMAX model structures give very similar and accurate results.

Table D.1 *Summary of identification results.*

Parameter	b_0	f_1	f_2
True value	1	-1.5	1.5
OE	0.6218	-1.0336	0.7389
	± 0.0382	± 0.0481	± 0.0391
Modified OE	0.9898	-1.5008	1.5035
	± 0.0084	± 0.0052	± 0.0079
ARMAX	1.0028	-1.5026	1.5052
	± 0.0077	± 0.0049	± 0.0074

5 An Alternative Box-Jenkins Model Structure

The trick to include a modified noise model in the output error model structure is of course also applicable to the Box-Jenkins model structure. The alternative form will in this case be

$$y(t) = \frac{B(q)}{F(q)}u(t) + \frac{F_a^*(q)C(q)}{F_a(q)D(q)}e(t) \tag{45}$$

with the corresponding predictor

$$\begin{aligned}\hat{y}(t|\theta) &= \frac{F_a(q)D(q)B(q)}{F_a^*(q)C(q)F(q)}u(t) + (1 - \frac{F_a(q)D(q)}{F_a^*(q)C(q)})y(t) \\ &= \frac{D(q)B(q)}{C(q)F_a^*(q)F_s(q)}u(t) + (1 - \frac{D(q)F_a(q)}{C(q)F_a^*(q)})y(t)\end{aligned} \tag{46}$$

An explicit expression for the gradient filters for this predictor can be derived quite similarly as in the output error case, albeit that the formulas will be even messier. For the sake of readability we skip the details.

6 Conclusions

In this paper we have proposed new versions of the well known output error and Box-Jenkins model structures that can be used also for identification of unstable systems. The new model structures are equivalent to the standard ones, as far as number of parameters and asymptotical results are concerned, but guarantee stability of the predictors.

References

[1] U. Forssell. *Properties and Usage of Closed-loop Identification Methods*. Licentiate thesis LIU-TEK-LIC-1997:42, Department of Electrical Engineering, Linköping University, Linköping, Sweden, September 1997.

[2] U. Forssell and L. Ljung. Closed-loop identification revisited – Updated version. Technical Report LiTH-ISY-R-2021, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998.

[3] M. Gevers. Towards a joint design of identification and control. In H. L. Trentelman and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 111–151. Birkhäuser, 1993.

[4] F. R. Hansen, G. F. Franklin, and R. Kosut. Closed-loop identification via the fractional representation: experiment design. In *Proceedings of the American Control Conference*, pages 1422–1427, Pittsburg, PA, 1989.

[5] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[6] L. Ljung. Information contents in identification data from closed loop operation. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 2242–2247, San Antonio, TX, December 1993.

[7] L. Ljung. Identification in closed loop: some aspects on direct and indirect approaches. In *Preprints of the 11th IFAC Symposium on System Identification*, volume 1, pages 141–146, Fukuoka, Japan, July 1997.

[8] L. Ljung and U. Forssell. Variance results for closed-loop identification methods. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 3, pages 2435–2440, San Diego, C.A., 1997.

[9] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall International, 1989.

[10] P. M. J. Van den Hof and R. J. P. Schrama. An indirect method for transfer function estimation from closed loop data. *Automatica*, 29:1523–1527, 1993.

[11] P. M. J. Van den Hof and R. J. P. Schrama. Identification and control — Closed-loop issues. *Automatica*, 31:1751–1770, 1995.

[12] P. M. J. Van den Hof, R. J. P. Schrama, R. A. de Callafon, and O. H. Bosgra. Identification of normalized coprime factors from closed-loop experimental data. *European Journal of Control*, 1:62–74, 1995.

Paper E

Maximum Likelihood Estimation of Models with Unstable Dynamics and Nonminimum Phase Noise Zeros

Authors: Urban Forssell and Håkan Hjalmarsson

Edited version of paper to appear in *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999.

Preliminary version published as Technical Report LiTH-ISY-R-2043, Department of Electrical Engineering, Linköping University, Linköping, Sweden.

Maximum Likelihood Estimation of Models with Unstable Dynamics and Nonminimum Phase Noise Zeros

Urban Forssell

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: ufo@isy.liu.se.

Håkan Hjalmarsson

Signal Processing Group,
Department of Signals, Sensors and Systems,
Royal Institute of Technology,
S-100 44 Stockholm, Sweden.
E-mail: hakan.hjalmarsson@s3.kth.se.

Abstract

Maximum likelihood estimation of single-input/single-output linear time-invariant dynamic models requires that the model innovation (the nonmeasurable white noise source that is assumed to be the source of the randomness of the system) can be computed from the observed data. For many model structures, the prediction error and the model innovation coincide and the prediction error can be used in maximum likelihood estimation. However, when the model dynamics and the noise model have unstable poles which are not shared or when the noise dynamics have unstable zeros this is not the case. One such example is an unstable output error model. In this contribution we show that in this situation the model innovation can be computed by noncausal filtering. Different implementations of the model innovation filter are also studied.

Keywords: Identification, Output error identification, Prediction error methods, Maximum likelihood estimators, Innovation, Spectral factorization

1 Introduction

In the prediction error approach (e.g., [5]) to system identification, the model parameters are determined by minimizing the prediction error, that is, the error between the true output and the optimal (w.r.t. the model), linear predictor based on past inputs and outputs. In this paper we consider causal, linear, and time-invariant models of the form

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \tag{1}$$

where $\{u(t)\}$ is the input, $\{y(t)\}$ is the output, and $\{e(t)\}$ is assumed to be a zero mean white noise sequence which will be called the model innovation. Furthermore, θ is a parameter vector and q is the time shift operator in discrete time. Throughout this paper we will assume that the data really has been generated by a model in the model structure we are considering, that is, for the structure (1) there is a parameter θ_0 such that the data are described exactly by (1):

$$y(t) = G(q, \theta_0)u(t) + H(q, \theta_0)e(t) \tag{2}$$

We will denote the corresponding model, the true model and the corresponding sequence $\{e(t)\}$ the true model innovation.

If $H^{-1}(q, \theta)G(q, \theta)$ and $H^{-1}(q, \theta)$ are stable, the optimal one-step-ahead predictor for (1) is, [5]

$$\hat{y}(t|\theta) = H^{-1}(q, \theta)G(q, \theta)u(t) + (1 - H^{-1}(q, \theta))y(t) \tag{3}$$

The prediction error $\varepsilon(t, \theta) = y(t) - \hat{y}(t|\theta)$ is

$$\varepsilon(t, \theta) = H^{-1}(q, \theta)(y(t) - G(q, \theta)u(t)) \tag{4}$$

For the true model it then holds that the prediction error is equal to the true model innovation

$$\varepsilon(t, \theta_0) = e(t) \tag{5}$$

This property underlies the nice consistency properties of the prediction error method. It also implies that there is a close connection between maximum likelihood estimation and the prediction error method. If the probability distribution function (PDF) of $e(t)$ is known and the log-likelihood function of this PDF is used as criterion in the prediction error method, the maximum likelihood estimate of the parameters is obtained. (Further details will be given below.)

When either $H^{-1}(q, \theta)G(q, \theta)$ or $H^{-1}(q, \theta)$ is unstable, (3) is not the optimal predictor and, consequently, the property (5) does not hold. This was shown in [1] and in this paper we will further study this situation, when it occurs, and what it implies in terms of the properties of the estimates and the applicability of the prediction error/maximum likelihood method.

To illustrate the ideas we will for simplicity concentrate on the problem of identification of unstable output error models. Apart from discussing the theoretical issues alluded to above we will also study different implementations of the model innovation filter and present a simulation study. Here we compare three different methods that are applicable to output error identification of unstable systems. The first one was originally derived in [1] and relies on causal filtering operations only. This method circumvents the problem of unstable predictors by modifying the output error model structure. The other two method are novel and utilize noncausal filtering. Since the "predictors" in these methods are computed using noncausal filtering and are based on future signals, the prediction error $\varepsilon(t, \theta)$ should be interpreted not as the "prediction error" but rather as the model innovation. These methods therefore should be seen as maximum likelihood methods rather than prediction error methods.

We also mention that with minor changes the methods for parameter estimation presented in this paper can also be used in other applications, such as methods for tuning of controller parameters based on direct criterion minimization (e.g., [3]) when the controller is nonminimum phase and/or unstable.

2 Prediction Error Identification

Given a parameterized predictor model $\hat{y}(t|\theta)$, where θ is assumed to range over some subset $D_{\mathcal{M}}$ of $\mathbb{R}^d$ ($d = \dim \theta$), and measured data

$$Z^N = \{y(1), u(1), \ldots, y(N), u(N)\} \tag{6}$$

the prediction error estimate is found as, [5]

$$\hat{\theta}_N = \arg \min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N) \tag{7}$$

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \ell\left(\varepsilon_F(t, \theta)\right) \tag{8}$$

$$\varepsilon_F(t, \theta) = L(q, \theta)\varepsilon(t, \theta) \tag{9}$$

Here $\ell(\cdot)$ is a suitably chosen function. A standard choice is $\ell(x) = \frac{1}{2}x^2$ which gives the least-squares method. For maximum likelihood estimation $\ell(x) = -\log(f_e(x))$, where $f_e(\cdot)$ is the PDF of the true model innovation. $\varepsilon_F(t, \theta)$ is the filtered prediction error corresponding to the predictor (3). $L(q, \theta)$ is a stable (possibly parameter dependent) prefilter that, e.g., can be used to affect the frequency domain fit of the resulting estimate.

Typically one finds the estimate $\hat{\theta}_N$ using a gradient search algorithm. This is a crucial step in the identification but a detailed discussion of these issues goes beyond the scope of this paper. Instead we refer to [5] for further details.

Note that the model structure (cf. (1))

$$y(t) = G(q,\theta)u(t) + L^{-1}(q,\theta)H(q,\theta)e(t) \tag{10}$$

has the optimal predictor

$$\hat{y}(t|\theta) = L(q,\theta)H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - L(q,\theta)H^{-1}(q,\theta))y(t) \tag{11}$$

so the prediction error is

$$\varepsilon(t,\theta) = L(q,\theta)H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{12}$$

This is equivalent to $\varepsilon_F(t,\theta)$ defined in (9) when the prediction error $\varepsilon(t,\theta)$ is determined from (1). This shows that we may view $L(q,\theta)$ either as a prefilter or as a noise model depending on what is most natural in the application at hand.

In this paper we will mainly consider all-pass filters $L(q,\theta)$. If a quadratic criterion is used this implies that the resulting estimate asymptotically (as $N \to \infty$) will be the same as the one obtained with $L(q,\theta) = 1$. This follows since from standard prediction error theory we have that

$$\hat{\theta}_N \to \arg\min_{\theta \in D_{\mathcal{M}}} \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}\Phi_{\varepsilon_F}(\omega)d\omega \tag{13}$$

with probability 1 as $N \to \infty$. Here $\Phi_{\varepsilon_F}(\omega)$ is the spectrum of the filtered prediction error $\varepsilon_F(t,\theta) = L(q,\theta)\varepsilon(t,\theta)$ and, if $L(q,\theta)$ is all-pass, $\Phi_{\varepsilon_F}(\omega) = c \cdot \Phi_\varepsilon(\omega)$ for some positive constant c, which does not affect the frequency domain fit of the limit estimate.

3 The Problem of Unstable Predictors

We will now discuss how the prefilter $L(q,\theta)$ should be parameterized so that (11) is the optimal predictor for the model (1) in case (3) is unstable. The material in this section can be seen as a generalization of the results in [1].

A rather general parameterization of the model (1) is, [5]

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{14}$$

where $A(q)$, $B(q)$, $C(q)$, $D(q)$, and $F(q)$ are polynomials in q^{-1}. Without loss of generality $A(q)$, $C(q)$, $D(q)$, and $F(q)$ can be taken monic. Using the generic formula (3) we may write the "predictor" for this model as

$$\hat{y}(t|\theta) = \frac{D(q)}{C(q)}\frac{B(q)}{F(q)}u(t) + (1 - \frac{A(q)D(q)}{C(q)})y(t) \tag{15}$$

The corresponding "prediction error" is

$$\varepsilon(t,\theta) = \frac{D(q)}{C(q)}(A(q)y(t) - \frac{B(q)}{F(q)}u(t)) \tag{16}$$

For stability we have to require that the zeros of the $C(q)$- and $F(q)$-polynomials lie strictly *inside* the unit circle. In this case (15) is the optimal one-step-ahead predictor for (14) and (16) the corresponding prediction error.

In practical implementations of the prediction error method the zeros of the noise model (i.e., the zeros of $C(q)$) are usually constrained to be inside the unit circle. ($C(q)$ is then said to be minimum phase.) If the true model innovation is filtered through some filter with zeros outside the unit circle, the minimum phase equivalent of this filter will be identified and a consistent estimate of the noise spectrum obtained. It follows from the spectral factorization theorem that this is no limitation for identification based on second order statistics, that is, with quadratic criteria. This situation changes if other criteria than a quadratic are used, as typically is the case in maximum likelihood estimation. Restricting to minimum phase noise models in such cases would imply that consistency would be lost. For the moment, however, we will assume a quadratic criterion and a minimum phase $C(q)$-polynomial. It then remains to study the case when $F(q)$ has zeros outside the unit circle.

The solution, derived in [1], is to modify the predictor as follows. Let $F_n(q)$ denote the nonminimum phase part and $F_m(q)$ the minimum phase part of $F(q)$, $F(q) = F_n(q)F_m(q)$. Also let $F_n^*(q)$ denote the monic polynomial whose zeros are equal to the zeros of $F_n(q)$ reflected into the unit disc. Then the optimal predictor for (14) is (given that $C(q)$ is minimum phase)

$$\hat{y}(t|\theta) = \frac{D(q)}{C(q)}\frac{B(q)}{F_n^*(q)F_m(q)}u(t) + (1 - \frac{F_n(q)A(q)D(q)}{F_n^*(q)C(q)})y(t) \tag{17}$$

Clearly all filters in (17) are stable. To see that (17) is indeed optimal, notice that the prediction error can be written

$$\varepsilon(t,\theta) = \frac{F_n(q)}{F_n^*(q)}\frac{D(q)}{C(q)}(A(q)y(t) - \frac{B(q)}{F(q)}u(t)) \tag{18}$$

When evaluated for θ_0 we get

$$\varepsilon(t,\theta_0) = \frac{F_n(q)}{F_n^*(q)}e(t) \tag{19}$$

and since $F_a(q)/F_a^*(q)$ is an all-pass filter this means that the prediction error is white (for $\theta = \theta_0$) and hence uncorrelated with past inputs and outputs. The orthogonality principle now proves that (17) is optimal.

As can be seen, the key feature of this solution is that the prediction error is a second order equivalent to the true model innovation (filtered through an all-pass filter). This is also evident if we compare (18), (16), and (12) since this will show that (18) results if the model (14) is applied with the all-pass prefilter

$$L(q,\theta) = \frac{F_n(q)}{F_n^*(q)} \tag{20}$$

At this point we should also remark that a similar solution is possible when models with nonminimum phase zeros are encountered. The predictor would then use a factorization of $C(q)$ into one minimum phase and one nonminimum phase part. (Note, though, that this is not necessary or even desirable if a quadratic criterion is used.)

In the sequel we will for simplicity mainly discuss identification of unstable output error models. This corresponds to the case when $A(q) = C(q) = D(q) = 1$. In this situation there is a nice and simple interpretation of (17) in terms of the Kalman filter (cf. [1]): Realize the output error model

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t) \tag{21}$$

in state-space form:

$$\begin{aligned} x(t+1) &= \mathcal{A}x(t) + \mathcal{B}u(t) \\ y(t) &= \mathcal{C}x(t) + e(t) \end{aligned} \tag{22}$$

The steady state Kalman filter predictor is

$$\begin{aligned} \hat{x}(t+1) &= (\mathcal{A} - \mathcal{K}\mathcal{C})\hat{x}(t) + \mathcal{B}u(t) + \mathcal{K}y(t) \\ \hat{y}(t|\theta) &= \mathcal{C}\hat{x}(t) \end{aligned} \tag{23}$$

where $\mathcal{K}$ is determined from an algebraic Riccati equation in the usual way:

$$\mathcal{K} = \frac{\mathcal{A}\Pi\mathcal{C}^T}{(R + \mathcal{C}\Pi\mathcal{C}^T)} \tag{24}$$

$$\Pi = \mathcal{A}\Pi\mathcal{A}^T - \mathcal{K}(R + \mathcal{C}\Pi\mathcal{C}^T)\mathcal{K}^T \tag{25}$$

Here R is the variance of the model innovation $e(t)$. Now, if $\mathcal{A}$ is stable the solution is $\mathcal{K} = 0$ and (23) corresponds to the standard output error predictor (cf. (15))

$$\hat{y}(t|\theta) = \frac{B(q)}{F(q)}u(t) \tag{26}$$

However if $\mathcal{A}$ is not stable, the solution to the Kalman predictor problem is a nonzero $\mathcal{K}$ which makes (23) exactly equal to (cf. (17))

$$\hat{y}(t|\theta) = \frac{B(q)}{F_n^*(q)F_m(q)}u(t) + (1 - \frac{F_n(q)}{F_n^*(q)})y(t) \tag{27}$$

It can thus be seen that the optimal steady state Kalman filter performs a *canonical* factorization of the prediction error spectrum into one stable and one anti-stable part. This also follows from the general Kalman/Wiener filtering theory (e.g., [4]) and applies as long as the (spectrum of) the prediction error does not have zeros on the unit circle. When there are zeros on the unit circle, there exists no asymptotically stable stationary optimal predictor and a time-varying Kalman filter has to be used. We will now study methods employing noncausal filtering to handle the unstable predictors. These methods can be interpreted as methods that gives different factorizations of the prediction error spectrum than the canonical one.

