System Identification without Lennart Ljung: what would have been different?

Michel Gevers

To Lennart, for a friendship of more than thirty years,
that has never faltered.

Abstract: This chapter presents a personal view on the development of identification theory in the control community, starting from the year 1965. We show how two landmark papers, (Ho and Kalman, 1965) and (Åström and Bohlin, 1965), gave birth to two main streams of research that have dominated the development of system identification over the last fourty years. The Ho-Kalman paper gave a first solution to state-space realization theory, which led to stochastic realization, and much later to subspace identification. The Åström-Bohlin paper laid the foundations for the highly successful Prediction Error methods based on parametric input-output models. The chapter highlights the key influence of Lennart Ljung on the development of Prediction Error Identification; it shows how his seminal contributions have profoundly changed the community's view on identification from a search for the elusive "true system" to a goal-oriented design problem.

1 Introduction

The development of identification theory in the control literature followed on the heels of the development of model-based control design around 1960. Up until the late 1950's, much of control design relied on Bode, Nyquist and Nichols charts, or on step response analyses. These techniques were limited to control design for single-input single-output (SISO) systems. Around 1960, Kalman introduced the state-space representation and laid the foundations for state-space based optimal filtering and optimal control theory, with Linear Quadratic (LQ) optimal control as a cornerstone for model-based control design.

The availability of these model-based control design techniques put pressure on the scientific community to extend the fields of application of "modern" control design beyond the realm of mechanical, electrical and aerospace applications, for which reliable models were easily available. Thus the need arose to develop data-based techniques that would allow one to develop dynamical models for such diverse fields as process control, environmental systems, biological and biomedical systems, transportation systems, etc.

Much of the early work on identification was developed by the statistics, econometrics and time series communities. Even though the statistical theory of parameter estimation has its roots in the work of Gauss (1809) and Fisher (1912), most of the theory of stationary stochastic processes was developed during the period 1920 to 1970. We shall not describe this work here, because we want to focus on the engineering views and developments of system identification. An excellent review of the history of system identification and time series analysis in the statistics community can be found in Deistler (2002).

Although a lot of results had already been established in the statistics and econometrics literature, one can view 1965 as the birthyear for identification theory in the control community, with the publication of two seminal papers, Ho and Kalman (1965) and Åström and Bohlin (1965). These two papers paved the way for the development of the two mainstream identification techniques that still dominate the field today: subspace identification and prediction error identification. The former is based on projection techniques in Euclidean space, the latter is based on the minimization of a parameter dependent criterion.

The Ho-Kalman paper provided the first solution to the determination of a minimal state-space representation from impulse response data. The solution of this deterministic realization problem was later extended by Akaike (1974) and others to stochastic realization, where a Markovian model is obtained for a purely random process on the basis of covariance data. This technology, based on canonical correlation analysis, was extended in the early nineties to processes that also contain a measurable (control) input, and was then rebaptized as subspace state-space identification.

The Åström-Bohlin paper introduced into the control community the Maximum Likelihood framework that had been developed by time series analysts for the estimation of the parameters of difference equation models. These were known in the statistical literature by such esoteric names as ARMA (AutoRegressive Moving Average) or ARMAX model (AutoRegressive Moving Average with eXogeneous inputs). These models, and the Maximum Likelihood framework, were there to stay, since they gave rise to the immensely successful Prediction Error Identification framework.

In 1970, Box and Jenkins published their book "Time series analysis, forecasting and control", Box and Jenkins (1970), which gave a major impetus to applications of identification. Indeed, the book gave a rather complete recipe for identification, all the way from initial data analysis to the estimation of a model. In the spirit of the time series analysis methods of the time, it relied on correlation analysis for the determination of model structure. For about 15 years, it remained the major high quality reference book on system identification. Other important references of the time were the survey paper Åström and Eykhoff (1971) and the special issue on system identification and time series analysis published by the IEEE Transactions on Automatic Control in December 1974. The Åström

and Eykhoff survey was to be used by many young researchers of the time as a stepping stone for future work. It explained the state of the art as much as it displayed some of the important open questions of the time. One of these was the identification of closed-loop systems, for which the Hankel-based projection methods (based on cross-correlation information) had been shown to fail.

From about the mid-seventies, the Prediction Error (PE) framework completely dominated identification theory and, perhaps more importantly, identification applications. Just about all of the activity at that time was focused on the search for the "true system", i.e. it dealt with questions of identifiability, convergence to the "true parameters", and asymptotic normality of the estimated parameters. Much of that activity dealt with identifiability problems for multivariable systems and for closed-loop systems.

Around 1976 the first attempts were made to view system identification as an approximation problem, in which one searches for the best possible approximation of the "true system" within some model class: Ljung (1976); Anderson et al. (1978); Ljung and Caines (1979). The prevailing view changed consequently from a search for the "true system" to a search for and characterization of the "best approximation". Hence, the characterization of the model errors (bias error and variance error) became the focal point of research. For control engineers, the object of primary interest is the model, in particular the transfer function model, rather than the parameters which are just a vehicle for the description of this model. As it turns out, the research on bias and variance error moved remarkably swiftly from the characterization of parameter errors to that of transfer function errors, thanks to some remarkable analysis of Ljung based on the idea of letting the model order go to infinity: Ljung (1985); Wahlberg and Ljung (1986).

The work on bias and variance analysis of identified models of the eighties then led, almost naturally, to a new perspective in which identification became viewed as a "design problem". With an understanding of the effect of the experimental conditions, the choice of model structure, and the choice of criterion on the quality of the identified model, one can tune these design variables towards the objective for which the model is being identified: Gevers and Ljung (1986). The book "System identification: Theory for the user", Ljung (1987), has had a profound impact on the engineering community of system identifiers. It squarely put forward the view of system identification as a design problem, in which the model use plays a central role. This viewpoint clearly distinguishes the field from the statistical literature on system identification and time series analysis, where the prevailing view is that the model must "explain" the data as best as possible.

The observation that the model quality can be tuned, through the choice of appropriate design variables, towards the eventual objective for which the model is being built opened the way to a flood of new activity that took place in the nineties and continues up to this day. The major application of this new paradigm is the situation where a model is built with the view of designing a model-based

controller. Thus, identification for control has blossomed, since around 1990. Because that topic embraces many aspects of identification and robust control theory, it has also opened or reopened new research interest in areas such as experiment design, closed-loop identification, frequency domain identification, uncertainty estimation, and data-based robust control analysis and design.

The present chapter attempts to exhibit both the continuity and the motivation for the developments that took place in system identification in the last fourty years, and also the significant new departures and insights that came as the result of some important breakthroughs. In doing so, this chapter will show how Lennart Ljung was responsible for several of these breakthroughs.

