A Personal View of the Development of System Identification

A 30-year Journey Through an Exciting Field

MICHEL GEVERS

In this article I describe the development of system identification in the control community as I have observed it over the last 30 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. This article should not be read as a survey of the subject but rather as a story told through the eyes 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 sociotechnological environment created by the evolution of the neighboring fields of science and by the demands of the applications world as well as the creative role played by a few individuals who suddenly make it possible to venture into a completely new direction or to establish a useful link with another field of science that sheds totally new light on the subject. This article attempts to exhibit both the continuity and the motivation for developments in system identification in the last 30 years and also the significant new departures and insights that came as the result of some important breakthroughs. To see the evolution of system identification in this sociohistorical perspective has always helped me a great deal. The main reason for writing this *story* is to share my experience with others, in the hope that this might also help them, and particularly newcomers to the field, understand the connections between successive developments of system identification. Because this article is not meant to be a survey but rather the story of my own journey through system identification, some topics of importance in system identification and their contributors are mentioned only briefly or are simply omitted because they have not played a major part in my own experience as a researcher. These contributions will be part of somebody else's story.

EARLY HISTORY

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 1950s, 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 the cornerstone of model-based control design.

The availability of model-based control-design techniques created a desire in the scientific and engineering community to extend the fields of application of modern control design beyond the realm of mechanical, electrical, and aerospace applications, for which reasonably accurate low-dimensional dynamical models could be obtained from first principles. The need arose, therefore, to develop data-based techniques for developing dynamical models for diverse applications such as process control, environmental systems, biological and biomedical systems, and transportation systems.

Much of the early work on identification was developed by the statistics, econometrics, and time-series communities. Although 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 between 1920 and 1970; an insightful and comprehensive review of the history of system identification and time-series analysis in the statistics, econometrics and time-series communities can be found in [1].

Although many results on system identification had already been established in the statistics and econometrics literature, 1965 can be viewed as the birth year for identification theory in the control community due to the publication of the seminal papers [2] and [3]. These papers paved the way for the development of the two mainstream identification techniques that dominate the field today, namely subspace identification and prediction-error identification. The former is based on projection techniques in Euclidean space, while the latter is based on minimizing a parameter-dependent criterion.

The Ho-Kalman paper [2] 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 extended in the early 1970s to a stochastic realization problem, where a Markovian model is obtained for a purely random process on the basis of covariance data [4], [5]. This stochastic realization technology, based on canonical correlation analysis, was extended in the early 1990s to processes that involve a measured (control) input, becoming known as subspace state-space identification. The identification work of the early nineties will be described later in this article.

1066-033X/06/$20.00©2006IEEE

The Åström-Bohlin paper [3] introduced into the control community the maximum likelihood framework that had been developed by time-series analysts for estimating the parameters of difference equation models [6], [7]. These models, which were known in the statistical literature by esoteric names such as ARMA (autoregressive moving average) or ARMAX model (autoregressive moving average with exogeneous inputs), later gave rise to the immensely successful prediction-error identification framework.

In 1970, Box and Jenkins published their book [8], which gave a major impetus to applications of identification. Indeed, [8] gave a rather complete recipe for identification, all the way from initial data analysis to the identification of a model and its validation. In the spirit of the time-series analysis methods of the time, [8] relied heavily on correlation analysis to determine model structure. For about 15 years, [8] remained the major high quality reference book on system identification. Important references of this time period also include the survey paper [9] as well as the special issue on system identification and time series analysis published by *IEEE Transactions on Automatic Control* in December 1974. The Åström and Eykhoff survey [9], which was used by many young researchers as a stepping stone for future work, explained the state of the art as much as it displayed some of the important open questions of the time. One of these questions was the identification of closed-loop systems, for which Hankel-based projection methods based on cross-correlation information had been shown to fail [10].

From the mid-1970s, the prediction-error framework came to completely dominate identification theory and, perhaps more importantly, identification applications. Much of the research activity focused on identifiability problems for both multivariable systems and closed-loop systems. The key identifiability problem for these two classes of systems was to find conditions on the parameterization and the experimental conditions under which the estimated model would converge to a unique representation of the true system. Just about all of the activity of that time period focused on the search for the *true system*, addressing questions of identifiability, convergence to the *true parameters*, statistical efficiency, and asymptotic normality of the estimated parameters.

Around 1976 the first attempts were made to view system identification as an approximation theory, in which one searches for the best possible approximation of the true system within a given model class [11]–[13]. The prevailing view in the identification community thus changed from a search for the true system to a search for and characterization of the *best approximation*. Hence, characterization of model errors became the focal point of research. For control engineers, the object of primary interest is the model, rather than the parameters, which are just a vehicle for describing the model. An important breakthrough came when Ljung introduced the concept of bias and variance error for an estimated transfer function [14].