4 Maximum Likelihood Estimation

Suppose that the PDF $f_e(\cdot)$ of $e(t)$ in the model (1) is known and that we would like to use maximum likelihood estimation. We then need a way to compute the model innovation. In some cases the prediction error coincides with the model innovation, as we saw in the previous section. However, if the noise model has nonminimum phase zeros or if the model dynamics has unstable poles that are not shared with the noise model the prediction error is an all-pass filtered version of the model innovation. Hence, the prediction error cannot be used for maximum likelihood estimation (unless the PDF is Gaussian since then the maximum likelihood criterion is quadratic).

In this section we will show that by instead using noncausal filters, the model innovation $e(t)$ can indeed be obtained. Since noncausal filtering is required this method is limited to off-line identification. For simplicity we will only discuss the case of maximum likelihood estimation of unstable output error models.

Consider the model (21) and assume that $F(q)$ has some zero(s) outside the unit circle. Furthermore, suppose that through some feedback mechanism (not necessarily linear) $\{y(t)\}$ and $\{u(t)\}$ are stationary and bounded even though the open loop system (21) is unstable. From (21) it follows that the signals $y(t)$, $u(t)$, and $e(t)$ in the system satisfy

$$F(q)y(t) - B(q)u(t) = F(q)e(t) \tag{28}$$

If $F(q)$ is nonminimum phase it is not possible to obtain $e(t)$ from $F(q)y(t) - B(q)u(t)$ by stable, causal filtering. However, no constraint is imposed, unless

$F(q)$ has some zero(s) on the unit circle, if noncausal filtering is applied! Hence by applying the noncausal filter $1/F(q)$ to $F(q)y(t) - B(q)u(t)$, $e(t)$ is obtained modulo transient effects. This gives

$$e(t) = \frac{1}{F(q)}\left(F(q)y(t) - B(q)u(t)\right) \tag{29}$$

$$= y(t) - \frac{B(q)}{F(q)}u(t) \tag{30}$$

where $1/F(q)$ is a noncausal filter. This suggests a way to implement maximum likelihood estimation of unstable output error models using the method (7)-(9) (with $L(q,\theta) = 1$): Use noncausal filtering!

Consider the predictor (26):

$$\hat{y}(t|\theta) = \frac{B(q)}{F(q)}u(t)$$

This can also be written

$$\hat{y}(t|\theta) = \frac{1}{F_n(q)}\frac{B(q)}{F_m(q)}u(t) = \frac{B(q)}{F_m(q)}\frac{1}{F_n(q)}u(t) \tag{31}$$

using the factorization $F(q) = F_n(q)F_m(q)$. Thus we can generate the signal $\hat{y}(t|\theta)$ by first computing

$$\hat{y}_{causal}(t|\theta) = \frac{B(q)}{F_m(q)}u(t) \tag{32}$$

followed by backwards filtering of the reversed sequence $\{\hat{y}^R_{causal}(t|\theta)\}_{t=1}^N$ through the filter $\frac{1}{F_n(q)}$ giving

$$\hat{y}^R(t|\theta) = \frac{1}{F_n(q)}\hat{y}^R_{causal}(t|\theta) \tag{33}$$

The desired signal $\hat{y}(t|\theta)$ is then finally obtained by reversing the sequence $\{\hat{y}^R(t|\theta)\}_{t=1}^N$. It is also clear from (31) that the filtering operations can change places so that the noncausal filter is applied before the causal. Algorithmically these two versions are equivalent; in practice the results will differ due to unknown initial conditions in both the causal and noncausal filtering. This will be further discussed below. Here we will assume that the causal filtering is performed first followed by the noncausal filtering.

An alternative to using (31) is to rewrite the predictor as

$$\hat{y}(t|\theta) = \left(\frac{P(q)}{F_n(q)} + \frac{Q(q)}{F_m(q)}\right)u(t) \tag{34}$$

that is, to rewrite the predictor filter using a partial fraction expansion in one stable part and one anti-stable part. The anti-stable filter should of course be applied using backwards filtering, just as in the previous method. The polynomials $P(q)$ and $Q(q)$ in (34) are found by solving the Diophantine equation

$$P(q)F_m(q) + Q(q)F_n(q) = B(q) \tag{35}$$

This variant requires more computations compared to factoring the denominator into stable and anti-stable factors, but on the other hand both filters should be applied directly to the input signal and not to some intermediate signal like in the previous case. Also in this case will transients due to unknown initial conditions in both the causal and the anti-causal filter deteriorate the results, unless some measures are taken to circumvent this.

We remind that since (31) and (34) utilize noncausal filtering the error signal $y(t) - \hat{y}(t|\theta)$ (neglecting transient effects) should be regarded as the model innovation, not the prediction error. The two methods (31) or (34) (together with (7)-(9)) are therefore maximum likelihood methods rather than prediction error methods.

Forward/backward filtering is common in off-line signal processing applications, like zero-phase filtering using IIR filters (cf. the function `filtfilt` in Matlab's Signal Processing Toolbox, [7]) and implementation of noncausal Wiener filters (e.g., [4]), but has not been widely used in identification applications even though these typically are also off-line applications.

As we have mentioned, the effects of unknown initial states in the different filters will deteriorate the results of the identification and it is desirable to have some means for circumventing this. In [2] methods for eliminating the transients in forward/backward filtering are derived. The idea is to match the output of a forward/backward implementation with that of a backward/forward implementation of the same filter. This leads to good compromise solutions in many signal processing problems. Unfortunately these methods are not well suited for the identification problem. The reason is mainly that in prediction error identification it is more important that the criterion function is kept small than that all predicted outputs are kept small. However, with some changes the ideas in [2] could perhaps be useful in this case too. This is a topic for future research.

5 Simulation Example

To illustrate the applicability of the proposed methods for output error identification of unstable systems we will in this section present a small simulation study. We will refer to the prediction error method using (27) as Method 1 and the two maximum likelihood methods which employ noncausal filtering as in (31) and (34) will be called Method 2 and 3, respectively. The estimates are found using (7)-(9) where $\ell(x) = \frac{1}{2}x^2$ and $L(q, \theta) = 1$.

In the Matlab implementations of the methods we have tried to reduce the negative effects of transients by matching the initial conditions in the predictor filters to the measured outputs, as far as possible. For Method 1 this is done by ensuring that the first few samples of the predicted output is equal the the corresponding samples of the measured output. Methods 2 and 3 involve both forwards and backwards filtering and transient effects will be present both at the beginning and at the end of the resulting sequence of predicted outputs. It is relatively straightforward to remove the transients at one end, the beginning say, but it is not clear how to achieve transient free results both at the beginning and the end. The implementation we have used removes the transients at the end for Method 2 and at the beginning for Method 3.

The true model – to be identified – is given by

$$y(t) = \frac{b_0}{1 + a_1 q^{-1} + a_2 q^{-2}} u(t) + e(t) \tag{36}$$

with $b_0 = 1$, $a_1 = -2.5$, and $a_2 = 1$. This system is unstable with poles in 2 and 0.5. To generate identification data we simulated this system using the feedback law

$$u(t) = r(t) - 1.5y(t-1) + 0.75y(t-2) \tag{37}$$

which places both closed-loop poles in 0.5. In the Monte Carlo simulation we used independent, zero mean, Gaussian white noise reference and noise signals $\{r(t)\}$ and $\{e(t)\}$ with variances 1 and 0.01, respectively. The Monte Carlo simulation comprised 100 different runs, in each run $N = 200$ data samples were used. The search was initialized at the point $\hat{\theta}^{(0)} = \begin{bmatrix} -2.48 \ 0.98 \ 1.02 \end{bmatrix}^T$ in Methods 1,2, and 3, for fair comparisons between the methods. Apart from the methods presented above, we also tried the standard prediction error method with an output error model structure and a second-order ARMAX model structure. Here we used the standard routines for output error and ARMAX identification available in Matlab's System Identification Toolbox, [6]. It should be noted that all these methods, except the standard output error method, should under ideal conditions give consistent estimates of the true system, despite the feedback.

In Table E.1 we have summarized the results of the identification, the numbers shown are the estimated parameter values together with their standard deviations.

From Table E.1 we see that the standard output error method gives completely useless results, as could be expected since the system is unstable. Methods 1, 2, and 3 all give estimates of roughly the same quality. The ARMAX method gives the best results in this simulation although the differences are not that big. A

Table E.1 *Summary of identification results.*

Parameter	a_1	a_2	b_0
True value	-2.5	1	1
OE	-0.3499	0.0490	-0.3123
	$\pm$0.0344	$\pm$0.0357	$\pm$0.0091
Method 1	-2.4848	0.9831	1.0034
	$\pm$0.0049	$\pm$0.0067	$\pm$0.0091
Method 2	-2.5130	1.0043	1.0133
	$\pm$0.0134	$\pm$0.0089	$\pm$0.0130
Method 3	2.5141	1.0074	1.0103
	$\pm$0.0088	$\pm$0.0134	$\pm$0.0131
ARMAX	-2.4974	0.9983	0.9989
	$\pm$0.0132	$\pm$0.0087	$\pm$0.0124

possible explanation is that transient effects is better taken care of in the ARMAX function than our implementations of Methods 1, 2, and 3.

To further illustrate the performance of these methods we calculated the average number of flops required for each method. The numbers are shown in Table E.2. In this example Method 2 required the least number of flops. Methods 1 and

Table E.2 *Average Number of Flops.*

Method	Flops
OE	$1.6537 \cdot 10^5$
Method 1	$1.7861 \cdot 10^5$
Method 2	$1.2161 \cdot 10^5$
Method 3	$1.4758 \cdot 10^5$
ARMAX	$4.1605 \cdot 10^5$

3 were slightly worse. Note especially that Method 1 and the standard output error method required roughly the same number of flops, despite the more involved gradient calculations in Method 1. The relatively small difference between this methods can partly be ascribed to be due to convergence problems for the standard output error method since the true system is unstable. Estimating the ARMAX model required significantly more computations than the other methods. Partly this difference is due to the more advanced initialization procedure used in the routine, but this should not make that big a difference as we obtained.

6 Conclusions

We have discussed estimation of models where the dynamic model is unstable and/or the noise model has nonminimum phase zeros. In such situations the predictor has to be modified and the prediction error does no longer coincide with the model innovation. The main contribution was to show that by instead using noncausal stationary filters, the model innovation can be computed if the transfer functions from the observed data to the model innovation does not have any poles on the unit circle. The advantage of using these noncausal filters is that since the model innovation is obtained, maximum likelihood identification can be performed. Hence, maximum likelihood identification of systems with unstable dynamics or nonminimum phase noise zeros is possible.

References

[1] U. Forssell and L. Ljung. Identification of unstable systems using output error and Box-Jenkins model structures. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3932–3937, Tampa, FL, 1998. To appear in IEEE Transactions on Automatic Control.

[2] F. Gustafsson. Determining the initial states in forward-backward filtering. *IEEE Transactions on Signal Processing*, 46:988–992, 1996.

[3] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin. Iterative feedback tuning: Theory and applications. *IEEE Control Systems*, 18:26–41, 1998.

[4] T. Kailath. *Lectures on Wiener and Kalman Filtering*. Springer-Verlag, 1981.

[5] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[6] L. Ljung. *System Identification Toolbox – User's Guide*. The MathWorks, Inc., Cochituate Place, Natick, MA. USA, 1995.

[7] The MathWorks, Inc., Cochituate Place, Natick, MA. USA. *Signal Processing Toolbox – User's Guide*, 1993.

Paper F

Time-domain Identification of Dynamic Errors-in-variables Systems Using Periodic Excitation Signals

Authors: Urban Forssell, Fredrik Gustafsson, and Tomas McKelvey

Edited version of paper to appear in *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999.

Preliminary version published as Technical Report LiTH-ISY-R-2044, Department of Electrical Engineering, Linköping University, Linköping, Sweden.

Time-domain Identification of Dynamic Errors-in-variables Systems Using Periodic Excitation Signals

Urban Forssell, Fredrik Gustafsson, and Tomas McKelvey

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: {ufo,fredrik,tomas}@isy.liu.se.

Abstract

The use of periodic excitation signals in identification experiments is advocated. With periodic excitation it is possible to separate the driving signals and the disturbances, which for instance implies that the noise properties can be independently estimated. In the paper a nonparametric noise model, estimated directly from the measured data, is used in a compensation strategy applicable to both least-squares and total least-squares estimation. The resulting least-squares and total least-squares methods are applicable in the errors-in-variables situation and give consistent estimates regardless of the noise. The feasibility of the idea is illustrated in a simulation study.

Keywords: Identification, Identification algorithms, Least-squares estimation, Excitation, Dynamic systems

1 Introduction

One of the most important steps in the identification process is the experiment design. This involves, for example, deciding what signals to measure, choosing the sampling interval, and designing the excitation signals. In this paper we will advocate the use of periodic excitation.

Periodic excitation has up to this point mostly been used in frequency domain identification (e.g, [6, 5]), but offer several interesting and useful advantages compared to nonperiodic (random) excitation also in time domain identification. The main advantages with periodic excitation in time domain identification are:

- Data reduction. By averaging over M periods the amount of data is reduced M times.

- Improved signal-to-noise ratio. By averaging over M periods the noise variance is lowered by a factor M. This will have important consequences for both the numerical properties of the estimation algorithms and the statistical properties of the estimates.

- Separation of driving signals and noise. With periodic excitation all nonperiodic signal variations over the periods will be due to random disturbances, noise. This means, for instance, that we can estimate the noise level and thus compute a priori bounds for the least-squares cost function used in the identification.

- Independent estimation of nonparametric noise models. Since we can separate the signals from the noise it is possible to independently estimate the noise properties. Such a noise model can be used as a prewhitening filter applied before the estimation or as a tool for model validation.

In this paper we will study how to identify dynamic errors-in-variables systems using time-domain data. This is a problem that has received considerable interest in the literature, see, e.g., [9, 1, 12] and the more recent [13, 2]. With periodic excitation a number of possibilities opens up for constructing simple, efficient methods that solves this problem. We will study some of them in this contribution. In particular, compensation methods for least-squares and total least-squares estimation that can handle also the errors-in-variables problem will be presented. The idea used is similar to the bias-correction technique studied in for instance [10, 12, 13]. Compared to the methods studied in these references, the proposed methods have the advantage of giving consistent estimates regardless of the properties of the noise.

2 Problem Formulation

Consider a linear, time-invariant, discrete-time system

$$y(t) = G(q)u(t) = \sum_{k=0}^{\infty} g_k q^{-k} u(t) \tag{1}$$

where $u(t) \in \mathbb{R}$ is the input, $y(t) \in \mathbb{R}$ is the output, and q^{-1} is the delay operator ($q^{-k}u(t) = u(t-k)$). We will assume that the order of the system is finite, so that the system can be represented as

$$y(t) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) + b_0 u(t-n_k) + b_{n_b} u(t-n_k-n_b) \tag{2}$$

Here we have explicitly included the possibility of a delay n_k. The transfer operator $G(q)$ in this case becomes

$$G(q) = \frac{q^{-n_k} B(q)}{A(q)} \tag{3}$$

$$B(q) = b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \tag{4}$$

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \tag{5}$$

The problem we consider is how to identify $G(q)$ using noisy measurements of $y(t)$ and $u(t)$. Our measured data can thus be described by

$$Z_m^N = \{z_m(1), \ldots, z_m(N)\} \tag{6}$$

$$z_m(t) = z(t) + w(t) \tag{7}$$

$$z(t) = \begin{bmatrix} y(t) & u(t) \end{bmatrix}^T, \ w(t) = \begin{bmatrix} w_y(t) & w_u(t) \end{bmatrix}^T \tag{8}$$

We will also use the notation $y_m(t) = y(t) + w_y(t)$ and $u_m(t) = u(t) + w_u(t)$. The unknown signals $w_y(t)$ and $w_u(t)$ act as noise sources on the measured output and input, respectively. We will make the following assumptions about the signals $z(t)$ and $w(t)$:

A1 $u(t)$ is periodic with period P, $P \geq 2\bar{n} + 1$ where $\bar{n}$ is an a priori given upper bound on the system order.

A2 $u(t)$ is persistently exciting of order $\bar{n}$.

A3 $z(t)$ and $w(t)$ are jointly quasi-stationary.

A4 $w(t)$ has the property

$$\lim_{M\to\infty} \frac{1}{M} \sum_{k=0}^{M-1} w(t+kP) = 0 \quad \forall t \tag{9}$$

A5 $z(t)$ and $w(t)$ are uncorrelated.

Assumption A2 is required in order to uniquely identify the system. Assumptions A1 and A4 enable us to use simple averaging to remove the noise. In a stochastic setting we assume A4 to hold with probability 1. Assumption A5 implies that sample means of products of $z(t)$ and $w(t-k)$ tend to zero as the number of samples tends to infinity. In addition to the assumptions listed above, it is also assumed that an integer number of periods has been measured, that is $N = M \cdot P,\ M \geq 1$.

3 Averaging

An important step in the identification is to average the measured data. Define the averaged input and output as

$$\bar{u}(t) = \frac{1}{M} \sum_{k=0}^{M-1} u_m(t + kP),\ t \in [1, P] \tag{10}$$

$$\bar{y}(t) = \frac{1}{M} \sum_{k=0}^{M-1} y_m(t + kP),\ t \in [1, P] \tag{11}$$

From assumption A4 it follows that $\bar{u}(t) \to u(t),\ t \in [1, P]$ and $\bar{y}(t) \to y(t),\ t \in [1, P]$ as M tends to infinity. $\bar{u}(t)$ and $\bar{y}(t)$ are thus consistent estimates of noise free signals $u(t)$ and $y(t)$, respectively. In [4] this is used to derive simple, consistent methods for the identification of errors-in-variables systems. The idea in [4] was that as M tends to infinity, the noise will average out and we are effectively identifying a noise-free system. In this paper we will not generally assume that the number of periods tends to infinity, which makes the problem significantly harder.