2 The milestone papers

2.1 Deterministic realization theory

In 1965, Ho and Kalman (1965) provided a first solution to a challenging system theoretical problem that became known as the state-space realization problem. It can be stated as follows.

Construct a minimal state-space realization

$$\left\{ \begin{array}{lcl} x_{t+1} & = & Ax_t + Bu_t \\ y_t & = & Cx_t \end{array} \right.$$

for an input-output model described by its impulse response matrices (also called Markov parameters) $H_k \in \mathbb{R}^{p\times m}$*:*

$$y_t = \sum_{k=1}^{\infty} H_k u_{t-k}.$$

The problem is thus to replace the infinite description

$$H(z) = \sum_{k=1}^{\infty} H_k z^{-k}$$

by a finite description with $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$, such that

$$H(z) = C(zI - A)^{-1}B$$

with $\dim(A)$ minimal. This problem can be split up into two parts: (i) find the McMillan degree of $H(z)$, which is then the minimal dimension of A; (ii) compute the matrices A, B, C. The key tool for the solution of this problem is the

Hankel matrix, and its factorization into the product of an infinite observability matrix times an infinite controllability matrix:

$$\mathcal{H} = \begin{bmatrix} H_1 & H_2 & H_3 & H_4 & \dots \\ H_2 & H_3 & H_4 & H_5 & \dots \\ H_3 & H_4 & H_5 & H_6 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix}$$

If the McMillan degree of $H(z)$ is n, then

1. rank $\mathcal{H} = n$
2. $\exists\, A \in \mathbb{R}^{n\times n}, B \in R^{n\times m}, C \in \mathbb{R}^{p\times n}$ such that $H_k = CA^{k-1}B$.

It took years of research to go from the theoretical results described in Ho and Kalman (1965) to a numerically reliable realization algorithm. However, all the key insights were present in the 1965 paper, and they were to have a profound impact on linear system theory, and on realization and identification theory.

2.2 The Maximum Likelihood framework

In complete contrast to the state-space formulation of Ho and Kalman, the landmark paper Åström and Bohlin (1965) introduced the Maximum Likelihood method for estimating the parameters of input-output models in ARMAX form:

$$A(z^{-1})y_t = B(z^{-1})u_t + \lambda C(z^{-1})e_t$$

where $\{e_t\}$ is a sequence of independent identically distributed Normal $(0, 1)$ random variables. The Maximum Likelihood (ML) method was well known and had been widely studied in mathematical statistics and time series analysis. However, what is remarkable about the Åström-Bohlin paper is that the authors not only gave a complete algorithmic derivation of the ML identification method for ARMAX models, but also presented all analysis results that were available at that time, such as the consistency, asymptotic efficiency and asymptotic normality of the parameter estimates, the persistence of excitation conditions on the input signal in connection with the order of the model, the model order validation on the basis of the whiteness of the residuals, etc.

The concepts and notations introduced by Åström and Bohlin in 1965 have been with us for almost 40 years now. Indeed, the following household notations of the identification community can all be found in this milestone paper:

- the residuals $C(z^{-1})\varepsilon_t = A(z^{-1})y_t - B(z^{-1})u_t$
- the cost criterion $V(\theta) = \frac{1}{2}\sum_{t=1}^{N} \varepsilon_t^2$

- the parameter estimate $\hat{\theta} = \arg\min V(\theta)$
- the white noise variance estimate $\hat{\lambda}^2 = \frac{2}{N} V(\hat{\theta})$.

The publication of Åström and Bohlin (1965) gave rise to a flurry of activity in parametric identification. It also established the basis for the adoption of the Prediction Error framework. The step from Maximum Likelihood to Prediction Error essentially consists of observing that, under an assumption of white Gaussian noise in the ARMAX model, the maximization of the likelihood function of the observations is equivalent to the minimization of the sum of the prediction errors. The Prediction Error framework then consists of adopting the minimization of a norm of the prediction errors as a reasonable criterion for parameter estimation, even in the absence of any known probability distribution for the observations. Such suggestion had already been made by Mr. Gauss himself, Gauss (1809), as observed in the fascinating paper Åström (1980).

3 From deterministic to stochastic realization

The combination of the deterministic realization theory based on the factorization of the Hankel matrix, and of the theory of Markovian and innovations representations gave rise to the stochastic theory of minimal realizations. The stochastic realization problem can be stated as follows.

Given the covariance sequence $\{R_k, k = 1, 2, \ldots, \infty\}$ *of a zero-mean stochastic process* $\{y_t\}$*, where* $R_k \triangleq E\{y_t y_{t-k}^T\}$*, find a minimal Markovian representation for the process* $\{y_t\}$*, of the form*

$$\left\{ \begin{array}{lcl} x_{t+1} & = & Ax_t + Gw_t \\ y_t & = & Cx_t + v_t \end{array} \right. \tag{1}$$

where $\begin{pmatrix} w_t \\ v_t \end{pmatrix}$ *is a zero-mean white noise sequence with covariance matrix*

$$W = E\left\{ \begin{pmatrix} w_t \\ v_t \end{pmatrix} \begin{pmatrix} w_t \\ v_t \end{pmatrix}^T \right\} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}.$$

This problem amounts to finding state-space matrices $\{A, G, C\}$ *with* $n = dim(A)$ *minimal, and the elements* Q, S, R *of the covariance matrix W such that the covariance of the output of (1) is exactly* R_k*.*

Observe that the covariance of the output of the Markovian representation (1) is given by $R_k = CA^{k-1}N$ with $N = A\Pi C^T + GS$ for $k > 0$, and $R_0 = C\Pi C^T + R$, where Π is the state covariance: $\Pi \triangleq E\{x_t x_t^T\}$.

The stochastic realization problem was studied very intensively during the early seventies in connection with innovations theory and spectral factorization theory: Akaike (1974); Gevers and Kailath (1973); Faurre (1976). The first step of the solution consists of observing that the Hankel matrix made up of the covariance sequence can be factored as

$$\mathcal{H} = \begin{bmatrix} R_1 & R_2 & R_3 & \dots \\ R_2 & R_3 & R_4 & \dots \\ R_3 & R_4 & R_5 & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \begin{bmatrix} N & AN & A^2N & A^3N \dots \end{bmatrix}$$

where $R_k = CA^{k-1}N$ with N defined as above. The Ho-Kalman algorithm mentioned above allows one to determine the minimal dimension n as the rank of $\mathcal{H}$, and to compute the matrices C, A and N from the factorization of $\mathcal{H}$.