The work on bias and variance analysis of identified models during the 1980s led, almost naturally, to a new perspective in which identification became viewed as a *design problem*. By understanding the effect of experimental conditions, model structure, and choice of criterion on the bias and variance errors of the identified model, it is possible to tune these design variables toward the objective for which the model is being identified [15], [16]. In this regard, the book [17] has had a profound impact on the engineering community of system identifiers. In particular, [17] squarely put forth the view of system identification as a design problem, in which the intended use of the model use plays a central role. This viewpoint clearly distinguishes the engineering approach to system identification from the statistical approach to system identification and time-series analysis, where the prevailing view is that the model must explain the data as well as possible.

The observation that model quality can be influenced by choosing design variables to reflect the objective for which the model was being built opened the way to a flood of new activity in the 1990s, which continues up to this day. The major application of this new paradigm is identification for the purpose of model-based control design. Consequently, identification for control has blossomed since its beginnings around 1990. Because identification for control embraces many aspects of identification and robust control theory, it has activated or reactivated research areas such as experiment design, closed-loop identification, frequency-domain identification, uncertainty estimation, and data-based robust control analysis and design.

THE MILESTONE PAPERS

Deterministic Realization Theory

In 1965, [2] provided a first solution to a challenging system-theoretic problem that became known as the *state-space realization problem*. It can be stated as follows.

Construct a minimal state-space realization

$$x_{t+1} = Ax_t + Bu_t,$$
$$y_t = Cx_t,$$

for the input-output model

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

described by its impulse response matrices (also called Markov parameters) $H_k \in \mathbb{R}^{p\times m}$.

The problem is to replace the infinite description

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

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

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

and A has minimal dimension. This problem can be divided into two parts,

namely, find the McMillan degree of $H(z)$ [18], which is then the minimal dimension of A, and compute the matrices A, B, C. The key tool for solving this problem is the Hankel matrix $\mathcal{H}$, whose factorization into the product of an infinite observability matrix and infinite controllability matrix is given by

$$\mathcal{H} = \begin{bmatrix} H_1 & H_2 & H_3 & H_4 & \dots \\ H_2 & H_3 & H_4 & H_5 & \dots \\ H_3 & H_4 & H_5 & H_6 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix}. \tag{1}$$

The Ho-Kalman realization method [2] is based on the following important properties: If the McMillan degree of $H(z)$ is n, then

1) rank $\mathcal{H} = n$, and
2) There exist $A \in \mathbb{R}^{n\times n}$, $B \in R^{n\times m}$, and $C \in \mathbb{R}^{p\times n}$ such that $H_k = CA^{k-1}B$ for all $k \geq 1$.

It took years of research to go from the theoretical results described in [2] to a numerically reliable realization algorithm [19], [20]. However, all of the key insights were present in the 1965 paper [2], and they were to have a profound impact on linear system theory in general and on realization and identification theory in particular.

The Maximum Likelihood Framework

In complete contrast to the state-space formulation of Ho and Kalman, the landmark 1965 paper of Åström and Bohlin [3] 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 zero-mean, unit-variance Gaussian random variables. The maximum likelihood method had been widely studied in mathematical statistics, including applications to time-series models [6]. The Åström-Bohlin paper [3] not only gave a complete algorithmic derivation of ML identification for ARMAX models but also summarized all of the analysis results that were available at that time, such as consistency, asymptotic efficiency, and asymptotic normality of the parameter estimates, persistence of excitation conditions on the input signal in connection with the order of the model, and model-order validation on the basis of the whiteness of the residuals.

The concepts and notation introduced in [3] have been with us for almost 40 years now. The following household notation of the identification community can be found in [3]:

- 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 $\hat{\lambda}^2 = (2/N)V(\hat{\theta})$.

The publication of [3] gave rise to a flurry of activity in parametric identification and established the basis for the prediction-error framework. The step from maximum likelihood to prediction error essentially consists of observing that, under the assumption of white Gaussian noise in the ARMAX model, maximization of the likelihood function of the observations is equivalent to minimizing the sum of the squared prediction errors. The prediction-error framework consists of adopting the minimization of a norm of the prediction errors as the criterion for parameter estimation, even when the probability distribution for the observations is unknown. As observed in [21], this idea had already been suggested by Gauss himself [22].