4 Estimating the Noise Statistics

Let $\bar{z}(t) = \begin{bmatrix} \bar{y}(t) & \bar{u}(t) \end{bmatrix}$. By periodically continuing $\bar{z}(t)$ outside $t = [1, P]$ we can estimate the noise $w(t)$ as

$$\hat{w}(t) = z_m(t) - \bar{z}(t),\ t \in [1, N] \tag{12}$$

A consistent estimate of the covariance function

$$R_{ww}(k) = Ew(t)w^T(t + k) \tag{13}$$

can now be computed as

$$\hat{R}_{ww}(k) = \frac{1}{(M-1)P} \sum_{t=1}^{MP} \hat{w}(t)\hat{w}^T(t+k) \tag{14}$$

where the convention is that all signals outside the interval $t = [1, MP]$ are replaced by 0. In practice for large data sets, the covariance function should be computed using FFT, see [8]. It is important to note that we have used P degrees of freedom for estimating the mean, so the proper normalization to get an unbiased estimate is $MP - P = (M-1)P$. How many periods do we need then? The rather precise answer provided in [7] is $M \geq 4$. The asymptotic properties $N = MP \to \infty$ of the estimate are then independent of how the excitation is divided into M and P.

An unbiased estimate of the spectrum of $w(t)$ is obtained by the periodogram

$$\hat{\Phi}_w(\omega) = \sum_{k=-MP+1}^{MP-1} \hat{R}_{ww}(k) e^{-i\omega k} \tag{15}$$

This can be used for prewhitening of $w(t)$ prior to the estimation. It turns out that the poor variance properties of (15) does not diminish its usefulness for prewhitening. An example of this will be shown in Section 11. We also mention that $\hat{\Phi}_w(\omega)$ can be estimated very efficiently using FFT directly from the original data.

5 Least-squares Estimation Using Periodic Data

Consider the linear regression model

$$\hat{y}_m(t,\theta) = \varphi^T(t)\theta \tag{16}$$

$$\varphi(t) = \left[-y_m(t-1), \ldots, -y_m(t-n_a), u_m(t-n_k), \ldots, u_m(t-n_k-n_b)\right]^T \tag{17}$$

$$\theta = \left[a_1, \ldots, a_{n_a}, b_0, \ldots, b_{n_b}\right]^T \tag{18}$$

The least-squares (LS) estimate of θ using N data samples can be written

$$\hat{\theta}_N = R_N^{-1} f_N \tag{19}$$

$$R_N = \frac{1}{N}\sum_{t=1}^{N} \varphi(t)\varphi^T(t), \;\; f_N = \frac{1}{N}\sum_{t=1}^{N} \varphi(t)y(t) \tag{20}$$

Introduce the notation

$$\varphi_z(t) = \left[-y(t-1), \ldots, -y(t-n_a), u(t-n_k), \ldots, u(t-n_k-n_b)\right]^T \tag{21}$$

$$\varphi_w(t) = \left[-w_y(t-1), \ldots, -w_y(t-n_a), w_u(t-n_k), \ldots, w_u(t-n_k-n_b)\right]^T \tag{22}$$

Since $z(t)$ and $w(t)$ are uncorrelated we have that

$$\lim_{N\to\infty} R_N = R = R_z + R_w \tag{23}$$

$$\lim_{N\to\infty} f_N = f = f_z + f_w \tag{24}$$

$$R_z = E\varphi_z(t)\varphi_z^T(t), \quad R_w = E\varphi_w(t)\varphi_w^T(t) \tag{25}$$

$$f_z = E\varphi_z(t)y(t), \quad f_w = E\varphi_w(t)w_y(t) \tag{26}$$

If, indeed

$$y_m(t) = \varphi^T(t)\theta + e(t) \tag{27}$$

where $e(t)$ is white noise with variance λ_0, then the least-squares estimate is consistent with asymptotic covariance matrix

$$N \operatorname{Cov}\hat{\theta} \approx \lambda_0 R^{-1} \tag{28}$$

However, with colored noise and/or noisy measurements of the input this is no longer true and the least-squares estimate will be biased.

Let $\bar{R}_P$ and $\bar{f}_P$ be defined similar to R_P and f_P, respectively, except that averaged data is used. We then have that

$$\lim_{P\to\infty} \bar{R}_P = \bar{R} = R_z + \frac{1}{M}R_w \tag{29}$$

$$\lim_{P\to\infty} \bar{f}_P = \bar{f} = f_z + \frac{1}{M}f_w \tag{30}$$

The M normalization is due to the averaging which decreases the noise variance with a factor of M. The least-squares estimate using averaged data will still be unbiased if the true system is given by (27), but the asymptotic covariance matrix changes to

$$MP \operatorname{Cov}\hat{\theta}_P \approx M\frac{\lambda_0}{M}\bar{R}^{-1} = \lambda_0\bar{R}^{-1} \tag{31}$$

The scaling factor is thus the same, but R is replaced by $\bar{R}$, and since $R \geq \bar{R}$ this means that the asymptotic covariance increases with averaged data.

6 Improving the Accuracy

If we have periodic excitation and if (27) holds then we can recover the original information in R_w and f_w using the nonparametric noise model (14). The idea is to construct nonparametric estimates $\hat{R}_w^{np}$ and $\hat{f}_w^{np}$ of R_w and f_w, respectively, from $\hat{R}_{ww}(k),\ k = 0, 1, \ldots$ and compensate for the missing terms in $\bar{R}$ and $\bar{f}$. As pointed out before, these estimates use $(M-1)P$ degrees of freedom. Note also that $\bar{R}_P$ and $\bar{f}_P$ already contain estimates of R_w and f_w, respectively. These have P degrees of freedom (averages over P samples), and are functions of the sample mean $\bar{w}(t)$. This is important since the nonparametric estimates are based on the second-order properties of $w(t)$, and thus these two estimates of R_w and f_w are uncorrelated, and even independent if Gaussian noise is assumed. This implies that we can compensate the least-squares quantities obtained from averaged data

$$R_P^c = \bar{R}_P + \hat{R}_w^{np}, \quad f_P^c = \bar{f}_P + \hat{f}_w^{np} \tag{32}$$

and recover all $MP = (M-1)P + P$ degrees of freedom. This is further discussed in [3].

7 Consistent Least-squares Estimation of Errors-in-variables Systems

A similar idea can be used to remove the bias in the least-squares estimate due to (colored) noise $w(t)$ acting on the input and the output: we simply have to subtract away the terms in $\bar{R}_P$ and $\bar{f}_P$ that are due to the noise $w(t)$ using the nonparametric estimates $\hat{R}_w^{np}$ and $\hat{f}_w^{np}$. By equating the degrees of freedom it can be shown that

$$\hat{R}_z = \bar{R}_P - \frac{1}{M-1}\hat{R}_w^{np}, \quad \hat{f}_z = \bar{f}_P - \frac{1}{M-1}\hat{f}_w^{np} \tag{33}$$

are consistent estimates of R_z and f_z, respectively. We have thus removed all effect of the noise $w(t)$ in $\bar{R}_P$ and $\bar{f}_P$ by a simple subtraction operation and the resulting least squares estimate

$$\hat{\theta}_P = \hat{R}_z^{-1}\hat{f}_z \tag{34}$$

will be consistent regardless of $w(t)$. The method (33)-(34) will be referred to as the compensated least-squares (CLS) method. Due to its simplicity and general applicability, this method is a very interesting alternative to other methods that are applicable in the errors-in-variables situation. Note that with the CLS method no iterations are required to find the estimate – a clear advantage compared to most other errors-in-variables methods which frequently use singular value decompositions (SVDs) and to most other time-domain identification schemes which often use Gauss-Newton type search algorithms for finding the estimates.

8 The Total Least-squares Solution

For simplicity, assume that $n_a = n_b = n$ and $n_k = 0$. The relation (2), $t \in [1, N]$, can in this case be restated as

$$T^N \theta' = 0 \tag{35}$$

where

$$T^N = \begin{bmatrix} -Y^N & U^N \end{bmatrix} \tag{36}$$

$$Y^N = \begin{bmatrix} y(n+1) & \dots & y(1) \\ \vdots & \ddots & \vdots \\ y(N) & \dots & y(N-n) \end{bmatrix} \tag{37}$$

$$U^N = \begin{bmatrix} u(n+1) & \dots & u(1) \\ \vdots & \ddots & \vdots \\ u(N) & \dots & u(N-n) \end{bmatrix} \tag{38}$$

$$\theta' = \begin{bmatrix} 1, a_1, \dots, a_n, b_0, \dots, b_n \end{bmatrix}^T \tag{39}$$

The nontrivial right null space of the data matrix T^N describes the system.

With noisy measurements $z_m(t)$ of $z(t)$, a total least-squares (TLS) solution is natural to apply. Denote the noisy variant of T^N by T_m^N. The total least-squares solution $\hat{\theta}'_{TLS}$ can be stated as

$$T_{TLS}^N = \arg\min_T \|T_m^N - T\|_F^2 \tag{40}$$

subject to

$$T_{TLS}^N \hat{\theta}'_{TLS} = 0 \tag{41}$$

The solution is easily calculated by a singular value decomposition of the data matrix T_m^N. Introduce the error

$$W^N = T_m^N - T^N \tag{42}$$

With periodic data, averaged over M periods ($N = MP$), we have that

$$\|W^P\|_F^2 \to 0, \quad \text{as } M \to \infty \tag{43}$$

using Assumption A4. Under these conditions the TLS estimate is a consistent estimate of the system.

Let R'_w be the covariance matrix of

$$\varphi'_w(t) = \begin{bmatrix} -w_y(t), \dots, -w_y(t-n), w_u(t), \dots, w_u(t-n) \end{bmatrix}^T \tag{44}$$

(cf. (22)). To improve the efficiency of the total least squares estimator one can use R'_w which lead to the generalized total least-squares (GTLS) solution [11]. The GTLS solution $\hat{\theta}'_{GTLS}$ is

$$T^N_{GTLS} = \arg\min_T \|(T^N_m - T)(R'_w)^{-1/2}\|^2_F \tag{45}$$

subject to

$$T^N_{GTLS}\hat{\theta}'_{GTLS} = 0 \tag{46}$$

To understand the effect of the scaling $(R'_w)^{-1/2}$ it is instructive to study the product of T^N_m and its transpose. Introduce the notation

$$R'_N = \frac{1}{N}(T^N_m)^T T^N_m \tag{47}$$

$$R'_z = E\varphi'_z(t)(\varphi'_z(t))^T \tag{48}$$

$$\varphi'_z(t) = \big[-y(t), \dots, -y(t-n), u(t), \dots, u(t-n)\big]^T \tag{49}$$

The solution to the TLS (GTLS) problem is given by the right null space of T^N_m or alternatively by the null space of R'_N. Using Assumption A5 we see that

$$R'_N \to R'_z + R'_w, \quad \text{as } N \to \infty \tag{50}$$

If we include the scaling $(R'_w)^{-1/2}$ the covariance matrix R'_w is replaced by the identity matrix, which does not affect the directions of the singular vectors of R'_z. This means that the GTLS solution can be computed by finding the singular vector corresponding to the smallest singular value of the matrix $(R'_w)^{-T/2}R'_N(R'_w)^{-1/2}$. If the true R'_w is known, or if a consistent estimate of it can be computed, the GTLS estimator is consistent even if the variance of the noise $w(t)$ does not tend to zero.

The point is now that with periodic excitation we can obtain a consistent estimate of R'_w very easily using the nonparametric noise model (14). In the rest of the paper we shall refer to this variant of the general algorithm as the GTLS algorithm.

9 A Compensation Method for Total Least-squares Estimation

Let $\bar{R}'_P$ be defined as R'_P except that periodic data is used. With periodic data, averaged over M periods, we have that

$$\bar{R}'_P \to R'_z + \frac{1}{M}R'_w, \quad \text{as } N \to \infty \tag{51}$$

Let $\hat{R}'^{np}_w$ be a nonparametric estimate of R'_w obtained using (14). Similar arguments as in Section 7 will show that

$$\bar{R}'_P - \frac{1}{M-1}\hat{R}'^{np}_w \tag{52}$$

is a consistent estimate of R'_z. This holds regardless of the noise, which implies that the total least-squares estimator with the compensation (52) gives consistent estimates regardless of the noise even though the number of periods, M, does not tend to infinity. This method will be referred to as the compensated total least-squares (CTLS) estimator.

10 Prewhitening of the Noise

As mentioned in Section 4 the spectrum of the noise signal $w(t)$ can be estimated very efficiently using FFT when periodic data is used. Similarly we can prefilter the data very easily in the frequency domain simply by multiplying the Fourier transformed data sequences and the inverse of a square-root factor of the estimated noise spectrum. The corresponding time-domain signals are then obtained through IFFT. To preserve the relation between $u(t)$ and $y(t)$ it is important that $u_m(t)$ and $y_m(t)$ are prefiltered using the same filter. We thus have the two choices: either we compute the prefilter that will whiten the noise in $y_m(t)$ ($w_y(t)$), or compute the prefilter that will whiten the noise in $u_m(t)$ ($w_u(t)$). In many cases it is most natural to whiten the noise on the output, but in other cases the choice is more arbitrary. The former would for instance be the case if we know that the measurement noise on $y(t)$ and $u(t)$ is negligible, so that $w(t)$ basically is due to process noise acting on the output, or if the system operates in closed-loop and the measurement noise is negligible, which typically leads to similar spectra of $w_u(t)$ and $w_y(t)$. If the choice is less obvious, one can benefit from whitening the noise which has the highest variance. This will of course distort the spectrum of the other noise signal, but since the variance is smaller the net effect will be positive.

Prewhitening of the noise can also be used to derive simplified estimation algorithms. Consider for instance the least-squares estimator (19)-(20). If $w_y(t)$ is white noise and if $w_y(t)$ is uncorrelated with $w_u(t-n_k-k)$, $k \geq 0$, then f_w defined in (26) will be zero. This means that the CLS algorithm (33)-(34) can be simplified since the second compensation in (33) may be skipped.

11 Example

Consider the system

$$\begin{aligned} y(t) &= 1.5y(t-1) - 0.7y(t-2) \\ &\quad + u(t-1) + 0.5u(t-2) + v(t) \qquad (53) \\ v(t) &= -0.9v(t-1) + e(t) \qquad (54) \end{aligned}$$

where $e(t)$ is white Gaussian noise with variance σ_e^2. Apart from the noise $v(t)$, (53) is of the form (2) with $a_1 = -1.5$, $a_2 = 0.7$, $b_0 = 1$, $b_1 = 0.5$, $n_a = 2$, $n_b = 2$, and $n_k = 1$. This system was simulated using the control law

$$u(t) = r(t) - 0.25y(t) \qquad (55)$$

where $r(t)$ is a periodic reference signal with period P. In the simulations $r(t)$ was taken as a unit binary random signal. We also added colored measurement noise on both $y(t)$ and $u(t)$. These noise sources were independent but with equal spectra. The measurement noises were realized as Gaussian, white noise sequences filtered through a second-order, high-pass Butterworth filter with cut-off frequency 0.3. The variance of the white noise was σ_n^2.

A number of identification approaches were considered:

1. **LS, LS-A, LS-AF** least-squares estimation using raw data, averaged data, and averaged and prefiltered data, respectively.
2. **CLS, CLS-F** least-squares estimation with compensation, cf. Eqs. (33)-(34), using averaged data and averaged and prefiltered data, respectively.
3. **TLS, TLS-A, TLS-AF** Total least-squares estimation using raw data, averaged data, and averaged and prefiltered data, respectively.
4. **GTLS, GTLS-F** Generalized total least-squares estimation with estimated noise statistics using averaged data and averaged and prefiltered data, respectively.
5. **CTLS, CTLS-F** Total least-squares estimation with compensation, cf. Eq. (52), using averaged data and averaged and prefiltered data, respectively.

With colored noise on both the input and the output the LS and TLS method will be biased. LS-A, LS-AF, TLS-A, and TLS-AF will be consistent as the number of periods tends to infinity, otherwise these methods will also give biased results. The other methods give consistent estimates regardless of the noise and the number of periods used (as long as $M \geq 4$).

In the simulation we used $P = 64$, $M = 32$, $\sigma_e^2 = 0.09$, $\sigma_n^2 = 0.01$. In the prefiltering of the data using a nonparametric noise model we chose to whiten the data on the output. The results of a Monte Carlo simulation consisting of 16 different runs are summarized in Table F.1. The numbers shown are the means and standard deviations of the estimated parameter values for each method.

Table F.1 *Summary of identification results.*

Parameter	a_1	a_2	b_0	b_1
True value	-1.5000	0.7000	1.0000	0.5000
LS	-1.0799	0.2033	0.9969	0.9158
	±0.1340	±0.1079	±0.1037	± 0.0634
TLS	-0.9891	-0.5340	1.7171	3.2591
	±1.6012	±1.5505	±3.0464	±4.6528
LS-A	-1.4798	0.6766	0.9963	0.5240
	±0.0251	±0.0247	±0.0565	±0.0637
CLS	-1.4991	0.6992	0.9970	0.5058
	±0.0246	±0.0238	±0.0586	±0.0647
TLS-A	-1.5026	0.6874	1.0412	0.5245
	±0.0259	±0.0244	±0.0656	±0.0670
GTLS	-1.5095	0.7087	1.0209	0.4788
	±0.0266	±0.0252	±0.0602	±0.0666
CTLS	-1.5102	0.7131	1.0170	0.4698
	±0.0255	±0.0235	±0.0636	±0.0634
LS-AF	-1.4992	0.6991	0.9991	0.4994
	±0.0030	±0.0028	±0.0116	±0.0126
CLS-F	-1.5006	0.7006	0.9988	0.4984
	±0.0032	±0.0029	±0.0118	±0.0127
TLS-AF	-1.5002	0.7001	1.0013	0.4981
	±0.0032	±0.0029	±0.0118	±0.0127
GTLS-F	-1.5009	0.7008	1.0013	0.4957
	±0.0033	±0.0030	±0.0118	±0.0128
CTLS-F	-1.5013	0.7012	1.0013	0.4944
	±0.0032	±0.0029	±0.0118	±0.0126

Studying Table F.1, we can first note that the LS and TLS methods perform very badly, while the results are quite good when averaged data is used. Focusing on the proposed methods, CLS and CTLS, we see that these compare well with the other methods, both with averaged data and with averaged and prefiltered data. In this example, the improvement in the accuracy with prefiltered data is substantial despite the poor variance properties of the periodogram (15). This holds for all methods, as can be seen from Table F.1.