There are various ways of performing the second step, which consists of computing the missing elements G, Q, R, S of the Markovian representation (1) from C, A, N and the output variance R_0. One way is to use a particular version of the Markovian representation called an *innovations model*:

$$\left\{ \begin{array}{rcl} \xi_{t+1} & = & A\xi_t + K\varepsilon_t \\ y_t & = & C\xi_t + \varepsilon_t \end{array} \right. \tag{2}$$

where $\{\varepsilon_t\}$ is a white noise sequence with covariance matrix $\Sigma = E\{\varepsilon_t \varepsilon_t^T\}$. Denoting $\hat{\Pi} = E\{\xi_t \xi_t^T\}$ the covariance of the state of this model, and imposing that the covariance sequence of the output of this innovations model (2) is $\{R_k\}$, one gets the following three constraints on the unknown quantities $\hat{\Pi}$, K and Σ:

$$\left\{ \begin{array}{rcl} \hat{\Pi} & = & A\hat{\Pi}A^T + K\Sigma K^T \\ N & = & A\hat{\Pi}C^T + K\Sigma \\ R_0 & = & C\hat{\Pi}C^T + \Sigma \end{array} \right.$$

The first (Lyapunov) equation follows directly from the state equation of the Markovian model, while the other two constraints are imposed by the matching of the output covariance R_k: see above. The solution of this second step is obtained via an associated Riccati equation.

An interesting aspect of the solution of the stochastic realization problem given by Akaike was the definition of the state of the innovations model as the set of canonical correlations obtained by projecting the space of future outputs onto the space of past outputs Akaike (1975). This insight formed the basis for the later work on *subspace identification*; it also gave rise to extensive studies of the interface between past and future observation spaces, called minimal splitting subspaces: see e.g. Katayama and Picci (1999).

Another important outcome of the stochastic realization and innovations theories of the seventies was the covariance equivalence established between the Markovian realization (1) and its innovations realization (2). In particular, this means that, even if a Markovian model (1) has been constructed as a *physical model* for y_t, driven by two independent white noise sources w_t and v_t (i.e. $S = 0$, say), it can be rewritten as the Markovian innovations model (2) driven by a unique white noise source ε_t with the same output covariance $\{R_k\}$. The same equivalence applies to models with deterministic inputs. This means that any (physical) state-space model

$$\left\{ \begin{array}{lcl} x_{t+1} & = & Ax_t + Bu_t + Gw_t \\ y_t & = & Cx_t + v_t \end{array} \right.$$

can be rewritten as an *innovations model* for y_t driven by a unique noise source ε_t:

$$\left\{ \begin{array}{lcl} \xi_{t+1} & = & A\xi_t + Bu_t + K\varepsilon_t \\ y_t & = & C\xi_t + \varepsilon_t. \end{array} \right.$$

The input-output equation of this state space innovations model is:

$$\begin{array}{rcl} y_t & = & C(zI - A)^{-1}Bu_t + [C(zI - A)^{-1}K + I]\varepsilon_t \\ & = & G(z)u_t + H(z)\varepsilon_t \end{array}$$

where

$$G(z) = C(zI - A)^{-1}B,\ H(z) = I + C(zI - A)^{-1}K.$$

Observe that the transfer functions $G(z)$ and $H(z)$ have the same poles, and that $H(z)$ is monic. Thus, this model is equivalent to an ARMAX model:

$$A(z^{-1})y_k = B(z^{-1})u_k + C(z^{-1})\varepsilon_k.$$

This theory established a nice link between Markovian models obtained from 'first principles modeling', their corresponding state-space innovations models, and the ARMAX input-output models used in Maximum Likelihood (ML) or Prediction Error (PE) identification. It also gave a solid theoretical justification for the use of ARMAX models to represent stationary linear Markov processes.

4 The golden years: 1975-1985

4.1 The big cleanup

The years 1975 to 1985 saw a frantic activity in system identification in the engineering community. The parametric input/ouput methods, in large part because increased computer speed and the development of special purpose identification

software made it more and more feasible to perform the iterative minimization of a cost criterion over a range of possible model structures. This is the period where many authors were putting their name on a new combination of model structure and method, with the inevitable claim about the supremacy of their new combination over existing methods. New "methods" were pouring out constantly in the scientific journals.

Some solid cleaning work was required, and it was one of L. Ljung's major contributions to perform this cleaning work. It consisted of clearly separating two independent concepts: the choice of a parametric model structure, which was seen as just a vehicle for computing predictions and hence parameter dependent prediction errors; and the choice of an identification criterion, which was to be chosen as some positive function of these prediction errors and hence of the parameter vector: see e.g. Ljung (1978). All existing parametric identification methods could then be seen as particular cases of this prediction error framework.

Thus Ljung introduced the generic input-output model structure

$$y_t = G(z,\theta)u_t + H(z,\theta)e_t \tag{3}$$

with $G(z,\theta)$ and $H(z,\theta)$ parametrized as rational transfer functions, and with e_t white noise. All commonly used model structures, such as ARX, ARMAX, OE, BJ, FIR, were special cases of the structure (3). From such model structure, one can derive the parameter-dependent one-step ahead prediction $\hat{y}_{t|t-1}(\theta)$, and hence the one-step ahead prediction error $\varepsilon_t(\theta) = y_t - \hat{y}_{t|t-1}(\theta)$. From a set of N data, Z^N, and hence of N prediction errors, $\{\varepsilon_t(\theta), t = 1,\ldots,N\}$, one can then define a criterion

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} l(\varepsilon_t(\theta)), \tag{4}$$

where $l(.)$ is a positive scalar-valued function. The minimization of $V_N(\theta, Z^N)$ with respect to θ in some domain D_θ then yields the parameter estimate:

$$\hat{\theta}_N = \arg\min_{\theta\in\mathcal{D}_\theta} V_N(\theta, Z^N). \tag{5}$$

The methodological work of Ljung culminated with the publication of his book, Ljung (1987), which has become the standard reference book in system identification. The impact of his book was greatly amplified by the simultaneous production by L. Ljung of the MATLAB identification toolbox, whose tools were directly connected to the theory developed in the book. Ljung's book was complemented by that of Stoica and Söderström, who adopted the same clear distinction between choice of model structure and choice of criterion; their book focused less on design issues, but more on analysis and on alternative criteria, in particular criteria based on correlation methods and instrumental variables: see Söderström and Stoica (1989).

4.2 Breakthroughs for MIMO and for closed-loop systems

During that same period, some important theoretical breakthroughs were made in two directions. The first was the elucidation of the manifold structure of multi-input multi-output systems. The second consisted of a range of identifiability results for linear systems identified under closed-loop conditions.

Many authors contributed to the solution of both problems, and we cannot possibly name them all. Let us just mention that the manifold structure of MIMO systems, together with the key role of the Kronecker (or structure) indices, was elucidated in Hazewinkel and Kalman (1976) and Clark (1976). Subsequently, many authors worked on methods for the estimation of these structure indices; others studied the relationship between the canonical (or pseudo-canonical) forms in state-space and in ARMA form: Deistler and Gevers (1981); Van Overbeek and Ljung (1982); Picci (1982); Rissanen (1982); Wertz et al. (1982).