FROM DETERMINISTIC TO STOCHASTIC REALIZATION THEORY

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

Let $\{y_t\}$ be a zero-mean stationary vector stochastic process whose covariance sequence $\{R_k\}_{k=1}^{\infty}$ is given, with $R_k \triangleq E\{y_t y_{t-k}^T\}$. Then find a minimal Markovian representation for $\{y_t\}$ of the form

$$x_{t+1} = Ax_t + Gw_t \tag{2}$$

$$y_t = Cx_t + v_t, \tag{3}$$

where $\binom{w_t}{v_t}$ is a zero-mean stationary white noise sequence with covariance matrix

$$W = E\left\{\begin{pmatrix} w_t \\ v_t \end{pmatrix}\begin{pmatrix} w_t \\ v_t \end{pmatrix}^T\right\} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}. \tag{4}$$

This problem amounts to finding state-space matrices A, G, C, where the dimension of A is minimal, and the blocks Q, S, R of the covariance matrix W so that the covariance of the output of (2), (3) is given by $\{R_k\}_{k=1}^{\infty}$.

Observe that the covariance of the output y_t of the Markovian representation (2) and (3) is given by $R_k = CA^{k-1}N$, where $N = A\Pi C^T + GS$ for $k \geq 1$, and $R_0 = C\Pi C^T + R$, where $\Pi \triangleq E\{x_t x_t^T\}$ is the state covariance matrix.

The stochastic realization problem was studied intensively during the early 1970s in connection with innovations theory and spectral factorization theory [5], [23], [24]. The first step of the solution consists of observing that the Hankel matrix $\tilde{\mathcal{H}}$ of the covariance sequence can be factored as

$$\tilde{\mathcal{H}} = \begin{bmatrix} R_1 & R_2 & R_3 & R_4 & \dots \\ R_2 & R_3 & R_4 & R_5 & \dots \\ R_3 & R_4 & R_5 & R_6 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \begin{bmatrix} N & AN & A^2N & A^3N & \dots \end{bmatrix}, \tag{5}$$

where $R_k = CA^{k-1}N$ with N defined as above. The similarity between the factorizations (1) and (5) shows that the Ho-Kalman solution to the deterministic realization problem [5] can be used to determine the minimal dimension n as the rank of $\tilde{\mathcal{H}}$ and to compute C, A, and N from the factorization of $\tilde{\mathcal{H}}$.

There are various ways to perform the second step, which consists of computing the remaining elements G, Q, R, S of the Markovian representation (2) from C, A, N and the output variance R_0. One way to compute these elements is to use a specialized Markovian representation, known as the *innovations model*, of the form

$$\xi_{t+1} = A\xi_t + K\varepsilon_t, \tag{6}$$

$$y_t = C\xi_t + \varepsilon_t, \tag{7}$$

where $\{\varepsilon_t\}$ is a stationary vector white noise sequence with covariance matrix $\Sigma = E\{\varepsilon_t\varepsilon_t^T\}$.

Denoting $\hat{\Pi} = E\{\xi_t\xi_t^T\}$ and requiring that the covariance sequence of the output y_t of this innovations model (6), (7) is $\{R_k\}$ yields three constraints on the unknown quantities $\hat{\Pi}$, K, and Σ of the form

$$\hat{\Pi} = A\hat{\Pi}A^T + K\Sigma K^T, \tag{8}$$

$$N = A\hat{\Pi}C^T + K\Sigma, \tag{9}$$

$$R_0 = C\hat{\Pi}C^T + \Sigma. \tag{10}$$

Observe that the Lyapunov equation (8) follows directly from (6), while constraints (9) and (10) are imposed by matching the output covariance R_k. Substituting K and Σ from (9) and (10) into (8) yields a Riccati equation for $\hat{\Pi}$ of the form

$$\hat{\Pi} = A\hat{\Pi}A^T + (N - A\hat{\Pi}C^T) \times (R_0 - C\hat{\Pi}C^T)^{-1}(N - A\hat{\Pi}C^T)^T. \tag{11}$$

By solving (11) for $\hat{\Pi}$, the remaining elements K and Σ are obtained from $\hat{\Pi}$ using (9) and (10). The stochastic realization problem, as well as the properties of its associated Riccati equation (11), was extensively studied through the late 1960s and 1970s in connection with the spectral factorization problem; see [24]. A comprehensive treatment of the stochastic realization problem can be found in [25], while a tutorial presentation is given in [26].

An interesting aspect of the stochastic realization solution given by Akaike in [27] is the definition of the state of the innovations model as the set of canonical correlations obtained by projecting the vector space spanned by the present and future outputs onto the vector space spanned by the present and past outputs. Stated otherwise, the state of the innovations model forms a basis for predicting future output signals from past signals. These canonical correlations can be computed by a singular value decomposition. This insight formed the basis for later work on subspace identification and gave rise to extensive studies of the interface between these spaces of future and past observations [25], [28].