12 Conclusions

We have studied the problem of identifying dynamic errors-in-variables systems using periodic excitation signals. Two new algorithms, the CLS and the CTLS algorithms, have been presented that gives consistent estimates regardless of the noise on the input and output. With the CLS algorithm the estimate is found without iterations by solving a standard least-squares problem, which can be done very efficiently using FFT. This method can therefore be an interesting alternative to existing time- and frequency-domain methods for this problem. The CTLS algorithm is an alternative to the GTLS algorithm where the noise statistics are estimated from data. The performance of the CTLS method and the GTLS method is similar.

References

[1] B. D. O. Anderson. Identification of scalar errors-in-variables models with dynamics. *Automatica*, 21:709–716, 1985.

[2] C. T. Chou and M. Verhaegen. Subspace algorithms for the identification of multivariable errors-in-variables models. *Automatica*, 33:1857–1869, 1997.

[3] F. Gustafsson and J. Schoukens. Utilizing periodic excitation in prediction error based system identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3926–3931, Tampa, FL, 1998.

[4] T. McKelvey. Periodic excitation for identification of dynamic errors-in-variables systems operating in closed loop. In *Proceedings of the 13th IFAC World Congress*, volume J, pages 155–160, San Francisco, CA, 1996.

[5] J. Schoukens, P. Guillaume, and R. Pintelon. Design of broadband excitation signals. In K. Godfrey, editor, *Perturbation Signals for System Identification*, pages 126–159. Prentice-Hall, 1993.

[6] J. Schoukens and R. Pintelon. *Identification of Linear Systems. A Practical Guideline to Accurate Modeling.* Pergamon Press, 1991.

[7] J. Schoukens, R. Pintelon, G. Vandersteen, and P. Guillaume. Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 33:1073–1086, 1997.

[8] J. Schoukens, Y. Rolain, F. Gustafsson, and R. Pintelon. Fast calculations of linear and non-linear least-squares estimates for system identification. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3408–3410, Tampa, Florida, 1998.

[9] T. Söderström. Identification of stochastic linear systems in presence of input noise. *Automatica*, 17:713–725, 1981.

[10] P. Stoica and T. Söderström. Bias correction in least-squares identification. *International Journal of Control*, 35:449–457, 1982.

[11] S. Van Huffel and J. Vandewalle. Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error. *SIAM Journal on Matrix Analysis and Applications*, 10:294–315, 1989.

[12] W.-X. Zheng and C.-B. Feng. Unbiased parameter estimation of linear systems in the presence of input and output noise. *International Journal of Adaptive Control and Signal Processing*, 3:231–251, 1989.

[13] W.-X. Zheng and C.-B. Feng. Identification of a class of dynamic errors-in-variables models. *International Journal of Adaptive Control and Signal Processing*, 6:431–440, 1992.

Paper G

Identification for Control: Some Results on Optimal Experiment Design

Authors: Urban Forssell and Lennart Ljung

Edited version of paper originally published in *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 3384-3389, Tampa, FL, 1998.

Submitted to *Automatica*.

Identification for Control: Some Results on Optimal Experiment Design

Urban Forssell and Lennart Ljung

Division of Automatic Control,
Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: ufo@isy.liu.se, ljung@isy.liu.se.

Abstract

The problem of designing the identification experiments to make them maximally informative with respect to the intended use of the model is studied. A focus will be on model based control and we show how to choose the feedback regulator and the spectrum of the reference signal in case of closed-loop experiments. A main result is that when only the misfit in the dynamics model is penalized and when both the input and the output power are constrained then the optimal controller is given by the solution to a standard LQ problem. When only the input power is constrained, it is shown that open-loop experiments are optimal. Some examples are also given to exemplify the theoretical results.

1 Introduction

The validation step in the identification procedure should be devoted to evaluating the model with respect to its intended use; simulation, prediction, control, etc. To make fair judgments it is important to know what factors affect the model's performance and how these factors can be manipulated. In this paper we focus on experiment design issues and especially how one should choose the feedback regulator in closed-loop experiments. The aim will be to design the experiments so as to minimize the variance errors in the identified models. This route was also taken in [2] and [4], and the presentation here will be inspired by these papers. See also [7].

The main drawback with the optimal solutions that will be derived is that they typically will depend on unknown quantities, like the unknown system that we are trying to identify. Clearly this will lead to infeasible designs. This problem is inevitable and shared with any optimal design problem. A way around this is to use iterations and to replace the true system in the optimal solutions with the current estimate of the system.

Despite this problem the optimal solutions derived here give important and useful insights into questions like "Should open-loop or closed-loop experiments be used?" and "To which frequency ranges should the input signal power be concentrated?".

The rest of the presentation will unveil as follows. Section 2 contains some background material on the identification method used and the assumptions made on the underlying system. The notation will also be introduced in this section. In Section 3 we provide a motivation for the type of criterion we will consider. A key issue is that the criterion will reflect the variance error introduced when using identified models in the control design. Section 4 contains the main results and in Section 5 these results are applied to some well known control design schemes like internal model control and model reference control. Conclusions are given in Section 6.

2 Preliminaries

To evaluate a model's performance one needs some notion of the "best" or "true" model. We will assume that the true system is single-input, single-output (SISO) and given by

$$y(t) = G_0(q)u(t) + v(t), \quad v(t) = H_0(q)e(t) \tag{1}$$

Here $\{e(t)\}$ is white noise with variance λ_0. To simplify the notation we introduce the transfer matrix

$$T_0(q) = \begin{bmatrix} G_0(q) & H_0(q) \end{bmatrix}^T \tag{2}$$

Given measured data $\{y(1), u(1), \dots, y(N), u(N)\}$ and a parameterized model structure

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t) \tag{3}$$

we can estimate the model by the straightforward fit:

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta) \tag{4}$$

$$V_N(\theta) = \sum_{t=1}^{N} \varepsilon^2(t, \theta) \tag{5}$$

$$\varepsilon(t, \theta) = y(t) - \hat{y}(t|\theta) = H^{-1}(q, \theta)(y(t) - G(q, \theta)u(t)) \tag{6}$$

For more details on this identification method see, e.g., [7]. The model structure (3) may also be represented as

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \tag{7}$$

In the sequel $G(q, \theta)$ and $H(q, \theta)$ will be called the dynamics model and the noise model, respectively. Let $\hat{G}_N(q) = G(q, \hat{\theta}_N)$ and $\hat{H}_N(q) = H(q, \hat{\theta}_N)$ and also introduce

$$\hat{T}_N(q) = \begin{bmatrix} \hat{G}_N(q) & \hat{H}_N(q) \end{bmatrix}^T \tag{8}$$

The error between the true transfer matrix $T_0(q)$ and the estimated one $\hat{T}_N(q)$ will be denoted $\tilde{T}_N(q)$,

$$\tilde{T}_N(q) = T_0(q) - \hat{T}_N(q) = \begin{bmatrix} \tilde{G}_N(q) & \tilde{H}_N(q) \end{bmatrix}^T$$

with an obvious notation for $\tilde{G}_N(q)$ and $\tilde{H}_N(q)$. As the number of data tends to infinity and the model order n increases we have that (cf. [5], [9])

$$\text{Cov } \hat{T}_N(e^{i\omega}) \approx \frac{n}{N} P(\omega) \tag{9}$$

$$= \frac{n}{N} \Phi_v(\omega) \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{ue}(-\omega) & \lambda_0 \end{bmatrix}^{-T} \tag{10}$$

In this paper we will assume that the input is generated as

$$u(t) = r(t) - K(q)y(t) \tag{11}$$

where $K(q)$ is a linear regulator. The reference signal $r(t)$ is assumed uncorrelated with the disturbance $e(t)$, i.e., the cross spectrum $\Phi_{re}(\omega) = 0$. This means that we can split the input spectrum into two parts, one that is due to $r(t)$ and one that is due to $e(t)$:

$$\Phi_u(\omega) = \Phi_u^r(\omega) + \Phi_u^e(\omega) \tag{12}$$

By combining (1) and (11) we see that the input can be written

$$u(t) = S_0(q)r(t) - K(q)S_0(q)H_0(q)e(t) \tag{13}$$

Consequently

$$\Phi_u^r(\omega) = |S_0(e^{i\omega})|^2 \Phi_r(\omega) \tag{14}$$
$$\Phi_u^e(\omega) = |K(e^{i\omega})|^2 |S_0(e^{i\omega})|^2 |H_0(e^{i\omega})|^2 \lambda_0 \tag{15}$$
$$\Phi_{ue}(\omega) = -K(e^{i\omega}) S_0(e^{i\omega}) H_0(e^{i\omega}) \lambda_0 \tag{16}$$

With these definitions we may rewrite $P(\omega)$ as

$$\begin{aligned} P(\omega) &= \Phi_v(\omega) \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{ue}(-\omega) & \lambda_0 \end{bmatrix}^{-T} \\ &= \frac{\Phi_v(\omega)}{\lambda_0 \Phi_u(\omega) - |\Phi_{ue}(\omega)|^2} \begin{bmatrix} \lambda_0 & -\Phi_{ue}(\omega) \\ -\Phi_{ue}(-\omega) & \Phi_u(\omega) \end{bmatrix}^T \\ &= \frac{|H_0(e^{i\omega})|^2}{\Phi_u^r(\omega)} \begin{bmatrix} \lambda_0 & -\Phi_{ue}(\omega) \\ -\Phi_{ue}(-\omega) & \Phi_u^r(\omega) + \Phi_u^e(\omega) \end{bmatrix}^T \end{aligned} \tag{17}$$

Here we have used the fact that

$$\Phi_v(\omega) = \lambda_0 |H_0(e^{i\omega})|^2 \tag{18}$$

A remark regarding (17) is that this result holds not only for the direct method, considered here, but also for the so called indirect and joint input-output approaches to closed-loop identification. See, e.g., [1] and [3]. Therefore all the results below will also apply to these approaches as well.

For future use we also note that the output spectrum can be written

$$\Phi_y(\omega) = |G_0(e^{i\omega})|^2 \Phi_u^r(\omega) + |S_0(e^{i\omega})|^2 \Phi_v(\omega) \tag{19}$$

3 Measuring the Performance Degradation

Our focus here will be on model-based control. It is then natural to evaluate the model's performance by comparing the outputs of the closed-loop system resulting from control design using the "true" models and the estimated ones, respectively. In general this output will be a function of both $G(q)$ and $H(q)$ and some driving signal $w(t)$:

$$y(t) = f(T(q))w(t) \tag{20}$$

The "true" or "best" signal is

$$y_0(t) = f(T_0(q))w(t) \tag{21}$$

while the one resulting from the estimated models is

$$\hat{y}_N(t) = f(\hat{T}_N(q))w(t) \tag{22}$$

A measure of the performance degradation can be obtained by computing the mean square error

$$E\tilde{y}_N^2(t) = E(y_0(t) - \hat{y}_N(t))^2 \tag{23}$$

Using first-order approximations we may write

$$\tilde{y}_N(t) = \tilde{T}_N^T(q)F(q)w(t) \tag{24}$$

where

$$F(q) = \frac{\partial}{\partial T} f(T)\Big|_{T=T_0(q)} \tag{25}$$

From the definition of the spectrum it follows that

$$E\tilde{y}_N^2(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{\tilde{y}_N}(\omega)d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr}\left[\Pi(\omega)C(\omega)\right] d\omega \tag{26}$$

where

$$\Pi(\omega) = E\tilde{T}_N(e^{-i\omega})\tilde{T}_N^T(e^{i\omega}) \tag{27}$$

$$C(\omega) = F(e^{i\omega})\Phi_w(\omega)F^T(e^{-i\omega}) \tag{28}$$

Thus, if we want to minimize the performance degradation due to the use of identified models, a natural criterion to minimize would be

$$\bar{J} = \int_{-\pi}^{\pi} \operatorname{tr}\left[\Pi(\omega)C(\omega)\right] d\omega \tag{29}$$

Let us comment on the matrix functions $\Pi(\omega)$ and $C(\omega)$. First, if the bias errors are negligible then $\Pi(\omega)$ will be the (transpose of the) covariance matrix of the estimate $\hat{T}_N(e^{i\omega})$,

$$\Pi(\omega) = \left[\mathrm{Cov}\ \tilde{T}_N(e^{i\omega})\right]^T \tag{30}$$

The error made in this approximation should be small if the model structure is reasonably flexible. Using the standard result (9) we thus have

$$\Pi(\omega) \approx \frac{n}{N}P^T(\omega) = \frac{n}{N}\Phi_v(\omega)\begin{bmatrix}\Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{ue}(-\omega) & \lambda_0\end{bmatrix}^{-1} \tag{31}$$

Second, the weighting function $C(\omega)$ should reflect the relative importance of a good fit in $G(e^{i\omega})$ and $H(e^{i\omega})$, respectively, for different frequencies. We shall assume that $C(\omega)$ is Hermitian and positive semi-definite as in (28). In general we may write

$$C(\omega) = \begin{bmatrix} C_{11}(\omega) & C_{12}(\omega) \\ C_{12}(-\omega) & C_{22}(\omega) \end{bmatrix} \tag{32}$$

If we neglect the scaling n/N (which will have the affect that the prejudice against high order models is reduced [7]), the criterion (29) simplifies to

$$J = \int_{-\pi}^{\pi} \operatorname{tr}\left[P(\omega)C(\omega)\right] d\omega \tag{33}$$

This criterion will be used here.

4 Main Results

In this section we will consider experiment design problems where the goal is to minimize J with respect to the design variables, which we take as $K(e^{i\omega})$ and $\Phi_r(\omega)$. Other equivalent choices are also possible, e.g., $\Phi_u(\omega)$ and $\Phi_{ue}(\omega)$ or $K(e^{i\omega})$ and $\Phi_u^r(\omega)$. To make the designs realistic we will also impose constraints on the input power or the output power, or both. We will also discuss the case where the power of the reference signal is bounded. Without loss of generality we will take the upper bounds to be unity. This is done to simplify the notation and can always be achieved by suitable scaling. In the sequel we also will omit the arguments ω and $e^{i\omega}$ whenever there is no risk of confusion.

First consider the case where penalty only is put on the misfit in G. This corresponds to $C_{12} = C_{22} = 0$. We then have the following result which to our knowledge is new.

Theorem 1 *The solution to the problem*

$$\min_{K,\Phi_r} \left\{ \int_{-\pi}^{\pi} \operatorname{tr}\left[PC\right] d\omega : \; C_{12} = 0, \; C_{22} = 0; \right.$$
$$\left. \int_{-\pi}^{\pi} \{\alpha\Phi_u + (1-\alpha)\Phi_y\} d\omega \le 1, \; \alpha \in [0,1] \right\} \tag{34}$$

is

$$K^{opt} = \arg\min_K \int_{-\pi}^{\pi} \frac{\alpha|K|^2 + (1-\alpha)}{|1+G_0K|^2} \Phi_v d\omega \tag{35}$$

$$\Phi_r^{opt} = \mu\sqrt{\Phi_v C_{11}} \frac{|1+G_0K^{opt}|^2}{\sqrt{\alpha + (1-\alpha)|G_0|^2}} \tag{36}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} \{\alpha\Phi_u + (1-\alpha)\Phi_y\} d\omega = 1 \tag{37}$$

Remark: It should be noted that equation (35), defining the optimal controller K^{opt}, is a standard LQ problem. If G_0 and Φ_v were known this could be easily solved to give an explicit expression for K^{opt}.