As for the identifiability of closed-loop systems, one of the earliest solutions was provided by the famous Swedish trio made up of Gustavsson, Ljung and Söderström[1], all PhD students of K.J. Åström at the time: see Söderström et al. (1976). They showed, among other things, that in many situations of practical interest, the direct application of a PE identification method to input-output data allows one to identify the open-loop plant, despite the presence of a feedback controller. Other closed-loop identifiability results covered indirect methods (in which the closed-loop transfer function is identified first, and the plant model is then derived from it using knowledge of the controller) and the so-called 'joint input-output method': see Ng et al. (1977); Gustavsson et al. (1977); Sin and Goodwin (1980); Anderson and Gevers (1982).

4.3 System identification viewed as approximation

For most of the sixties and seventies, the prevailing assumption was that the system was in the model set: $\mathcal{S} \in \mathcal{M}$. Thus, the focus of research was on questions of convergence to the *true system* and of statistical efficiency of the parameter estimates. In the mid-seventies, the first attempts were made to view system identification in the context of *approximation*: Ljung (1976); Anderson et al. (1978); Ljung and Caines (1979). This marked the beginning of an entirely new era, in which the elusive search for a linear time-invariant 'true system' was progressively abandoned to give way to the search for a 'best approximate model' within some a priori chosen model set $\mathcal{M}$. With the idea of model approximation came of course the idea of model error, and hence the desire to characterize this model error.

[1] Over a 4-year period they jointly published no less than 6 important papers on various aspects of system identification.

4.4 The birth of θ^*

In statistics, the natural way to analyse estimation errors is through the concepts of bias and variance error of the parameter estimate. However, in the context of model sets that do not contain the true system, the concept of parameter error becomes meaningless; since there are no 'true parameters', the bias error in the parameter estimate does not make sense. The object of interest is the transfer function, not the parameters that are used to represent it. By observing that, under reasonable conditions, the parameter estimate $\hat{\theta}_N$ converges to a well-defined θ^*,

$$\lim_{N\to\infty} \hat{\theta}_N = \theta^* = \arg\min_{\theta\in\mathcal{D}_\theta} \lim_{N\to\infty} EV_N(\theta),$$

Ljung introduced the concept of bias error and variance error of the transfer function estimate by defining the following decomposition of the total transfer function error, at a given frequency ω:

$$\begin{aligned} G_0(e^{j\omega}) - G(e^{j\omega},\hat{\theta}_N) &= \underbrace{G_0(e^{j\omega}) - G(e^{j\omega},\theta^*)}_{bias\ \ error} \\ &+ \underbrace{G(e^{j\omega},\theta^*) - G(e^{j\omega},\hat{\theta}_N)}_{variance\ \ error} \end{aligned}$$

Within this framework, he and his collaborators derived approximate asymptotic expressions for the transfer function variance, and integral expressions for the transfer function bias that were to become the cornerstone of the next major phase of the development of system identification: the design phase; see Ljung (1985); Wahlberg and Ljung (1986).

5 Identification as a design problem

If identification is viewed as approximation, then one knows and accepts that the model is erroneous. If the model is to be used for a specific purpose (as is most often the case), then perhaps one can construct a model in such a way that the model errors do not penalize too much the goal for which the model is being built. This is the whole idea behind goal-oriented identification. More precisely, if one can understand the connection between

- identification design (experiment design, choice of model structure, choice of criterion) and model quality on the one hand,
- the effect of model quality on the intended model application on the other hand,

then one can formulate the identification problem as a *goal-oriented design problem*.

The first few steps of this new paradigm were laid in Gevers and Ljung (1986); Wahlberg and Ljung (1986); Ljung (1987). This new (engineering) way of looking at the identification problem opened up a vast new window of opportunities for research. In particular, this viewpoint was instrumental in the development, from around 1990, of a field that has seen an enormous activity ever since, both on the theoretical front and in practical applications, namely *identification for control*.

6 System identification in the nineties

At the triennial IFAC Symposium on System Identification held in Budapest in 1991, there was a feeling among many participants that most of the important problems in system identification had been solved and that the golden age of identification was over. This prediction proved to be wrong. The research on system identification was pulled all through the nineties essentially by two catalysts, whose first feeble signs emerged around 1990: subspace-based identification, and identification for control. In addition, important new research activity took place in frequency domain identification, in closed-loop identification, and in the development of a range of new methods for the quantification of model uncertainty.

6.1 Subspace-based identification

The reasons for the emergence of subspace identification are almost certainly to be found in the state of the art of identification of multivariable systems in the eighties. Even though the manifold structure of MIMO (multi-input multi-output) systems had been completely characterized in the late seventies, the practical problem of identifying MIMO systems remained wide open. Indeed, the estimation of the structure indices that characterize the parametrizations of multivariable systems remained very tricky, and led to ill-conditioned numerical procedures. Thus, there was a great incentive for the development of simple, albeit suboptimal procedures, based on the numerically robust Singular Value Decomposition and Least Squares techniques, that completely bypass the need for the estimation of structure indices. The development of subspace-based identification methods filled a much needed gap, because in that framework the handling of MIMO systems causes no additional difficulty.

A major difficulty was that the projection methods developed by Akaike, which were based on canonical correlation analysis, were not easily extendable to output data that contained, besides the stochastic components, also a contri-

bution due to a measured input. In the early nineties, several research teams managed to crack this nut; they provided several closely related solutions to the problem: see Larimore (1990); Van Overschee and De Moor (1994); Verhaegen (1994); Viberg (1995). These first few solutions opened the way for a lot of research on the properties of subspace-based identification, their connection with stochastic realization theory, and on improvements of the numerical procedures: see e.g. Chui and Maciejowski (1996); Katayama and Picci (1999); Chiuso and Picci (2002).

6.2 Identification for control

Identification for control has been the major outlet for the new paradigm of system identification as a design problem. The reasons for this are many: (i) in the systems and control community of system identification, control is very often the main motivation for model building; (ii) it has been observed that high performance control can often be achieved with very simple models, provided some basic dynamical features of the system are accurately reflected; (iii) a powerful robust control theory, based on models and uncertainty sets, had been developed all through the eighties but the models and uncertainty sets used were not data-based for lack of a proper theory; (iv) the identification for control research delivered iterative model and controller tuning tools that were intuitive, practical and easy to implement by the process engineers.

Whereas the building blocks for goal-oriented identification were laid around 1986, the first specific contributions in which identification and control design were looked upon as a combined design problem appeared only around 1990. The plenary Gevers (1991) at the 1991 IFAC Symposium on System Identification addressed many of the key issues about the interplay between the identification of a reduced order model and the design of a controller from such model; however, it was more an agenda for research than a presentation of solutions. Indeed, in 1990 there was very little understanding about the interplay between system identification and robust control. The two theories had been developed by two separate communities that had had very little contact with one another.