Proof *Replace the design variables K and Φ_r by the equivalent pair K and Φ_u^r. Then, by using* (12)*,* (17)*, and* (19)*, problem* (34) *can be rewritten as*

$$\min_{K,\Phi_u^r} \left\{ \int_{-\pi}^{\pi} \frac{\Phi_v}{\Phi_u^r} C_{11} d\omega : \right. \\ \int_{-\pi}^{\pi} \{(\alpha + (1-\alpha)|G_0|^2)\Phi_u^r + \frac{\alpha|K|^2 + (1-\alpha)}{|1+G_0K|^2}\Phi_v\} d\omega \leq 1, \\ \left. \alpha \in [0,1] \right\} \tag{38}$$

The criterion function is independent of K hence the optimal controller K^{opt} can be found by solving the LQ problem

$$\min_K \int_{-\pi}^{\pi} \frac{\alpha|K|^2 + (1-\alpha)}{|1+G_0K|^2}\Phi_v d\omega \tag{39}$$

(Here it is implicitly assumed that $y(t) = G_0(q)u(t) + v(t)$, $u(t) = -K(q)y(t)$, and $\alpha \in [0,1]$.) This proves (35)*. Define the constant γ as*

$$\gamma = 1 - \int_{-\pi}^{\pi} \frac{\alpha|K^{opt}|^2 + (1-\alpha)}{|1+G_0K^{opt}|^2}\Phi_v d\omega \tag{40}$$

Problem (38) *now reads*

$$\min_{\Phi_u^r} \{ \int_{-\pi}^{\pi} \frac{\Phi_v}{\Phi_u^r} C_{11} d\omega : \int_{-\pi}^{\pi} (\alpha + (1-\alpha)|G_0|^2)\Phi_u^r d\omega \leq \gamma \} \tag{41}$$

This problem has the solution (cf. [7], p. 376)

$$\Phi_u^r = \mu \sqrt{\frac{\Phi_v C_{11}}{(\alpha + (1-\alpha)|G_0|^2)}} \tag{42}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} ((\alpha + (1-\alpha)|G_0|^2)\Phi_u^r d\omega = \gamma \tag{43}$$

or in other words so that

$$\int_{-\pi}^{\pi} \{\alpha\Phi_u + (1-\alpha)\Phi_y\} d\omega = 1 \tag{44}$$

Consequently the optimal Φ_r is

$$\Phi_r^{opt} = \mu\sqrt{\Phi_v C_{11}} \frac{|1 + G_0 K^{opt}|^2}{\sqrt{\alpha + (1-\alpha)|G_0|^2}} \tag{45}$$

which ends the proof. □

We stress that the optimal controller K^{opt} in (35) can easily be found by solving the indicated discrete-time LQ problem (if G_0 and Φ_v were known). Among other things this implies that the optimal controller K^{opt} is guaranteed to stabilize the closed-loop system and be linear, of the same order as G_0. This is a clear advantage over the results reported in, e.g., [4] and also over the results in Theorem 3 below. Furthermore, the optimal controller is independent of C_{11} which also is quite interesting and perhaps somewhat surprising. This means that whatever weighting C_{11} is used in the design criterion, it is always optimal to use the LQ regulator (35) in the identification experiment.

From Theorem 1 we also see that closed-loop experiments are optimal as long as the is a constraint on the output power, i.e., as long as $\alpha \neq 1$. If $\alpha = 1$ then $K^{opt} = 0$ and the optimal input spectrum $\Phi_u^{opt} (= \Phi_r^{opt})$ becomes

$$\Phi_u^{opt} = \mu\sqrt{\Phi_v C_{11}} \tag{46}$$

This side result can in fact be strengthened [7]:

Theorem 2 *The problem*

$$\min_{K,\Phi_r} \left\{ \int_{-\pi}^{\pi} \operatorname{tr}[PC]\, d\omega : \int_{-\pi}^{\pi} \Phi_u d\omega \leq 1,\ C_{12} = 0 \right\} \tag{47}$$

has the solution

$$K^{opt} = 0 \quad \textit{(open-loop operation)} \tag{48}$$

$$\Phi_r^{opt} = \Phi_u^{opt} = \mu\sqrt{\Phi_v C_{11}} \tag{49}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} \Phi_u d\omega = 1 \tag{50}$$

Consider now the case where the power of the reference signal is constrained. The following theorem, which we believe is new, then states what the optimal controller and reference spectrum are. The expression for the optimal controller also holds for the case where the reference signal is fixed and pre-specified (which would be the case if data were collected under normal (controlled) operation of the plant).

Theorem 3 *The problem*

$$\min_{K,\Phi_r} \left\{ \int_{-\pi}^{\pi} \operatorname{tr}\left[PC\right] d\omega : \int_{-\pi}^{\pi} \Phi_r d\omega \le 1 \right\} \tag{51}$$

has the solution

$$K^{opt} = -\frac{C_{11}G_0^* + C_{12}H_0^*}{C_{11}|G_0|^2 + 2\operatorname{Re}(C_{12}G_0H_0^*) + C_{22}|H_0|^2} \tag{52}$$

$$\Phi_r^{opt} = \mu\sqrt{\frac{\Phi_v \cdot \det C \cdot |H_0|^2}{C_{11}|G_0|^2 + 2\operatorname{Re}(C_{12}G_0H_0^*) + C_{22}|H_0|^2}} \tag{53}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} \Phi_r d\omega = 1 \tag{54}$$

Remark: Note that K^{opt} in (52) only is a feasible design if K^{opt} is proper and stabilizes the closed-loop system. Unfortunately no such guarantees can be given, which is a clear disadvantage compared to the result in Theorem 1. The optimal solutions derived in [4] suffer from the same type of problems, as mentioned earlier. To circumvent this problem it was suggested in [4] that the so called Youla parameterization (e.g., [8]) should be used. This idea could lead to interesting extensions of the results in Theorem 3 also, although we will not study it any further here.

Proof *Introduce the following square-root factorization of C:*

$$C = \begin{bmatrix} C_1 & C_2 \\ 0 & C_3 \end{bmatrix} \begin{bmatrix} C_1 & C_2 \\ 0 & C_3 \end{bmatrix}^* \tag{55}$$

Clearly we have the relations

$$C_{11} = |C_1|^2 + |C_2|^2 \tag{56}$$
$$C_{12} = C_2 C_3^* \tag{57}$$
$$C_{22} = |C_3|^2 \tag{58}$$

With P given by (17) *we have that*

$$\operatorname{tr}\left[PC\right] = C_{22}|H_0|^2 + \Psi \tag{59}$$

where

$$\Psi = \frac{\Phi_v}{\Phi_u^r}\left[C_{11} + 2\,\mathrm{Re}(C_{12}^* K S_0 H_0) + C_{22}|K|^2|S_0|^2|H_0|^2\right] \tag{60}$$

$$= \frac{\Phi_v}{\Phi_u^r}\left[|C_1|^2 + |C_2 + C_3 K S_0 H_0|^2\right] \tag{61}$$

$$= \frac{\Phi_v}{\Phi_r}\left[|C_1(1+KG_0)|^2 + |C_2(1+KG_0) + C_3 K H_0|^2\right] \tag{62}$$

$$= \frac{\Phi_v}{\Phi_r}\left\|\begin{bmatrix} C_1 & C_2 \end{bmatrix} + K\begin{bmatrix} C_1 G_0 & C_2 G_0 + C_3 H_0 \end{bmatrix}\right\|^2 \tag{63}$$

There is no constraint on K so to minimize the criterion function we should, at each frequency, choose K as the least-squares estimate

$$K^{opt} = -\frac{\begin{bmatrix} C_1 & C_2 \end{bmatrix}\begin{bmatrix} C_1 G_0 & C_2 G_0 + C_3 H_0 \end{bmatrix}^*}{\left\|\begin{bmatrix} C_1 G_0 & C_2 G_0 + C_3 H_0 \end{bmatrix}\right\|^2} \tag{64}$$

$$= -\frac{(|C_1|^2 + |C_2|^2)G_0^* + C_2 C_3^* H_0^*}{(|C_1|^2 + |C_2|^2)|G_0|^2 + 2\,\mathrm{Re}(C_2 C_3^* G_0 H_0^*) + |C_3|^2|H_0|^2} \tag{65}$$

$$= -\frac{C_{11}G_0^* + C_{12}H_0^*}{C_{11}|G_0|^2 + 2\,\mathrm{Re}(C_{12}G_0 H_0^*) + C_{22}|H_0|^2} \tag{66}$$

With this choice of K it can be shown (by straightforward but tedious calculations) that

$$\mathrm{tr}\,[PC]\Big|_{K=K^{opt}} = C_{22}|H_0|^2 + \\ + \frac{\Phi_v}{\Phi_r}\frac{\det C \cdot |H_0|^2}{C_{11}|G_0|^2 + 2\,\mathrm{Re}(C_{12}G_0 H_0^*) + C_{22}|H_0|^2} \tag{67}$$

The first term is independent of Φ_r, hence we get (cf. [7], p. 376)

$$\Phi_r^{opt} = \mu\sqrt{\frac{\Phi_v \cdot \det C \cdot |H_0|^2}{C_{11}|G_0|^2 + 2\,\mathrm{Re}(C_{12}G_0 H_0^*) + C_{22}|H_0|^2}} \tag{68}$$

where μ is a constant, adjusted so that

$$\int_{-\pi}^{\pi} \Phi_r^{opt} d\omega = 1$$

This proves the theorem. □

Note that if C is singular the optimal Φ_r in (53) can be chosen at will, as long as

$$\int_{-\pi}^{\pi} \Phi_r^{opt} d\omega = 1 \tag{69}$$

This is due to the factor $\det C$ in (53) which vanishes when C is singular. Furthermore, when C is singular the optimal K in (52) simplifies to (cf. [6, 2])

$$K^{opt} = -\frac{C_{12}}{C_{12}G_0 + C_{22}H_0} \tag{70}$$

Theorem 4 *If C is singular then the problem*

$$\min_{K,\Phi_r} \left\{ \int_{-\pi}^{\pi} \operatorname{tr}[PC]\, d\omega : \textit{any constraint} \right\} \tag{71}$$

has the solution

$$K^{opt} = -\frac{C_{12}}{C_{12}G_0 + C_{22}H_0} \tag{72}$$

$$\Phi_r^{opt} \quad \textit{arbitrary} \tag{73}$$

as long as u given by

$$u(t) = r(t) + \frac{C_{12}(q)}{C_{12}(q)G_0(q) + C_{22}(q)H_0(q)} y(t) \tag{74}$$

is an admissible input.

5 Examples

5.1 Internal Model Control

In Internal Model Control (IMC) the control law is chosen as

$$u(t) = K(q)(r_c(t) - y(t)) = \frac{Q(q)}{1 - G(q)Q(q)}(r_c(t) - y(t)) \tag{75}$$

where Q is some stable transfer function and G a model of the true plant G_0. Here r_c denotes the reference signal to be used in the control application, not to be mixed up with the reference signal r used in the identification experiment. Note that the control design will not depend on the noise model. Also note that the designed closed-loop system is

$$G_c^{des}(q) = \frac{G(q)K(q)}{1 + G(q)K(q)} = G(q)Q(q) \tag{76}$$

For this 1-dof regulator G_c coincides with the complementary sensitivity function. From (76) we see that $Q = G_c/G$ should be large where a raise in the system's gain is desired.

Ideally, that is when $G = G_0$, the "best" or "true" IMC controller results:

$$K_0(q) = \frac{Q(q)}{1 - G_0(q)Q(q)} \tag{77}$$

In general the model G is obtained through identification using N data samples. In that case $G = \hat{G}_N$ and the corresponding controller is

$$\hat{K}_N(q) = \frac{Q(q)}{1 - \hat{G}_N(q)Q(q)} \tag{78}$$

Neglecting the response from the noise source e we obtain the "true" output y_0 as

$$y_0(t) = \frac{G_0(q)K_0(q)}{1 + G_0(q)K_0(q)} r_c(t) = G_0(q)Q(q)r_c(t) \tag{79}$$

The actual output resulting from use of the suboptimal controller (78) is

$$\hat{y}_N(t) = \frac{G_0(q)\hat{K}_N(q)}{1 + G_0(q)\hat{K}_N(q)} r_c(t) = \frac{G_0(q)Q(q)}{1 + \tilde{G}_N(q)Q(q)} r_c(t) \tag{80}$$

Using these expressions it can now be shown that the weighting matrix C (cf. Section 3) in this case is given by

$$C_{IMC} = \begin{bmatrix} |G_0|^2|Q|^4\Phi_{r_c} & 0 \\ 0 & 0 \end{bmatrix} \tag{81}$$

Suppose now that we want to solve the problem

$$\min_{K,\Phi_r} \left\{ \int_{-\pi}^{\pi} \operatorname{tr}\left[PC_{IMC}\right] d\omega : \int_{-\pi}^{\pi} \{\Phi_u + \Phi_y\} d\omega < \infty \right\} \tag{82}$$

where K and r refers to the set-up (11) used in the identification. Theorem 1 then states that the optimal strategy is to use closed-loop experiments with K and Φ_r chosen as

$$K^{opt} = \arg\min_K \int_{-\pi}^{\pi} \frac{|K|^2 + 1}{|1 + G_0K|^2} \Phi_v d\omega \tag{83}$$

$$\Phi_r^{opt} \propto \sqrt{\Phi_v \Phi_{r_c}} \cdot |G_0| \cdot |Q|^2 \cdot \frac{|1 + G_0K^{opt}|^2}{\sqrt{1 + |G_0|^2}} \tag{84}$$

A remark on this choice of reference spectrum is that (for fixed G_0 and Φ_v) Φ_r^{opt} will be large where either of Φ_{r_c} or $|Q| \approx |G_c^{des}/G_0|$, is large.

5.2 Generalized Minimum Variance Control

In Generalized Minimum Variance Control (GMVC) the goal is to make the closed-loop system behave like

$$y_0(t) = R(q)e(t) \tag{85}$$

This can be accomplished by using a regulator of the form

$$u(t) = -\frac{H_0(q) - R(q)}{G_0(q)R(q)} y(t) \tag{86}$$

The corresponding C-matrix is given by [2]:

$$C_{GMVC} = \lambda_0 \left| \frac{R}{H_0} \right|^2 \begin{bmatrix} \frac{|H_0 - R|^2}{|G_0|^2} & -\frac{H_0 - R}{G_0} \\ -\frac{H_0^* - R^*}{G_0^*} & 1 \end{bmatrix} \tag{87}$$

Note that C_{GMVC} is singular. Hence the choice

$$K_y^{opt} = -\frac{-(H_0 - R)/G_0}{-(H_0 - R) + H_0} = \frac{H_0 - R}{G_0 R} \tag{88}$$

minimizes

$$\bar{E}\tilde{y}_N^2(t) \tag{89}$$

subject to any constraint according to Theorem 4. In other words, the optimal controller is the true GMVC controller. This was first shown in [2].

5.3 Model Reference Control

In Model Reference Control (MRC) the closed-loop system should ideally behave like

$$y_0(t) = Q(q)r_c(t) + R(q)e(t) \tag{90}$$

for some desired Q and R[1]. This can be achieved by letting

$$u(t) = K_{r,0}(q)r_c(t) - K_{y,0}(q)y(t) \tag{91}$$

where

$$K_{r,0}(q) = \frac{Q(q)H_0(q)}{R(q)G_0(q)}, \quad K_{y,0}(q) = \frac{H_0(q) - R(q)}{G_0(q)R(q)} \tag{92}$$

[1] The case where $Q = 0$ is also termed generalized minimum variance control.

When using identified models $\hat{G}_N$ and $\hat{H}_N$ in the control design we will obtain

$$\hat{K}_{r,N}(q) = \frac{Q(q)\hat{H}_N(q)}{R(q)\hat{G}_N(q)}, \quad \hat{K}_{y,N}(q) = \frac{\hat{H}_N(q) - R(q)}{\hat{G}_N(q)R(q)} \tag{93}$$

giving the sub-optimal output $\hat{y}_N$. By studying the degradation signal $\tilde{y}_N = y_0 - \hat{y}_N$ it can be shown that the C-matrix in this case is given by

$$C_{MRC} = \left|\frac{R}{H_0}\right|^2 \begin{bmatrix} \frac{|Q|^2}{|G_0|^2}\Phi_{r_c} + \frac{|H_0-R|^2}{|G_0|^2}\lambda_0 & -\frac{H_0-R}{G_0}\lambda_0 \\ -\frac{H_0^*-R^*}{G_0^*}\lambda_0 & \lambda_0 \end{bmatrix} \tag{94}$$

Note that

$$C_{MRC} = \left|\frac{R}{H_0}\right|^2 \begin{bmatrix} \frac{|Q|^2}{|G_0|^2}\Phi_{r_c} & 0 \\ 0 & 0 \end{bmatrix} + C_{GMVC} \tag{95}$$

Consider now the design problem

$$\min_{K,\Phi_r} \left\{ \int_{-\pi}^{\pi} \operatorname{tr}\left[PC_{MRC}\right] d\omega : r(t) = r_c(t) \right\} \tag{96}$$

where K and r refers to (11). The constraint that $r = r_c$ means that the identification experiment should be performed using the same reference signal as in the control application, or in other words that the data should be collected under normal operation of the controlled plant.

Using the results of Theorem 3 it can be shown that the optimal controller is

$$K^{opt} = -\frac{1}{G_0}\frac{|Q|^2\Phi_{r_c} + (|R|^2 - H_0R^*)\lambda_0}{|Q|^2\Phi_{r_c} + |R|^2\lambda_0} \tag{97}$$

$$= -\frac{1}{G_0}\left(1 - \frac{H_0R^*}{|Q|^2\Phi_{r_c}/\lambda_0 + |R|^2}\right) \tag{98}$$

It is interesting to study how the optimal solution changes as a function of the control design variables Q, R, Φ_{r_c} and λ_0. Here we consider the following special cases:

- When $\Phi_{r_c}/\lambda_0 \to \infty$,

$$K^{opt} \to -\frac{1}{G_0} \tag{99}$$

  We recognize the right-hand side expression as the optimal controller when $C_{12} = 0$, $C_{22} = 0$, cf. (52).

- When $\Phi_{r_c}/\lambda_0 \to 0$,

$$K^{opt} \to -\frac{|R|^2 - H_0 R^*}{G_0 |R|^2} = \frac{H_0 - R}{G_0 R} = K_{y,0} \tag{100}$$

The true MRC feedback controller $K_{y,0}$ is thus optimal in the identification only when $\Phi_{r_c}/\lambda_0 = 0$ or when $Q = 0$ (compare with the results in [4, 2]).

6 Conclusions

We have studied the problem of optimal experiment design. A main result was that we, possibly for the first time, have been able to give explicit solution to a experiment design problem with constraints on both the input and output power. The optimal controller is in this case given by the solution to a standard LQ problem. Another main conclusion is that closed-loop experiments are optimal whenever the output variance is constrained, otherwise open-loop experiments are preferable.

References

[1] U. Forssell and L. Ljung. Closed-loop identification revisited – Updated version. Technical Report LiTH-ISY-R-2021, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998.

[2] M. Gevers and L. Ljung. Optimal experiment design with respect to the intended model application. *Automatica*, 22:543–554, 1986.