The robust control community had developed a robust analysis and design theory that was based on uncertainty descriptions which were not based on data but on prior assumptions. The identification community had delivered bias and variance error descriptions that were not explicit, and that were certainly not transferable at the time to the toolboxes used in robust control analysis and design. More importantly, neither community had given much attention to the interaction between model building and control design: what are the qualities that a model must possess (or, conversely, what are the plant-model errors that are acceptable) if the model is to be used for the design of a controller that must achieve a given level of performance on the plant?

In July 1992, the IEEE Transactions on Automatic Control devoted a special issue to 'system identification for robust control design'. In retrospect, and in keeping with the observation just made, such issue was perhaps premature, given the paucity of results that were available at that time. About half of the papers in that special issue did not really deal with identification for control, but with the estimation of uncertainty sets, without any account taken of control-oriented design issues. A few papers in that special issue did address the joint 'identification and control design' paradigm, notably Bayard et al. (1992) and Schrama (1992). They produced one of the first key results in identification for control: the necessity of an iterative scheme for the design of a control-oriented nominal model; this observation had in fact been first made in Liu and Skelton (1990).

The first half of the nineties produced a string of results on the design of control-oriented nominal models. These results were produced by different teams who used their favorite combinations of identification criteria and control design criteria: Schrama and Bosgra (1993); Lee et al. (1993); Åström and Nilsson (1994); Zang et al. (1995); de Callafon and Van den Hof (1997). They confirmed the necessity of using an iterative scheme of model updates and controller updates, and they produced significant evidence - but no hard proofs - about the advantages of performing the identification in closed loop, rather than in open loop, when the model is to be used for the design of a new controller. Of course, the closed-loop experimental conditions that produce a specific input signal spectrum can always be mimicked by an open loop experiment with the same input spectrum, but the advantage of the closed-loop experiment is that the appropriate input spectrum is automatically generated by the unknown system itself.

The iterative schemes of identification and control design had a remarkably fast transfer into the world of applications. There were two main reasons for this:

- whereas the industrial world was still living with the belief that one should 'open the loop' to perform a valid identification experiment, here was a new theory that showed the benefits of closed-loop identification; this came as welcome news to process control engineers who had never really liked the idea of opening the loop;

- in the process industry, thousands of measurements are flowing into the computer; here was a theory that showed how these data could be used for the design of a better controller.

The early work on identification for control focused on control-oriented identification criteria, i.e. on obtaining a nominal model whose bias error distribution was tuned for control design. This means that the nominal control performance obtained with the optimal controller computed from the nominal model is close to the actual control performance obtained with the same controller on the actual plant.

More recently, attention has shifted to the distribution of the variance error of the identified models, i.e. to the estimation of control-oriented uncertainty sets: Kosut and Anderson (1994); Mäkilä et al. (1995); de Vries and Van den Hof (1995); Bombois et al. (2000). The idea is that, since one can manipulate the shape of the model uncertainty set by the choice of the experimental conditions under which the new model is identified, one should attempt to obtain a model uncertainty set that is tuned for control design. Even though many new insights have been gained on the interplay between uncertainty sets estimated from data and corresponding sets of stabilizing controllers, there is at this point no clear view as to the most operational definition of a 'control-oriented uncertainty set'. One view is that the corresponding set of controllers achieving stability and the required performance with all models of that set should be as large as possible Gevers et al. (2003). Another view is that the worst-case performance achieved by some optimal robust controller with all models in this uncertainty set should be as close as possible to the performance achieved with the central (nominal) model.

The work on identification for control has had many beneficial side effects: it has forced the scientific community to reassess some 'truths' that had been considered to be firmly establised and to reopen some research questions that had been considered to be completely settled. Identification for control has triggered an enormous new research activity on the estimation of data-based uncertainty sets for identified models, as well as on the identification of systems operating in closed loop. An intense debate has been reopened on the relative merits of model or controller reduction versus the direct identification of a control-oriented restricted complexity model, leading to the notion of *near-optimal restricted complexity model* introduced in Hjalmarsson (2005), a must-read for any researcher in identification for control. Finally, identification for control has led to the recent rebirth of the topic of optimal experiment design, a subject that had been very active in the 1970's but which had been almost completely abandoned since then. The overview paper Gevers (2005) presents up-to-date results on the use of optimal experiment design in the context of identification for control, a topic that is also very present in Hjalmarsson (2005).

In the remainder of this chapter, we briefly review the research activities that were triggered by identification for control, as well as some other topics that have seen important developments in the last decade.

6.3 Quantification of model uncertainty

The demands of robust control theory for adequate uncertainty sets triggered a lot of interest for the estimation of uncertainty sets from data. It is fair to say that most of the robust control theory developed in the eighties had been based on a priori assumed uncertainty bounds on model errors and on the noise. In

the context of system identification, the estimation of model errors from data is of course of interest in its own right; indeed a reputable engineer should never deliver a product (a model, in this case) without a statement about its error margins. In this authors's opinion, this activity on estimation of uncertainty sets was erroneously put under the umbrella of 'identification for control'; indeed, in most of the work on data-based estimation of uncertainty sets of the last decade, the control objective is not taken into account in the identification design.

A wide range of new techniques were developed to provide error bounds on estimated models, in all kinds of shapes and frameworks, using time domain, frequency domain, H_∞, l_1, probabilistic, worst case, set membership and other methods: see Helmicki et al. (1991); Goodwin et al. (1992); Poolla et al. (1994); Smith and Dahleh (1994); de Vries and Van den Hof (1995); Kosut (1995); Hakvoort and Van den Hof (1997); Giarré et al. (1997). Even though it is now a few years old, probably one of the best presentations of these different methods is the survey paper Ninness and Goodwin (1995).

6.4 Closed-loop identification revisited

For some reason, the work of the seventies on closed-loop identification had stopped at the question of identifiability, i.e. under what conditions do the parameter estimates converge to the 'true' parameters in the case where the system is in the model set ($\mathcal{S} \in \mathcal{M}$). The influence of the experimental conditions on bias error distribution in the case of restricted complexity models, as well as on asymptotic variance, had not been addressed.

One of the important lessons that emerged from the study of the interplay between identification and control is the benefit of closed-loop identification when the model is to be used for control design. Until the late eighties, it was commonly accepted that closed-loop identification was preferably to be avoided. In identification for control with reduced order models, the required connection between the control performance criterion (obviously a closed-loop criterion) and the identification criterion, established the need for closed-loop identification, as described above. In the ideal context of optimal experiment design with full order models, optimality of closed-loop identification was actually established when the model is to be used for control design with a noise rejection objective: Gevers and Ljung (1986); Hjalmarsson et al. (1996); Forssell and Ljung (2000).

This observation triggered an important new activity in the design of special purpose closed-loop identification methods, the main goal pursued by these new methods being to obtain a better handle on the bias error in closed-loop identification: Hansen et al. (1989); Van den Hof and Schrama (1993); Van den Hof et al. (1995); Forssell and Ljung (1999).

6.5 Optimal experiment design for control

In the 1970's, optimal input design for system identification was an active area of research, with different quality measures being used for this optimal design: Mehra (1974); Zarrop (1979); Goodwin and Payne (1977). The questions at that time addressed open-loop identification and focused on quality measures of the parameter covariance matrix P_θ. That work ground to a halt for about fifteen years. The new paradigm of identification as a design problem has given it a new lease of life. The recent work on the connection between model uncertainty sets obtained by identification and corresponding sets of robust controllers have clearly put this subject in the limelight again: Gevers et al. (1999), Forssell and Ljung (2000), Hildebrand and Gevers (2003); Lindqvist and Hjalmarsson (2001); Barenthin et al. (2005); Bombois et al. (2004). The new emphasis is to establish a direct link between the experimental conditions under which a model is identified (together with its uncertainty set) and the performance of the controller that results from the use of the model and its estimated uncertainty set.

6.6 Frequency domain identification

Another area of important activity in the nineties has been frequency domain identification. In the identification community, there has historically been no more than a polite attention paid to frequency domain identification. Every now and then, some of the experts in the area published a paper whose main aim was to reassure the community with a message that sounded essentially like 'yes, indeed, we can also handle frequency domain data': Ljung and Glover (1981).

Things changed drastically at the end of the eighties, with a convergence of efforts arising from two completely different horizons.

- During the eighties, the robust control community had developed most of its analysis and design tools in the frequency domain; thus, there was a great demand for tools that would enable one to obtain frequency domain models and, even more importantly, frequency domain uncertainty descriptions from data. This led, at the end of the eighties, to the development of a number of new interpolation techniques that were using noisy pointwise frequency domain transfer function measurements as their data Parker and Bitmead (1987); Partington (1991); Gu and Khargonekar (1992);

- At about the same time, Pintelon and Schoukens had independently developed frequency domain identification techniques, essentially based on the Maximum Likelihood principle. With their instrumentation and measurement background, they were interested in methods that would deliver reliable models for devices under test, through the application of short input data sequences Pintelon and Schoukens (1990).

Aware of the interest for frequency domain identification emanating from the robust control community, Schoukens and Pintelon continued their work on frequency domain identification all through the nineties, with specific and important contributions on the use of periodic excitation and of maximally informative input signals. Their book Pintelon and Schoukens (2001) provides a comprehensive treatment of frequency domain identification. In Ljung (2006a), Ljung updates the comparative analysis of Ljung and Glover (1981) between time and frequency domain identification on the basis of the new understanding gained about frequency domain identification over the last twenty years.

6.7 Some other areas of activity in the last decade

The areas mentioned above have been areas of major activity in the nineties. With the exception of subspace-based identification, most of these areas have had their origin in the new engineering view of identification as a design problem. This has created a lot of interest in all aspects of *identification for control*, which turned out to require better methods for the quantification of uncertainty, a better understanding of the bias issues in closed-loop identification and the development of frequency domain identification methods that were as much as possible compatible with frequency domain based robust control analysis tools. But these were by no means the only areas of activity in system identification.

• Interesting progress has been made on the use of alternative basis functions (other than the shift operator) for the representation of input-output models, such as Laguerre, Kautz, and other generalized orthonormal basis functions. Such alternative bases can not only lead to more compact descriptions when some prior knowledge is available about the system, but they have also led to improved formulas for the estimation of the variance of black box transfer function models: Heuberger et al. (1995); Ninness et al. (1999a,b).

• Finally, this overview of recent activity would not be complete without a mention of the topic of identification of nonlinear systems. As is very nicely stated in Deistler (2002), *"identification of nonlinear systems" is like a statement about "non-elephant zoology"*. Indeed, the area is vast, and the approaches are many, ranging from semi-parametric identification using, e.g, neural network methods, to the development of ad hoc techniques for specific classes of nonlinear systems such as, e.g. compartmental models. The paper Sjöberg et al. (1995), even though no longer recent, is still a very valuable overview and introduction to the subject.

7 So what would have been different?

What would have been different without Lennart Ljung? Clearly, it is impossible to guess the direction that system identification would have taken without Lennart Ljung. Somebody else might have had the same ideas at about the same time; who knows? But clearly, Ljung is responsible for some key and judicious trajectory changes in the development of system identification theory, which have profoundly altered our ways of thinking about the topic. If I were to single out what I view as his most important contributions, I would list the following.

- His introduction of a solid identification framework with a clear separation between the parametric model, seen as a way of computing predictions, and a criterion, whose minimization yields an estimated model.
- His move from properties of parameter estimates to properties of transfer function estimates, with the introduction of the concept of bias and variance of transfer function estimates. A critical feature here was the introduction of θ^*, Ljung's baby.
- The formulation of system identification as a design problem, with the introduction of the objective as a weighting in the identification criterion.
- The parallel development of the theory and of the identification toolbox. The impact of this effort of Ljung on the spread of system identification in academia and industry has been immense.

8 Concluding remarks

I have presented my personal views on the development of system identification in the control community, as I have observed it over a period of thirty years, both as a student of the subject eager to learn and understand the work of my colleagues, and as an active participant in these developments. To paraphrase Ljung (2006b), I shall say "This chapter is not a survey": see Magritte (1929). Rather it should be read as a story told with the eyes and the prejudices of one of the actors in the field. I tend to believe that the way a particular field of science develops depends on a combination of two forces: the socio-technical environment created by the evolution of the neighbouring fields of science and by the demands of the applications world, and the creative role played by a few individuals who suddenly make it possible to venture into a totally new direction or to establish a useful link with another field of science that sheds new light on the subject. Clearly, if there is one person who has produced these creative moves in system identification over the last 30 years, it is Lennart Ljung.

Acknowledgments
This chapter presents research results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its author.

References

Akaike, H. (1974). Stochastic theory of minimal realization, *IEEE Trans. Automatic Control* **26**: 667–673.

Akaike, H. (1975). Markovian representation of stochastic processes by canonical variables, *SIAM Journal on Control and Optimization* **13**: 162–173.

Anderson, B. and Gevers, M. (1982). Identifiability of linear stochastic systems operating under linear feedback, *Automatica, 18* **2**: 195–213.

Anderson, B., Moore, J. and Hawkes, R. (1978). Model approximation via prediction error identification, *Automatica* **14**: 615–622.

Åström, K. (1980). Maximum likelihood and prediction error methods, *Automatica* **16**: 551–574.