[3] M. Gevers, L. Ljung, and P. Van den Hof. Asymptotic variance expression for closed-loop identification and their relevance in identification for control. In *Proceedings of the 11th IFAC Symposium on System Identification*, volume 3, pages 1449–1454, Fukuoka, Japan, 1997.

[4] H. Hjalmarsson, M. Gevers, and F. De Bruyne. For model-based control design, closed loop identification gives better performance. *Automatica*, 32:1659–1673, 1996.

[5] L. Ljung. Asymptotic variance expressions for identified black-box transfer function models. *IEEE Transactions on Automatic Control*, 30:834–844, 1985.

[6] L. Ljung. Parametric methods for identification of transfer functions of linear systems. In C. L. Leondes, editor, *Advances in Control*, volume 24. Academic Press, 1986.

[7] L. Ljung. *System Identification: Theory for the User.* Prentice-Hall, 1987.

[8] M. Vidyasagar. *Control System Synthesis: A Factorization Approach.* MIT Press, 1985.

[9] Y.-C. Zhu. Black-box identification of MIMO transfer functions: Asymptotic properties of prediction error models. *International Journal of Adaptive Control and Signal Processing*, 3:357–373, 1989.

Paper H

Asymptotic Variance Expressions for Identified Black-box Models

Author: Urban Forssell

Preliminary version published as Technical Report LiTH-ISY-R-2089, Department of Electrical Engineering, Linköping University, Linköping, Sweden.

Submitted to *Systems & Control Letters*.

Asymptotic Variance Expressions for Identified Black-box Models

Urban Forssell

Department of Electrical Engineering,
Linköping University,
S-581 83 Linköping, Sweden.
E-mail: ufo@isy.liu.se.

Abstract

The asymptotic probability distribution of identified black-box transfer function models is studied. The main contribution is that we derive variance expressions for the real and imaginary parts of the identified models that are asymptotic in both the number of measurements and the model order. These expressions are considerably simpler than the corresponding ones that hold for fixed model orders, and yet they frequently approximate the true covariance well already with quite modest model orders. We illustrate the relevance of the asymptotic expressions by using them to compute uncertainty regions for the frequency response of an identified model.

Keywords: Identification; Prediction error methods; Covariance; Uncertainty

1 Introduction

A general linear, discrete-time model of a time-invariant system can be written

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k) + v(t) \tag{1}$$

Here $\{y(t)\}$ is the output, $\{u(t)\}$ the input, and $\{v(t)\}$ an additive disturbance, whose character we will discuss later. With (1) we associate the transfer function

$$G(e^{i\omega}) = \sum_{k=1}^{\infty} g(k)e^{-ik\omega}, \quad -\pi \le \omega \le \pi \tag{2}$$

In this paper we will discuss the statistical properties of transfer function models of the form (2) when the impulse response coefficients $g(k)$ are determined or estimated using measured data $Z_N = \{y(1), u(1), \dots, y(N), u(N)\}$. The identification method we will consider is the classical prediction error method, e.g., [4].

The main goal of the paper is to derive explicit expressions for the covariance matrix

$$P(\omega) = E \begin{bmatrix} \mathrm{Re}[\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})] \\ \mathrm{Im}[\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})] \end{bmatrix} \begin{bmatrix} \mathrm{Re}[\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})] \\ \mathrm{Im}[\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})] \end{bmatrix}^T \tag{3}$$

where $\hat{G}_N(e^{i\omega})$ denotes the identified model obtained using N measurements, that are asymptotic both in the number of observed data and in the model order. Similar results have previously been given in [3]. There, however, the focus was on expressions for

$$P_*(\omega) = E|\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})|^2 \ (= \mathrm{tr}\, P(\omega)) \tag{4}$$

which is real-valued and hence does not bring any information about the phase of $\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})$. With an explicit expression for the covariance matrix $P(\omega)$ defined in (3) we can construct confidence ellipsoids for $\hat{G}_N(e^{i\omega})$ in the complex plane which is useful, e.g., for robust control design and analysis.

2 Preliminaries

Notation and Definitions

The delay operator is denoted by q^{-1},

$$q^{-1}u(t) = u(t-1) \tag{5}$$

and the set of integers by $\mathbb{Z}$. The Kronecker delta function $\delta_{(\cdot)}$ is defined as

$$\delta_k = \begin{cases} 1, & k = 0 \\ 0, & k \ne 0 \end{cases} \tag{6}$$

As we shall work with signals that may contain deterministic components, we will consider generalized covariances and spectra of the form

$$R_{yu}(k) = \bar{E}y(t)u(t-k) \tag{7}$$

$$\Phi_{yu}(\omega) = \sum_{k=-\infty}^{\infty} R_{yu}(k)e^{-ik\omega} \tag{8}$$

In (7) the symbol $\bar{E}$ is defined as

$$\bar{E}f(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Ef(t) \tag{9}$$

A filter $F(q)$,

$$F(q) = \sum_{k=0}^{\infty} f(k)q^{-k} \tag{10}$$

is said to be *stable* if

$$\sum_{k=0}^{\infty} |f(k)| < \infty \tag{11}$$

A set of filters $F(q,\theta)$; $\theta \in D_{\mathcal{M}}$,

$$F(q,\theta) = \sum_{k=0}^{\infty} f(k,\theta)q^{-k} \tag{12}$$

is said to be *uniformly stable* if

$$\sum_{k=0}^{\infty} \sup_{\theta\in D_{\mathcal{M}}} |f(k,\theta)| < \infty \tag{13}$$

Basic Prediction Error Theory

Consider the set of models

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t); \quad \theta \in D_{\mathcal{M}} \subset \mathbb{R}^d \tag{14}$$

where $D_{\mathcal{M}}$ is a compact and connected subset of $\mathbb{R}^d$ ($d = \dim\theta$), and where

$$G(q,\theta) = \sum_{k=1}^{\infty} g(k,\theta)q^{-k} \tag{15}$$

$$H(q,\theta) = 1 + \sum_{k=1}^{\infty} h(k,\theta)q^{-k} \tag{16}$$

(In (14) $\{e(t)\}$ is supposed to be a sequence of independent, identically distributed random variables of zero means, variances λ_0, and bounded fourth order moments.) The corresponding one-step-ahead predictor is

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t) \tag{17}$$

The prediction error is

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = H^{-1}(q,\theta)(y(t) - G(q,\theta)u(t)) \tag{18}$$

The standard least-squares prediction error estimate is found as

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta) \tag{19}$$

$$V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N} \frac{1}{2}\varepsilon^2(t,\theta) \tag{20}$$

We will denote the corresponding transfer function estimates, $\hat{G}_N(q)$ and $\hat{H}_N(q)$, respectively; $\hat{G}_N(q) = G(q,\hat{\theta}_N)$ and $\hat{H}_N(q) = H(q,\hat{\theta}_N)$.

Under weak regularity conditions we have (see, e.g., [4])

$$\hat{\theta}_N \to \theta^* \text{ with probability 1 as } N \to \infty \tag{21}$$

$$\theta^* = \arg\min_{\theta \in D_{\mathcal{M}}} \bar{V}(\theta) \tag{22}$$

$$\bar{V}(\theta) = \bar{E}\frac{1}{2}\varepsilon^2(t,\theta) \tag{23}$$

Further,

$$\sqrt{N}(\hat{\theta}_N - \theta^*) \in AsN(0, P_\theta) \tag{24}$$

$$P_\theta = R^{-1}QR^{-1} \tag{25}$$

$$R = \bar{V}''(\theta^*) \tag{26}$$

$$Q = \lim_{N\to\infty} E\ N \cdot V_N'(\theta^*)(V_N'(\theta^*))^T \tag{27}$$

Here prime and double prime denotes differentiation once and twice, respectively, with respect to θ.

Asymptotic Variances for Identified Models

The result (24)-(27) states that asymptotically, as N tends to infinity, the estimate $\sqrt{N}\cdot\hat{\theta}_N$ will have a normal distribution with mean $\sqrt{N}\cdot\theta^*$ and covariance matrix

P_θ. Using the Taylor expansion

$$\begin{bmatrix} \mathrm{Re}[\hat{G}_N(e^{i\omega}) - G(e^{i\omega},\theta^*)] \\ \mathrm{Im}[\hat{G}_N(e^{i\omega}) - G(e^{i\omega},\theta^*)] \end{bmatrix} = \begin{bmatrix} \mathrm{Re}[G'_\theta(e^{i\omega},\theta^*)] \\ \mathrm{Im}[G'_\theta(e^{i\omega},\theta^*)] \end{bmatrix} (\hat{\theta}_N - \theta^*) + o(|\hat{\theta}_N - \theta^*|) \tag{28}$$

where $G'_\theta(e^{i\omega},\theta^*)$ is a $1 \times d$ dimensional matrix being the derivative of $G(e^{i\omega},\theta)$ with respect to θ evaluated at $\theta = \theta^*$, we thus have

$$\sqrt{N} \begin{bmatrix} \mathrm{Re}[\hat{G}_N(e^{i\omega}) - G(e^{i\omega},\theta^*)] \\ \mathrm{Im}[\hat{G}_N(e^{i\omega}) - G(e^{i\omega},\theta^*)] \end{bmatrix} \in AsN(0, P(\omega)) \tag{29}$$

with

$$P(\omega) = \begin{bmatrix} \mathrm{Re}[G'_\theta(e^{i\omega},\theta^*)] \\ \mathrm{Im}[G'_\theta(e^{i\omega},\theta^*)] \end{bmatrix} P_\theta \begin{bmatrix} \mathrm{Re}[G'_\theta(e^{i\omega},\theta^*)] \\ \mathrm{Im}[G'_\theta(e^{i\omega},\theta^*)] \end{bmatrix}^T \tag{30}$$

The matrix $P(\omega)$ in (30) gives an expression for the sought covariance matrix (3). However, evaluation of (30) is complicated and leads to intractable expressions, which limits the usefulness of the result. From, e.g., [3] we know that it is possible to compute variance expressions that are asymptotic in the model order, i.e., in the dimensionality of θ. This type of expressions tend to be simpler and, hence, easier to work with and to interpret than those obtainable using the above technique. Furthermore, despite the fact that these results are asymptotic in both the number of measurements and the model order, they frequently give reasonable approximations of the variance even for fixed model orders. In the next section we will derive the corresponding expressions for (3).

3 Main Result

Introduce

$$T(q,\theta) = \begin{bmatrix} G(q,\theta) \\ H(q,\theta) \end{bmatrix} \tag{31}$$

and define

$$\chi_0(t) = \begin{bmatrix} u(t) \\ e(t) \end{bmatrix} \tag{32}$$

The spectrum of $\{\chi_0(t)\}$ is

$$\Phi_{\chi_0}(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{ue}(-\omega) & \lambda_0 \end{bmatrix} \tag{33}$$

Using (31) and (32) we can rewrite the model (14) as

$$y(t) = T^T(q,\theta)\chi_0(t) \tag{34}$$

Suppose that the parameter vector θ can be decomposed so that

$$\theta = \begin{bmatrix}\theta_1^T, \ \theta_2^T, \ \dots, \ \theta_n^T\end{bmatrix}^T \quad \dim \theta_k = s \quad \dim \theta = n \cdot s \tag{35}$$

We shall call n the order of the model (34). Suppose also that

$$\frac{\partial}{\partial \theta_k} T(q, \theta) = q^{-k+1} \frac{\partial}{\partial \theta_1} T(q, \theta) \triangleq q^{-k+1} Z(q, \theta) \tag{36}$$

where the matrix $Z(q, \theta)$ in (36) is of size $2 \times s$.

It should be noted that most polynomial-type model structures, like ARMAX, Box-Jenkins, etc., satisfy this shift structure. Thus (36) is a rather weak assumption. Consider, e.g., an ARX model

$$y(t) = \frac{B(q)}{A(q)} u(t) + \frac{1}{A(q)} e(t) \tag{37}$$

where

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n} \tag{38}$$

$$B(q) = 1 + b_1 q^{-1} + \cdots + b_n q^{-n} \tag{39}$$

In this case we have $\theta_k = \begin{bmatrix} a_k & b_k \end{bmatrix}^T$ and

$$Z(q, \theta) = q^{-1} \begin{bmatrix} -\frac{B(q)}{A^2(q)} & \frac{1}{A(q)} \\ -\frac{1}{A^2(q)} & 0 \end{bmatrix} \tag{40}$$

From (36) it follows that the $2 \times d$ dimensional matrix $T'_\theta(e^{i\omega}, \theta)$, being the derivative of $T(e^{i\omega}, \theta)$ with respect to θ, can be written

$$T'_\theta(e^{i\omega}, \theta) = e^{i\omega} Z(e^{i\omega}, \theta) W_n(e^{i\omega}) \tag{41}$$

where

$$W_n(e^{i\omega}) = \begin{bmatrix} e^{-i\omega} I & e^{-i2\omega} I & \cdots & e^{-in\omega} I \end{bmatrix} \tag{42}$$

with I being an $s \times s$ identity matrix. The following lemma, which also is of independent interest, will be used in the proof of the main result, Theorem 2, below.

Lemma 1 *Let $W_n(e^{i\omega})$ be defined by (42) and let $\{w(t)\}$ be an s-dimensional process with invertible spectrum $\Phi_w(\omega)$. Then*

$$\begin{aligned} &\lim_{n \to \infty} \frac{1}{n} W_n(e^{i\omega_1}) \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} W_n^T(e^{-i\xi}) \Phi_w(-\xi) W_n(e^{i\xi}) d\xi \right]^{-1} W_n^T(e^{-i\omega_2}) \\ &= \left[\Phi_w(-\omega_1) \right]^{-1} \delta_{(\omega_1 - \omega_2) \bmod 2\pi} \end{aligned} \tag{43}$$

Proof *From Lemma 4.3 in [7] (see also [5], Lemma 4.2) we have that*

$$\lim_{n\to\infty}\frac{1}{n}W_n(e^{i\omega_1})\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}W_n^T(e^{i\xi})\Phi_w(\xi)W_n(e^{-i\xi})d\xi\right]^{-T}W_n^T(e^{-i\omega_2})$$
$$= \left[\Phi_w(\omega_1)\right]^{-T}\delta_{(\omega_1-\omega_2)} \tag{44}$$

(Note especially the transposes, which are due to different definitions of correlation functions and spectra than the ones used here.) The result (43) *follows since* $\Phi_w^T(\omega) = \Phi_w(-\omega)$ *and* $W_n(e^{i(\omega+2\pi k)}) = W_n(e^{i\omega})\ \forall\ k \in \mathbb{Z}$. □ □

For the proof of Theorem 2 below we additionally need a number of technical assumptions which we now list. Let

$$\theta^*(n) = \arg\min_{\theta\in D_{\mathcal{M}}}\bar{V}(\theta) \tag{45}$$

(If the minimum is not unique, let $\theta^*(n)$ denote any, arbitrarily chosen minimizing element.) The argument n is added to emphasize that the minimization is carried out over nth order models. Define

$$\hat{\theta}_N(n,\delta) = \arg\min_{\theta\in D_{\mathcal{M}}}V_N(\theta,\delta,n) \tag{46}$$

$$V_N(\theta,\delta,n) = \frac{1}{2}\left[\frac{1}{N}\sum_{t=1}^{N}\varepsilon^2(t,\theta) + \delta|\theta-\theta^*(n)|^2\right] \tag{47}$$

where δ is a regularization parameter, helping us to select a unique minimizing element in (46) in case $\delta = 0$ leads to nonunique minima.

Assume that the true system can be described by

$$y(t) = G_0(q)u(t) + v(t), \quad v(t) = H_0(q)e(t) \tag{48}$$

where $\{e(t)\}$ is a sequence of independent, identically distributed random variables with zero means, variances λ_0, and bounded fourth order moments, and where $G_0(q)$ is stable and strictly causal and $H_0(q)$ is monic and inversely stable. From (48) it follows that the spectrum of the additive disturbance $\{v(t)\}$ is

$$\Phi_v(\omega) = \lambda_0|H_0(e^{i\omega})|^2 \tag{49}$$

It will be assumed that there exists a $\theta_0 \in D_{\mathcal{M}}$ such that

$$G(q,\theta_0) = G_0(q) \quad\text{and}\quad H(q,\theta_0) = H_0(q) \tag{50}$$

Further, suppose that the predictor filters

$$H^{-1}(q,\theta)G(q,\theta) \ \text{ and }\ H^{-1}(q,\theta) \tag{51}$$

are uniformly stable for $\theta \in D_{\mathcal{M}}$ along with their first, second, and third order derivatives.

Let

$$T_n^*(e^{i\omega}) = T(e^{i\omega}, \theta^*(n)) \tag{52}$$

$$\hat{T}_N(e^{i\omega}, n, \delta) = T(e^{i\omega}, \hat{\theta}_N(n, \delta)) \tag{53}$$

$$T_0(e^{i\omega}) = T(e^{i\omega}, \theta_0) \tag{54}$$

Assume that

$$\lim_{n\to\infty} n^2 E[\varepsilon(t, \theta^*(n)) - e(t)]^2 = 0 \tag{55}$$

which implies that $T_n^*(e^{i\omega})$ tends to $T_0(e^{i\omega})$ as n tends to infinity. Let $Z(e^{i\omega}, \theta)$ defined in (36) be denoted by $Z_0(e^{i\omega})$ when evaluated for $\theta = \theta_0$. Assume that

$$Z_0(e^{i\omega}) Z_0^T(e^{-i\omega}) \tag{56}$$

is invertible. Further assume that $R_u(k)$ and $R_{ue}(k)$ exist and that $R_{ue}(k) = 0$, $k < 0$. Finally, assume that

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{t=1}^{N} E\left[\frac{d}{d\theta}\varepsilon^2(t, \theta^*(n))\right] = 0 \quad (n \text{ fixed}) \tag{57}$$

We can now state the main result of the paper.

Theorem 2 *Consider the estimate* $\hat{T}_N(e^{i\omega}, n, \delta)$ *under the assumptions* (36) *and* (45)-(57). *Then*

$$\sqrt{N} \begin{bmatrix} \mathrm{Re}[\hat{T}_N(e^{i\omega}, n, \delta) - T_n^*(e^{i\omega})] \\ \mathrm{Im}[\hat{T}_N(e^{i\omega}, n, \delta) - T_n^*(e^{i\omega})] \end{bmatrix} \in AsN(0, \bar{P}(\omega, n, \delta)) \tag{58}$$
$$[\textit{as } N \to \infty \textit{ for fixed } n,\ \delta]$$

where

$$\lim_{\delta\to 0} \lim_{n\to\infty} \frac{1}{n} \bar{P}(\omega, n, \delta) = \begin{cases} \frac{1}{2}\Phi_v(\omega) \begin{bmatrix} \mathrm{Re}[\Phi_{\chi_0}^{-1}(\omega)] & \mathrm{Im}[\Phi_{\chi_0}^{-1}(\omega)] \\ -\mathrm{Im}[\Phi_{\chi_0}^{-1}(\omega)] & \mathrm{Re}[\Phi_{\chi_0}^{-1}(\omega)] \end{bmatrix}, & \textit{if } \omega \bmod \pi \neq 0 \\ \Phi_v(\omega) \begin{bmatrix} \mathrm{Re}[\Phi_{\chi_0}^{-1}(\omega)] & 0 \\ 0 & 0 \end{bmatrix}, & \textit{if } \omega \bmod \pi = 0 \end{cases} \tag{59}$$

The proof is given in Appendix A.

From (58)-(59) we conclude that, for $\omega \bmod \pi \neq 0$, the random variable

$$\sqrt{N}(\hat{T}_N(e^{i\omega}, n, \delta) - T_n^*(e^{i\omega})) \tag{60}$$

asymptotically has a *complex* normal distribution (e.g., [1]) with covariance matrix $\bar{P}_*(\omega, n, \delta)$ satisfying

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\bar{P}_*(\omega, n, \delta) = \Phi_v(\omega)\Phi_{\chi_0}^{-1}(-\omega) \tag{61}$$

Compare also with the results in [3]. (In connection to this the author would like to point out that the matrix on the right hand side of equation (3.22) in [3] should be transposed.) For $\omega \bmod \pi = 0$ this is no longer true, since then the diagonal blocks of

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\bar{P}(\omega, n, \delta) \tag{62}$$

are not equal. That the covariance for the imaginary part is zero in this case is very natural, since the transfer function is real-valued for $\omega \bmod \pi = 0$.

Further, with basically only notational differences the result in Theorem 2 also holds for multivariable models. The extension of the results in [3] to the multivariable case was given in [8]. We can also prove the result for polynomial models where the different polynomials are of unequal orders, or if other criteria than the quadratic one is used. In open loop the situation simplifies and, e.g., to prove the corresponding results for the model $\hat{G}_N(e^{i\omega}, n, \delta)$ in this case the conditions in Theorem 2 can be relaxed so that a fixed, arbitrary noise model $H(q, \theta) = H_*(q)$ can be used. See [3] for all this.

Let us return to the result in Theorem 2. A more explicit expression for (59) can be obtained if we note that

$$\Phi_{\chi_0}^{-1}(\omega) = \begin{bmatrix} 1/\Phi_u^r(\omega) & -\Phi_{ue}(\omega)/(\lambda_0\Phi_u^r(\omega)) \\ -\Phi_{eu}(\omega)/(\lambda_0\Phi_u^r(\omega)) & 1/\Phi_e^r(\omega) \end{bmatrix} \tag{63}$$

$$\Phi_u^r(\omega) = \Phi_u(\omega) - |\Phi_{ue}(\omega)|^2/\lambda_0 \tag{64}$$

$$\Phi_e^r(\omega) = \lambda_0 - |\Phi_{ue}(\omega)|^2/\Phi_u(\omega) \tag{65}$$

$\Phi_u^r(\omega)$ can be interpreted as the "noise-free" part of the input spectrum, i.e., that part of the total input spectrum $\Phi_u(\omega)$ that does not originate from $\{e(t)\}$ but from some external reference or set-point signal. In open loop we have $\Phi_{ue}(\omega) = 0$, which, e.g., implies that $\Phi_u^r(\omega) = \Phi_u(\omega)$, and the expression for $\Phi_{\chi_0}^{-1}(\omega)$ simplifies to

$$\Phi_{\chi_0}^{-1}(\omega) = \begin{bmatrix} 1/\Phi_u(\omega) & 0 \\ 0 & 1/\lambda_0 \end{bmatrix} \tag{66}$$

If we return to the general, closed loop situation we see from (63) that

$$\operatorname{Re}[\Phi_{\chi_0}^{-1}(\omega)] = \begin{bmatrix} 1/\Phi_u^r(\omega) & -\operatorname{Re}[\Phi_{ue}(\omega)/(\lambda_0\Phi_u^r(\omega))] \\ -\operatorname{Re}[\Phi_{eu}(\omega)/(\lambda_0\Phi_u^r(\omega))] & 1/\Phi_e^r(\omega) \end{bmatrix} \tag{67}$$

$$\operatorname{Im}[\Phi_{\chi_0}^{-1}(\omega)] = \begin{bmatrix} 0 & -\operatorname{Im}[\Phi_{ue}(\omega)/(\lambda_0\Phi_u^r(\omega))] \\ -\operatorname{Im}[\Phi_{eu}(\omega)/(\lambda_0\Phi_u^r(\omega))] & 0 \end{bmatrix} \tag{68}$$

Using (67) and (68) we can thus easily prove the following consequence of Theorem 2, dealing with the asymptotic distribution of the estimate $\hat{G}_N(e^{i\omega}, n, \delta)$.

Corollary 3 *Consider the situation in Theorem 2. Then*

$$\sqrt{N}\begin{bmatrix} \operatorname{Re}[\hat{G}_N(e^{i\omega}, n, \delta) - G_n^*(e^{i\omega})] \\ \operatorname{Im}[\hat{G}_N(e^{i\omega}, n, \delta) - G_n^*(e^{i\omega})] \end{bmatrix} \in AsN(0, P(\omega, n, \delta)) \tag{69}$$

$$[\text{as } N \to \infty \text{ for fixed } n,\, \delta]$$

where

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}P(\omega, n, \delta) = \begin{cases} \frac{1}{2}\Phi_v(\omega)\begin{bmatrix} 1/\Phi_u^r(\omega) & 0 \\ 0 & 1/\Phi_u^r(\omega) \end{bmatrix}, & \text{if } \omega \bmod \pi \neq 0 \\ \Phi_v(\omega)\begin{bmatrix} 1/\Phi_u^r(\omega) & 0 \\ 0 & 0 \end{bmatrix}, & \text{if } \omega \bmod \pi = 0 \end{cases} \tag{70}$$

From (69)-(70) we see that, for $\omega \bmod \pi \neq 0$, the random variable

$$\sqrt{N}(\hat{G}_N(e^{i\omega}, n, \delta) - G_n^*(e^{i\omega})) \tag{71}$$

asymptotically has a complex normal distribution with covariance matrix $P_*(\omega, n, \delta)$ satisfying

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}P_*(\omega, n, \delta) = \frac{\Phi_v(\omega)}{\Phi_u^r(\omega)} \tag{72}$$

This does not hold for $\omega \bmod \pi = 0$, but for all ω we nevertheless have that

$$\operatorname{tr}\left(\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}P(\omega, n, \delta)\right) = \frac{\Phi_v(\omega)}{\Phi_u^r(\omega)} \tag{73}$$

which ties in nicely with the results in [3], cf. (3) and (4).

An intuitive, but not formally correct, interpretation of the result (69)-(70) is that it gives the following convenient expression for the covariance matrix (3) (for the case $\omega \bmod \pi \neq 0$):

$$P(\omega) \approx \frac{n}{N}\frac{1}{2}\Phi_v(\omega)\begin{bmatrix} 1/\Phi_u^r(\omega) & 0 \\ 0 & 1/\Phi_u^r(\omega) \end{bmatrix} \quad \text{as } N,\ n \to \infty \tag{74}$$

From (74) we see that asymptotically, as both n and N tend to infinity, the real and imaginary parts of the G-estimate are uncorrelated with equal variance proportional to the number-of-parameters-to-number-of-measurements ratio (n/N) and to the noise-to-signal ratio ($\Phi_v(\omega)/\Phi_u^r(\omega)$).

4 Example

If we assume that the random variable

$$\begin{bmatrix} \mathrm{Re}[\hat{G}_N(e^{i\omega})] \\ \mathrm{Im}[\hat{G}_N(e^{i\omega})] \end{bmatrix} \tag{75}$$

has a normal distribution with covariance matrix $P(\omega)$, we may compute an $\alpha\%$ confidence ellipsoid for $\hat{G}_N(e^{i\omega})$ at a particular frequency ω_k, $\omega_k \bmod \pi \neq 0$, as the set of all $G(e^{i\omega_k})$ such that

$$\begin{bmatrix} \mathrm{Re}[G(e^{i\omega_k}) - \hat{G}_N(e^{i\omega_k})] \\ \mathrm{Im}[G(e^{i\omega_k}) - \hat{G}_N(e^{i\omega_k})] \end{bmatrix} P^{-1}(\omega_k) \begin{bmatrix} \mathrm{Re}[G(e^{i\omega_k}) - \hat{G}_N(e^{i\omega_k})] \\ \mathrm{Im}[G(e^{i\omega_k}) - \hat{G}_N(e^{i\omega_k})] \end{bmatrix}^T \leq C_\alpha \tag{76}$$

where C_α is defined through $\mathrm{Prob}(X \leq C_\alpha) = \alpha\%$, $X \in \chi^2(2)$. In this example we will construct such confidence ellipsoids for the frequency response of an identified 15th order ARX model of the system

$$y(t) = \frac{0.1589q^{-1} + 0.3042q^{-2} + 0.0323q^{-3}}{1 - 1.3219q^{-1} + 0.4618q^{-2} - 0.0408q^{-3}} u(t) + e(t) \tag{77}$$

for 40 distinct frequencies ω_k: 0 and 39 logarithmically spaced frequencies between 0.01 and π. (For $\omega_k \bmod \pi = 0$ the confidence region becomes an interval on the real axis. The details are omitted.) For the identification, a data set consisting of $N = 1000$ samples of $y(t)$ and $u(t)$ is generated by simulating the system using a unit binary random signal $\{u(t)\}$ and a uniformly $[-1, 1]$ distributed random signal $\{e(t)\}$, independent of $\{u(t)\}$.

We remark that a high-order ARX model can approximate any linear system arbitrarily well if the model order is sufficiently large, see, e.g., [4]. Hence we can quite safely neglect the bias error in this example and use the covariance matrix $P(\omega)$ to describe the model uncertainty. However, since we do not know $P(\omega)$ we have to replace it by some estimate $\hat{P}(\omega)$. Here we will first consider (29)-(30) and replace $P(\omega)$ by

$$\hat{P}(\omega) = \begin{bmatrix} \mathrm{Re}[G'_\theta(e^{i\omega}, \hat{\theta}_N)] \\ \mathrm{Im}[G'_\theta(e^{i\omega}, \hat{\theta}_N)] \end{bmatrix} \frac{1}{N} \hat{P}_\theta \begin{bmatrix} \mathrm{Re}[G'_\theta(e^{i\omega}, \hat{\theta}_N)] \\ \mathrm{Im}[G'_\theta(e^{i\omega}, \hat{\theta}_N)] \end{bmatrix}^T \tag{78}$$

where $\hat{P}_\theta$ denotes the estimated parameter covariance delivered by the identification routine. The resulting 95% confidence ellipsoids for $\hat{G}_N(e^{i\omega_k})$, $k \in [1, \ldots, 40]$, are shown in Figure 1. The solid line represents the true system.

The results in Figure 1 should be compared with the confidence ellipsoids shown in Figure 2 obtained using the following estimate of $P(\omega)$ (for $\omega \bmod \pi \neq 0$) derived from (74):

$$\hat{P}(\omega) = \frac{n}{N} \frac{1}{2} \hat{\lambda}_N |\hat{H}_N(e^{i\omega})|^2 \begin{bmatrix} 1/\Phi_u(\omega) & 0 \\ 0 & 1/\Phi_u(\omega) \end{bmatrix} \tag{79}$$

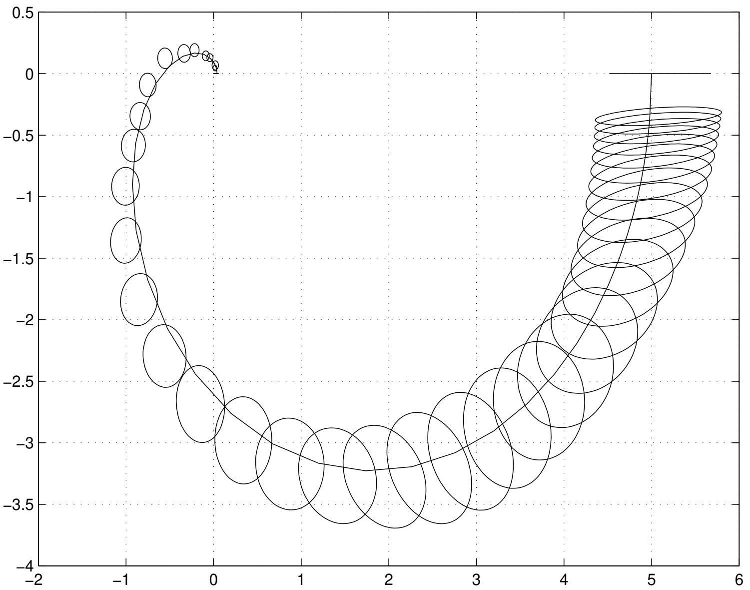

Figure 1 *Confidence ellipsoids obtained using expression* (78).

where $\hat{\lambda}_N$ denotes the variance of the prediction error $\varepsilon(t, \hat{\theta}_N)$. As can be seen, the confidence regions obtained using the expressions (78) and (79) (or (30) and (74), respectively) for the covariance matrix $P(\omega)$ are quite similar, except at low frequencies, excluding zero, where the asymptotic expression (79) over-estimates the variance of the imaginary part of $\hat{G}_N(e^{i\omega})$. (It should be noted that the expression (79) gives circles, since the real and imaginary parts asymptotically are uncorrelated with equal variance.) However, viewed over all frequencies, the discrepancy is surprisingly small considering that the simpler expression (79) assumes that $N \gg n \to \infty$, while $n = 15$ in this example.

To verify the correctness of the confidence regions in Figures 1 and 2 we also performed $M = 1000$ Monte Carlo simulations and derived empirical confidence regions by computing, for each frequency ω_k, the convex hull of the 950 ($= 95\% \cdot M$) points among the 1000 Monte Carlo estimates $\hat{G}_N^{(i)}(e^{i\omega_k})$ that were closest to the sample mean

$$\bar{G}_M(e^{i\omega_k}) = \frac{1}{M} \sum_{i=1}^{M} \hat{G}^{(i)}(e^{i\omega_k}) \tag{80}$$

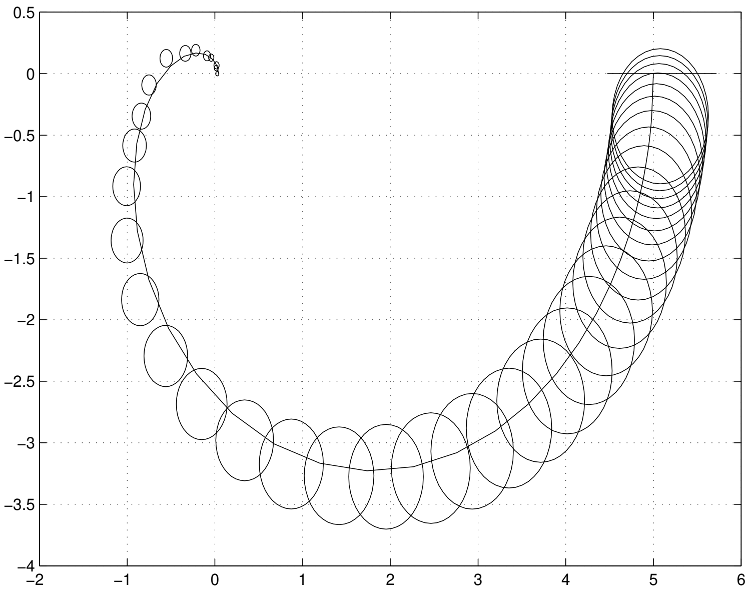

Figure 2 *Confidence ellipsoids obtained using expression* (79).

measured in the norm given by the inverse of the sample covariance matrix

$$\hat{P}_M(e^{i\omega_k}) = \frac{1}{M}\sum_{i=1}^{M}\begin{bmatrix}\mathrm{Re}[\hat{G}^{(i)}(e^{i\omega_k}) - \bar{G}(e^{i\omega_k})] \\ \mathrm{Im}[\hat{G}^{(i)}(e^{i\omega_k}) - \bar{G}(e^{i\omega_k})]\end{bmatrix}\begin{bmatrix}\mathrm{Re}[\hat{G}^{(i)}(e^{i\omega_k}) - \bar{G}(e^{i\omega_k})] \\ \mathrm{Im}[\hat{G}^{(i)}(e^{i\omega_k}) - \bar{G}(e^{i\omega_k})]\end{bmatrix}^T \quad (81)$$

The results thus obtained are shown in Figure 3. Clearly, the confidence ellipsoids in Figure 1 are rather good approximations of the "true" uncertainty regions in Figure 3. The plot in Figure 2 deviates more from the "true" one in Figure 3, but note that if we were to fit magnitude bounds on the uncertainty all three cases would give similar results. In particular, this shows that the asymptotic expression (74) can be relevant also with fixed model orders.