Åström, K. and Bohlin, T. (1965). Numerical identification of linear dynamic systems from normal operating records, *Proc. IFAC Symposium on Self-Adaptive Systems*, Teddington, UK.

Åström, K. and Eykhoff, P. (1971). System identification – A survey, *Automatica* **7**: 123–162.

Åström, K. and Nilsson, J. (1994). Analysis of a scheme for iterated identification and control, *Proc. IFAC Symp. on Identification*, Copenhagen, Denmark, pp. 171–176.

Barenthin, M., Jansson, H. and Hjalmarsson, H. (2005). Applications of mixed H_2 and H_∞ input design in identification, *16th IFAC World Congress on Automatic Control, paper 03882*, Prague.

Bayard, D., Yam, Y. and Mettler, E. (1992). A criterion for joint optimization of identification and robust control, *IEEE Trans. Automatic Control* **37**: 986–991.

Bombois, X., Gevers, M. and Scorletti, G. (2000). A measure of robust stability for an identified set of parametrized transfer functions, *IEEE Transactions on Automatic Control* **45**(11): 2141–2145.

Bombois, X., Scorletti, G., Gevers, M., Hildebrand, R. and Van den Hof, P. (2004). Cheapest open-loop identification for control, *CD-ROM Proc. 33rd IEEE Conf on Decision and Control*, The Bahamas, pp. 382–387.

Box, G. and Jenkins, G. (1970). *Time Series Analysis, forecasting and control*, Holden-Day, Oakland, California.

Chiuso, A. and Picci, G. (2002). Geometry of oblique splitting subspaces, minimality and hankel operators, *'Directions in Mathematical Systems Theory and Optimization', A. Rantzer and C. Byrnes Eds*, Springer-Verlag, New York.

Chui, N. and Maciejowski, J. (1996). Realization of stable models with subspace methods, *Automatica* **32**(11): 1587–1595.

Clark, J. (1976). The consistent selection of parametrizations in systems identification, *Proc. Joint Automatic Control Conference*, Purdue University.

de Callafon, R. and Van den Hof, P. (1997). Suboptimal feedback control by a scheme of iterative identification and control design, *Mathematical Modelling of Systems* **3**(1): 77–101.

de Vries, D. and Van den Hof, P. (1995). Quantification of uncertainty in transfer function estimation: a mixed probabilistic - worst-case approach, *Automatica* **31**: 543–558.

Deistler, M. (2002). System identification and time series analysis: Past, present, and future, *'Stochastic Theory and Control', Festschrift for Tyrone Duncan*, Springer, Kansas, USA, pp. 97–108.

Deistler, M. and Gevers, M. (1981). Some properties of the parametrization of arma systems with unknown order, *Multivariate Analysis* **11**: 474–484.

Faurre, P. (1976). Stochastic realization algorithms, *'System identification: Advances and Case Studies' R.K. Mehra and D.G. Lainiotis Eds.*, Academic Press, New York.

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

Forssell, U. and Ljung, L. (2000). Some results on optimal experiment design, *Automatica* **36**: 749–756.

Gauss, C. (1809). *Teoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, Reprinted translation: 'Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections', Dover, New York.

Gevers, M. (1991). Connecting identification and robust control: A new challenge, *Proc. IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, pp. 1–10.

Gevers, M. (2005). Identification for control: from the early achievements to the revival of experiment design, *European Journal of Control* **11**: 1–18.

Gevers, M., Bombois, X., Codrons, B., Bruyne, F. D. and Scorletti, G. (1999). The role of experimental conditions in model validation for control, *in* A. Garulli, A. Tesi and A. Vicino (eds), *Robustness in Identification and Control - Proc. of Siena Workshop, July 1998*, Vol. 245 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, pp. 72–86.

Gevers, M., Bombois, X., Codrons, B., Scorletti, G. and Anderson, B. (2003). Model validation for control and controller validation in a prediction error identification framework - Part I: theory, *Automatica* **39**(3): 403–415.

Gevers, M. and Kailath, T. (1973). An innovations approach to least-squares estimation, part vi: Discrete-time innovations representations and recursive estimation, *IEEE Transactions Automatic Control* **18**: 588–600.

Gevers, M. and Ljung, L. (1986). Optimal experiment designs with respect to the intended model application, *Automatica* **22**: 543–554.

Giarré, L., Milanese, M. and Taragna, M. (1997). H_∞ identification and model quality evaluation, *IEEE Trans. Automatic Control* **42**(2): 188–199.

Goodwin, G., Gevers, M. and Ninness, B. (1992). Quantifying the error in estimated transfer functions with application to model order selection, *IEEE Trans. Automatic Control* **37**: 913–928.

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

Gu, G. and Khargonekar, P. (1992). A class of algorithms for identification in H_∞, *Automatica* **28**: 299–312.

Gustavsson, I., Ljung, L. and Söderström, T. (1977). Identification of processes in closed loop - identifiability and accuracy aspects, *Automatica* **13**: 59–75.

Hakvoort, R. and Van den Hof, P. (1997). Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors, *IEEE Trans. Automatic Control* **42**(11): 1516–1528.

Hansen, F., Franklin, G. and Kosut, R. (1989). Closed-loop identification via the fractional representation: Experiment design, *Proc. American Control Conference* pp. 1422–1427.

Hazewinkel, M. and Kalman, R. (1976). On invariants, canonical forms and moduli for linear constant, finite-dimensional dynamical systems, *Lecture Notes in Economics and Mathematical Systems*, Vol. 131, Springer-Verlag, Berlin, pp. 48–60.

Helmicki, A., Jacobson, C. and Nett, C. (1991). Control oriented system identification: A worst-case/deterministic approach in H_∞, *IEEE Trans. Automatic Control* **AC-36**: 1163–1176.

Heuberger, P., Bosgra, O. and Van den Hof, P. (1995). A generalized orthonormal basis for linear dynamical systems, *IEEE Trans. Automatic Control* **AC-40**(3): 451–465.

Hildebrand, R. and Gevers, M. (2003). Identification for control: optimal input design with respect to a worst-case ν-gap cost function, *SIAM Journal on Control and Optimization* **41**(5): 1586–1608.

Hjalmarsson, H. (2005). From experiment design to closed-loop control, *Automatica* **41**: 393–438.

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

Ho, B. and Kalman, R. (1965). Effective construction of linear state-variable models from input-output functions, *Regelungstechnik* **12**: 545–548.

Katayama, T. and Picci, G. (1999). Realization of stochastic systems with exogenous inputs and subspace identification methods, *Automatica* **35**: 1635–1652.

Kosut, R. (1995). Uncertainty model unfalsification : a sytem identification paradigm compatible with robust control design, *Proc. 34th IEEE Conference on Decision and Control*, New Orleans,LA.