We can also mention that variants of this example have previously been used in [6] and [2] to illustrate the performances of methods for model error bounding. In [6] a worst-case deterministic approach is studied and in [2] a probabilistic method is considered. The approach above can be seen as a "trivial" variant of the probabilistic error bounding methods; "trivial", since we only use the available covariance information to compute the uncertainty regions. However, as we have seen, this approach can still lead to realistic, nonconservative uncertainty bounds at a low computational cost.

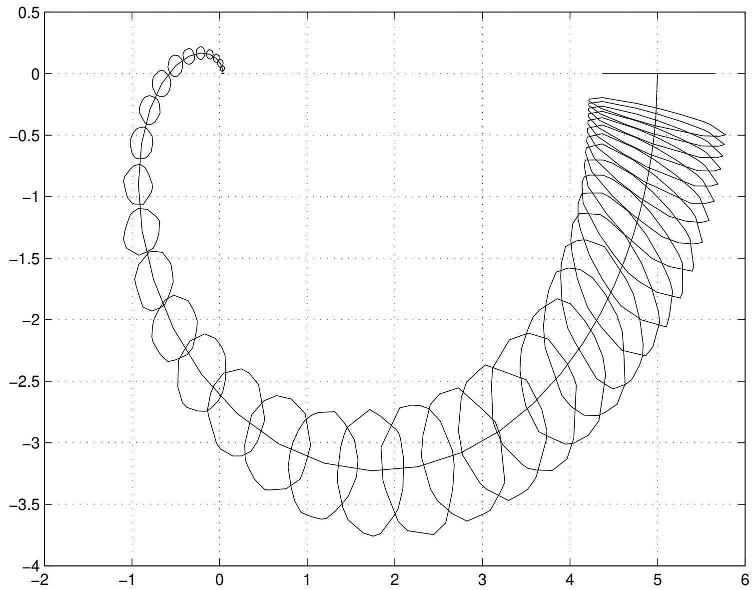

Figure 3 *Confidence regions obtained using Monte Carlo simulations.*

5 Summary

We have derived asymptotic expressions for the covariance of the real and imaginary parts of identified transfer function models that are asymptotic both in the number of measurements and in the model orders. The relevance of these expressions also with fixed model orders has been shown in a Monte Carlo simulation where confidence regions for an identified model are computed.

Acknowledgment

The author gratefully acknowledges F. Tjärnström who provided the Matlab code with which the figures were generated.

A Proof of Theorem 2

Let $\psi(t,\theta)$ of dimension $d \times 1$ be the gradient of the predictor $\hat{y}(t|\theta)$ and define

$$M_n(\theta) = \bar{E}\psi(t,\theta)\psi^T(t,\theta) \tag{A.1}$$

If we introduce

$$\chi(t,\theta) = \begin{bmatrix} u(t) \\ \varepsilon(t,\theta) \end{bmatrix} \tag{A.2}$$

and

$$\zeta(t,\theta) = H^{-1}(q,\theta)Z^T(q,\theta)\chi(t,\theta) \tag{A.3}$$

we can also write $\psi(t,\theta)$ as

$$\psi(t,\theta) = W_n^T(q)\zeta(t,\theta) \tag{A.4}$$

($W_n(q)$ is the time-domain counterpart of (42).) From Lemma A.4 in [3] we have

$$\sqrt{N}(\hat{\theta}_N(n,\delta) - \theta^*(n)) \in AsN(0, R(n,\delta)) \tag{A.5}$$

$$R(n,\delta) = \lambda_0\big[M_n(\theta^*(n)) + \delta I\big]^{-1} - \lambda_0\delta\big[M_n(\theta^*(n)) + \delta I\big]^{-2} + C_\delta \cdot \tilde{\rho}_n \tag{A.6}$$

where C_δ is a constant and $\|\tilde{\rho}_n\| \to 0$ as $n \to \infty$. Considering the Taylor expansion

$$\begin{bmatrix} \mathrm{Re}[\hat{T}_N(e^{i\omega},n,\delta) - T_n^*(e^{i\omega})] \\ \mathrm{Im}[\hat{T}_N(e^{i\omega},n,\delta) - T_n^*(e^{i\omega})] \end{bmatrix} = \begin{bmatrix} \mathrm{Re}[T'_\theta(e^{i\omega},\theta^*(n))] \\ \mathrm{Im}[T'_\theta(e^{i\omega},\theta^*(n))] \end{bmatrix} (\hat{\theta}_N(n,\delta) - \theta^*(n)) \\ + o(|\hat{\theta}_N(n,\delta) - \theta^*(n)|) \tag{A.7}$$

we thus see that the random variable

$$\sqrt{N} \begin{bmatrix} \mathrm{Re}[\hat{T}_N(e^{i\omega},n,\delta) - T_n^*(e^{i\omega})] \\ \mathrm{Im}[\hat{T}_N(e^{i\omega},n,\delta) - T_n^*(e^{i\omega})] \end{bmatrix} \tag{A.8}$$

will have asymptotic normal distribution with covariance matrix

$$\bar{P}(\omega,n,\delta) = \begin{bmatrix} \mathrm{Re}[T'_\theta(e^{i\omega},\theta^*(n))] \\ \mathrm{Im}[T'_\theta(e^{i\omega},\theta^*(n))] \end{bmatrix} R(n,\delta) \begin{bmatrix} \mathrm{Re}[T'_\theta(e^{i\omega},\theta^*(n))] \\ \mathrm{Im}[T'_\theta(e^{i\omega},\theta^*(n))] \end{bmatrix}^T \tag{A.9}$$

Denote the upper left block of $\bar{P}(\omega,n,\delta)$ by $\bar{P}_{1,1}(\omega,n,\delta)$:

$$\begin{aligned} \bar{P}_{1,1}(\omega,n,\delta) &= \mathrm{Re}[T'_\theta(e^{i\omega},\theta^*(n))]R(n,\delta)(\mathrm{Re}[T'_\theta(e^{i\omega},\theta^*(n))])^T \\ &= \frac{1}{2}\big[T'_\theta(e^{i\omega},\theta^*(n)) + T'_\theta(e^{-i\omega},\theta^*(n))\big]R(n,\delta) \times \\ &\times \frac{1}{2}\big[T'_\theta(e^{i\omega},\theta^*(n)) + T'_\theta(e^{-i\omega},\theta^*(n))\big]^T \end{aligned} \tag{A.10}$$

In view of (41) we have

$$
\begin{aligned}
&\frac{1}{n}\bar{P}_{1,1}(\omega,n,\delta)\\
&= \frac{1}{n}\frac{1}{4}(e^{i\omega}Z(e^{i\omega},\theta^*(n))W_n(e^{i\omega}) + e^{-i\omega}Z(e^{-i\omega},\theta^*(n))W_n(e^{-i\omega}))R(n,\delta)\times\\
&\times(W_n^T(e^{i\omega})Z^T(e^{i\omega},\theta^*(n))e^{i\omega} + W_n^T(e^{-i\omega})Z^T(e^{-i\omega},\theta^*(n))e^{-i\omega})\\
&= \frac{1}{n}\frac{1}{4}\big[e^{i\omega}Z(e^{i\omega},\theta^*(n))W_n(e^{i\omega})R(n,\delta)W_n^T(e^{i\omega})Z^T(e^{i\omega},\theta^*(n))e^{i\omega}\\
&+ e^{i\omega}Z(e^{i\omega},\theta^*(n))W_n(e^{i\omega})R(n,\delta)W_n^T(e^{-i\omega})Z^T(e^{-i\omega},\theta^*(n))e^{-i\omega}\\
&+ e^{-i\omega}Z(e^{-i\omega},\theta^*(n))W_n(e^{-i\omega})R(n,\delta)W_n^T(e^{i\omega})Z^T(e^{i\omega},\theta^*(n))e^{i\omega}\\
&+ e^{-i\omega}Z(e^{-i\omega},\theta^*(n))W_n(e^{-i\omega})R(n,\delta)W_n^T(e^{-i\omega})Z^T(e^{-i\omega},\theta^*(n))e^{-i\omega}\big]
\end{aligned}
\tag{A.11}
$$

Together with (A.6) this shows that we need to consider limits like

$$
\lim_{n\to\infty}\frac{1}{n}W_n(e^{i\omega_1})\lambda_0\big[M_n(\theta^*(n)) + \delta I\big]^{-1}W_n^T(e^{-i\omega_2})
\tag{A.12}
$$

and

$$
\lim_{n\to\infty}\frac{1}{n}W_n(e^{i\omega_1})\lambda_0\delta\big[M_n(\theta^*(n)) + \delta I\big]^{-2}W_n^T(e^{-i\omega_2})
\tag{A.13}
$$

Using Lemma 1 together with (A.1)-(A.4), where we note that $\varepsilon(t,\theta^*(n)) \to e(t)$ as $n\to\infty$, we have

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}W_n(e^{i\omega_1})\lambda_0\big[M_n(\theta^*(n)) + \delta I\big]^{-1}W_n(e^{-i\omega_2})\\
&= \lambda_0\big[Z_0^T(e^{-i\omega_1})S(-\omega_1)Z_0(e^{i\omega_1}) + \delta I\big]^{-1}\delta_{(\omega_1-\omega_2)\bmod 2\pi}
\end{aligned}
\tag{A.14}
$$

where

$$
S(\omega) = \frac{\Phi_{\chi_0}(\omega)}{|H_0(e^{i\omega})|^2}
\tag{A.15}
$$

Similarly,

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}W_n(e^{i\omega_1})\lambda_0\delta\big[M_n(\theta^*(n)) + \delta I\big]^{-2}W_n(e^{-i\omega_2})\\
&= \lambda_0\delta\big[Z_0^T(e^{-i\omega_1})S(-\omega_1)Z_0(e^{i\omega_1}) + \delta I\big]^{-2}\delta_{(\omega_1-\omega_2)\bmod 2\pi}
\end{aligned}
\tag{A.16}
$$

Now, consider the terms in (A.11) that originate from the first term in $R(n,\delta)$ (cf. (A.6)). Using the result (A.14) repeated times will show that as n tends to infinity these terms tend to

$$
\begin{aligned}
&\frac{1}{4}\Big[e^{i\omega}Z_0(e^{i\omega})\lambda_0\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{i\omega})e^{i\omega}\delta_{\omega \bmod \pi}\\
&+e^{i\omega}Z_0(e^{i\omega})\lambda_0\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega})e^{-i\omega}\\
&+e^{-i\omega}Z_0(e^{-i\omega})\lambda_0\big[Z_0^T(e^{i\omega})S(\omega)Z_0(e^{-i\omega})+\delta I\big]^{-1}Z_0^T(e^{i\omega})e^{i\omega}\\
&+e^{-i\omega}Z_0(e^{-i\omega})\lambda_0\big[Z_0^T(e^{i\omega})S(\omega)Z_0(e^{-i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega})e^{-i\omega}\delta_{\omega \bmod \pi}\Big]\\
&=\frac{1}{2}\Big[\frac{1}{2}\big\{Z_0(e^{-i\omega})\lambda_0\big[Z_0^T(e^{i\omega})S(\omega)Z_0(e^{-i\omega})+\delta I\big]^{-1}Z_0^T(e^{i\omega})\\
&+Z_0(e^{i\omega})\lambda_0\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega})\big\}\\
&+\frac{1}{2}\big\{Z_0(e^{i\omega})\lambda_0\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{i\omega})\\
&+Z_0(e^{-i\omega})\lambda_0\big[Z_0^T(e^{i\omega})S(\omega)Z_0(e^{-i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega})\big\}\delta_{\omega \bmod \pi}\Big]
\end{aligned}
\tag{A.17}
$$

If we apply the matrix inversion lemma

$$[A+BCD]^{-1}=A^{-1}-A^{-1}B[C^{-1}+DA^{-1}B]^{-1}DA^{-1} \tag{A.18}$$

to, e.g.,

$$\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega}) \tag{A.19}$$

we see that

$$
\begin{aligned}
&\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega})\\
&=\big[\delta^{-1}I-\delta^{-1}Z_0^T(e^{-i\omega})[\delta S^{-1}(-\omega)+Z_0(e^{i\omega})Z_0^T(e^{-i\omega})]^{-1}Z_0(e^{i\omega})\big]Z_0^T(e^{-i\omega})\\
&=\delta^{-1}Z_0^T(e^{-i\omega})-\delta^{-1}Z_0^T(e^{-i\omega})\big[\delta(Z_0(e^{i\omega})Z_0^T(e^{-i\omega}))^{-1}S^{-1}(-\omega)+I\big]^{-1}\\
&\approx\delta^{-1}Z_0^T(e^{-i\omega})-\delta^{-1}Z_0^T(e^{-i\omega})\big[I-\delta(Z_0(e^{i\omega})Z_0^T(e^{-i\omega}))^{-1}S^{-1}(-\omega)\big]\\
&=Z_0^T(e^{-i\omega})(Z_0(e^{i\omega})Z_0^T(e^{-i\omega}))^{-1}S^{-1}(-\omega)
\end{aligned}
\tag{A.20}
$$

where the approximate inequality in the third step holds for small δ. We thus see that, e.g.,

$$
\begin{aligned}
&Z_0(e^{i\omega})\lambda_0\big[Z_0^T(e^{-i\omega})S(-\omega)Z_0(e^{i\omega})+\delta I\big]^{-1}Z_0^T(e^{-i\omega})\\
&\qquad\rightarrow\lambda_0 S^{-1}(-\omega)=\Phi_v(\omega)\Phi_{\chi_0}^{-1}(-\omega) \text{ as } \delta\rightarrow\infty
\end{aligned}
\tag{A.21}
$$

It follows that the limit of (A.17) as δ tends zero is

$$
\begin{aligned}
&\frac{1}{2}\Phi_v(\omega)\left[\frac{1}{2}\big\{\Phi_{\chi_0}^{-1}(\omega)+\Phi_{\chi_0}^{-1}(-\omega)\big\}+\frac{1}{2}\big\{\Phi_{\chi_0}^{-1}(\omega)+\Phi_{\chi_0}^{-1}(-\omega)\big\}\delta_{\omega \bmod \pi}\right]\\
&=\frac{1}{2}\Phi_v(\omega)\left[\mathrm{Re}[\Phi_{\chi_0}^{-1}(\omega)]+\mathrm{Re}[\Phi_{\chi_0}^{-1}(\omega)]\delta_{\omega \bmod \pi}\right]
\end{aligned}
\tag{A.22}
$$

Next, we may use similar arguments as in [3] to prove that the terms in (A.11) of the form (A.16), originating from the second term in $R(n,\delta)$, and the terms in (A.11) originating from the third terms in $R(n,\delta)$ all tend to zero as δ tends to zero. In summary we thus have that

$$\begin{aligned}&\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\operatorname{Re}[T'_\theta(e^{i\omega},\theta^*(n))]R(n,\delta)(\operatorname{Re}[T'_\theta(e^{i\omega},\theta^*(n))])^T\\&=\frac{1}{2}\Phi_v(\omega)\left[\operatorname{Re}[\Phi_{\chi_0}^{-1}(\omega)]+\operatorname{Re}[\Phi_{\chi_0}^{-1}(\omega)]\delta_{\omega \bmod \pi}\right]\end{aligned} \tag{A.23}$$

Analogous calculations will also show that

$$\begin{aligned}&\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\operatorname{Re}[T'_\theta(e^{i\omega},\theta^*(n))]R(n,\delta)(\operatorname{Im}[T'_\theta(e^{i\omega},\theta^*(n))])^T\\&=\frac{1}{2}\Phi_v(\omega)\left[\operatorname{Im}[\Phi_{\chi_0}^{-1}(\omega)]-\operatorname{Im}[\Phi_{\chi_0}^{-1}(\omega)]\delta_{\omega \bmod \pi}\right]\end{aligned} \tag{A.24}$$

and

$$\begin{aligned}&\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\operatorname{Im}[T'_\theta(e^{i\omega},\theta^*(n))]R(n,\delta)(\operatorname{Im}[T'_\theta(e^{i\omega},\theta^*(n))])^T\\&=\frac{1}{2}\Phi_v(\omega)\left[\operatorname{Re}[\Phi_{\chi_0}^{-1}(\omega)]-\operatorname{Re}[\Phi_{\chi_0}^{-1}(\omega)]\delta_{\omega \bmod \pi}\right]\end{aligned} \tag{A.25}$$

which ends the proof. □

References

[1] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Holden-Day, 1981.

[2] R. G. Hakvoort and P. M. J. Van den Hof. Identification of probabilistic uncertainty regions by explicit evaluation of bias and variance errors. *IEEE Transactions on Automatic Control*, 42:1516–1528, 1997.

[3] L. Ljung. Asymptotic variance expressions for identified black-box transfer function models. *IEEE Transactions on Automatic Control*, 30:834–844, 1985.

[4] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[5] L. Ljung and Z.-D. Yuan. Asymptotic properties of black-box identification of transfer function. *IEEE Transactions on Automatic Control*, 30:514–530, 1985.

[6] B. Wahlberg and L. Ljung. Hard frequency-domain model error bounds from least-squares like identification techniques. *IEEE Transactions on Automatic Control*, 37:900–912, 1992.

[7] Z.-D. Yuan and L. Ljung. Black-box identification of multivariable transfer functions – Asymptotic properties and optimal input design. *International Journal of Control*, 40:233–256, 1984.

[8] Y.-C. Zhu. Black-box identification of MIMO transfer functions: Asymptotic properties of prediction error models. *International Journal of Adaptive Control and Signal Processing*, 3:357–373, 1989.