Kosut, R. and Anderson, B. (1994). Least-squares parameter set estimation for robust control design, *Proc. American Control Conference*, Baltimore,Maryland, pp. 3002–3006.

Larimore, W. (1990). Canonical variate analysis in identification, filtering, and adaptive control, *Proc. 29th IEEE Conf. on Decision and Control*, Honolulu, Hawaii, pp. 596–604.

Lee, W., Anderson, B., Kosut, R. and Mareels, I. (1993). A new approach to adaptive robust control, *Int. Journal of Adaptive Control and Signal Processing* **7**: 183–211.

Lindqvist, K. and Hjalmarsson, H. (2001). Identification for control: adaptive input design using convex optimization, *CD-ROM Porc. of 40th IEEE Conf. on Decision and Control*, Orlando, Florida.

Liu, K. and Skelton, R. (1990). Closed loop identification and iterative controller design, *29th IEEE Conf on Decision and Control* pp. 482–487.

Ljung, L. (1976). On consistency and identifiability, *Mathematical Programming Study* **5**: 169–190.

Ljung, L. (1978). Convergence analysis of parametric identification methods, *IEEE Trans. Automatic Control* **AC-23**: 770–783.

Ljung, L. (1985). Asymptotic variance expressions for identified black-box transfer function models, *IEEE Trans. Automatic Control* **AC-30**: 834–844.

Ljung, L. (1987). *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ.

Ljung, L. (2006a). Frequency domain versus time domain methods in system identification, *in* B. Francis and J. Willems (eds), *Control of Uncertain Systems: Modelling, Approximation and Design*, Springer Verlag, LNCIS Series, pp. 277–291.

Ljung, L. (2006b). Some aspects on nonlinear system identification, *CD-Rom Proc. of 14th IFAC Symposium on System Identification*, Newcastle, Australia, pp. 110–121.

Ljung, L. and Caines, P. (1979). Asymptotic normality of prediction error estimators for approximative system models, *Stochastics* **3**: 29–46.

Ljung, L. and Glover, K. (1981). Frequency domain versus time domain methods in system identification, *Automatica* **17**(1): 71–86.

Magritte, R. (1929). Ceci n'est pas une pipe, *Los Angeles County Museum of Art*, Los Angeles, USA.

Mäkilä, P., Partington, J. and Gustafsson, T. (1995). Worst-case control-relevant identification, *Automatica* **31**(12): 1799–1819.

Mehra, R. (1974). Optimal input signals for parameter estimation in dynamic systems - survey and new results, *IEEE Trans. on Automatic Control* **AC-19**(6): 753–768.

Ng, T., Goodwin, G. and Anderson, B. (1977). Identifiability of mimo linear dynamic systems operating in closed loop, *Automatica* **13**: 477–485.

Ninness, B. and Goodwin, G. (1995). Estimation of model quality, *Automatica* **31**(12): 32–74.

Ninness, B., Hjalmarsson, H. and Gustafsson, F. (1999a). Asymptotic variance expressions for output error model structures, *14th IFAC World Congress*, Beijing, P.R. China.

Ninness, B., Hjalmarsson, H. and Gustafsson, F. (1999b). On the fundamental role of orthonormal bases in system identification, *IEEE Transactions on Automatic Control* **44**(7): 1384–1407.

Parker, P. and Bitmead, R. (1987). Adaptive frequency response identification, *Proc. CDC* .

Partington, J. (1991). Robust identification and interpolation in H_∞, *International Journal of Control* **54**: 1281–1290.

Picci, G. (1982). Some numerical aspects of multivariable systems identification, *Math. Programming Studies* **18**: 76–101.

Pintelon, R. and Schoukens, J. (1990). Robust identification of transfer functions in the s- and z-domains, *IEEE Trans. on Instrumentation and Measurement* **39**(4): 565–573.

Pintelon, R. and Schoukens, J. (2001). *System Identification - A Frequency Domain Approach*, IEEE Press, New York.

Poolla, K., Khargonekar, P., Tikku, A., Krause, J. and Nagpal, K. (1994). A time-domain approach to model validation, *IEEE Trans. Automatic Control* **39**: 951–959.

Rissanen, J. (1982). Estimation of structure by minimal description length, *Circuits, Systems, and Signal Processing* **1**: 395–406.

Schrama, R. (1992). Accurate identification for control: The necessity of an iterative scheme, *IEEE Trans. on Automatic Control* **37**: 991–994.

Schrama, R. and Bosgra, O. (1993). Adaptive performance enhancement by iterative identification and control design, *Int. Journal of Adaptive Control and Signal Processing* **7**(5): 475–487.

Sin, K. and Goodwin, G. (1980). Checkable conditions for identifiability of linear systems operating in closed loop, *IEEE Trans. Automatic Control* **AC-25**: 722.

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennec, P.-Y., Hjalmarsson, H. and Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview, *Automatica* **31**(12): 1691–1724.

Smith, R. S. and Dahleh, M. (1994). *The Modeling of Uncertainty in Control Systems*, Lecture Notes in Control and Information Sciences, Volume 192, Springer-Verlag, London.

Söderström, T., Ljung, L. and Gustavsson, I. (1976). Identifiability conditions for linear multivariable systems operating under feedback, *IEEE Trans. Automatic Control* **AC-21**: 837–840.

Söderström, T. and Stoica, P. (1989). *System Identification*, Prentice-Hall International, Hemel Hempstead, Hertfordshire.

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

Van den Hof, P., Schrama, R., de Callafon, R. and Bosgra, O. (1995). Identification of normalized coprime plant factors from closed-loop experimental data, *European Journal of Control* **1**: 62–74.

Van Overbeek, A. and Ljung, L. (1982). On-line structure selection for multivariable state-space models, *Automatica* **18**: 529–544.

Van Overschee, P. and De Moor, B. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems, *Automatica* **30**(1): 75–93.

Verhaegen, M. (1994). Identification of the deterministic part of mimo state space models given in innovations form from input-output data, *Automatica* **30**(1): 61–74.

Viberg, M. (1995). Subspace-based methods for the identification of linear time-invariant systems, *Automatica* **31**(12): 1835–1851.

Wahlberg, B. and Ljung, L. (1986). Design variables for bias distribution in transfer function estimation, *IEEE Trans. Automatic Control* **AC-31**: 134–144.

Wertz, V., Gevers, M. and Hannan, E. (1982). The determination of optimum structures for the state-space representation of multivariable stochastic processes, *IEEE Transactions on Automatic Control* **27**(6): 1200–1211.

Zang, Z., Bitmead, R. and Gevers, M. (1995). Iterative weighted least-squares identification and weighted LQG control design, *Automatica* **31**(11): 1577–1594.

Zarrop, M. (1979). *Optimal experiment design for dynamic system identification*, Lecture Notes in Control and Information Sciences, Vol. 21, Springer Verlag, Berlin, New York.