Another outcome of the stochastic realization and innovations theories of the 1970s was the covariance equivalence between the Markovian realization (2), (3) and its innovations realization (6), (7). This equivalence implies that a Markovian model (2), (3), which may have been constructed as a first principles model based on the laws of physics, and where y_t is driven by two independent white noise sources w_t and v_t (thus with $S = 0$), can be rewritten as the Markovian innovations model (6), (7) driven by a single white noise source ε_t that yields the same output covariance $\{R_k\}$ as the Markovian model. This equivalence also applies to models with deterministic inputs, that is, every state-space model

$$x_{t+1} = Ax_t + Bu_t + Gw_t,$$

$$y_t = Cx_t + v_t,$$

possibly derived from physics laws, can be rewritten for all u_t as an innovations model driven by a single noise source ε_t, that is,

$$\xi_{t+1} = A\xi_t + Bu_t + K\varepsilon_t, \tag{12}$$

$$y_t = C\xi_t + \varepsilon_t. \tag{13}$$

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

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

where

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

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

Observe that the transfer function matrices $G(z)$ and $H(z)$ have the same poles, and that the independent term in the series expansion of $H(z)$ is the identity. The model (12), (13) is therefore equivalent to the ARMAX model

$$A(z^{-1})y_k = B(z^{-1})u_k + C(z^{-1})\varepsilon_k, \tag{14}$$

where $A(z^{-1})$, $B(z^{-1})$, and $C(z^{-1})$ are now matrix polynomials in the delay operator z^{-1} with $A(z^{-1})$ and $C(z^{-1})$ monic, that is,

$$A(z^{-1}) = I + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a},$$

$$B(z^{-1}) = b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b},$$

$$C(z^{-1}) = I + c_1 z^{-1} + \cdots + c_{n_c} z^{-n_c}.$$

This theory established a link between a Markovian model obtained from first principles modeling driven by possibly independent noise sources w_t and v_t, the equivalent state-space innovations model driven by a single noise source ε_t, and the corresponding ARMAX model used in maximum likelihood or prediction-error identification. The theory also gave a solid theoretical justification for the use of ARMAX models for representing stationary linear Markov processes, whether these processes are physically driven by one or several noise sources.

THE GOLDEN YEARS: 1975–1985

The Big Cleanup

The years 1975–1985 saw frantic activity in system identification in the engineering

community. The methods based on a prediction-error criterion together with input-output models completely took over the field, at the expense of methods based on realization theory. Their theoretical superiority over stochastic realization methods was based on the statistical properties of the parametric estimates: not only are prediction-error methods asymptotically efficient (since their covariance achieves the Cramér-Rao bound), but the asymptotic accuracy can also be evaluated. The main reason for the growing appeal of prediction-error methods, however, was that increased computer speed and the development of special purpose identification software made it more feasible to iteratively minimize a cost criterion over a range of possible model structures. During this period many authors put their names on new combinations of model structures and methods, with claims about the supremacy of their new combination over existing methods. New "methods" appeared constantly in the scientific journals.

Some solid cleaning was required, and it was one of L. Ljung's major contributions to perform this function. His contribution was to separate two independent concepts, the choice of a parametric model structure, which provided a vehicle for computing predictions and hence parameter-dependent prediction errors, and the choice of an identification criterion, which was a nonnegative function of the prediction errors and hence of the parameter vector [29]. All existing parametric identification methods could then be seen as particular cases of this prediction-error framework.

To do this, [29] introduced the generic input-output model structure

$$y_t = G(z, \theta)u_t + H(z, \theta)e_t, \qquad (15)$$

where $G(z, \theta)$ and $H(z, \theta)$ are parameterized rational transfer functions and e_t is white noise. All commonly used model structures were special cases of the generic structure (15). To estimate θ from (15), one can derive the parameter-dependent one-step-ahead prediction $\hat{y}_{t|t-1}(\theta)$ and hence the one-step-ahead prediction error $\varepsilon_t(\theta) = y_t - \hat{y}_{t|t-1}(\theta)$. Next, given a set Z^N of N data and hence of N prediction errors, one can define a criterion

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} l(\varepsilon_t(\theta)), \qquad (16)$$

where $l(\cdot)$ is a nonnegative scalar-valued function. Minimizing $V_N(\theta, Z^N)$ with respect to θ over a domain $\mathcal{D}_\theta$ then yields the parameter estimate

$$\hat{\theta}_N = \arg\min_{\theta \in \mathcal{D}_\theta} V_N(\theta, Z^N). \qquad (17)$$

This work culminated in the publication in 1987 of [17], which has become the standard reference book on system identification. The book has had an enormous impact on the engineering community, both as a theoretical basis and as a guide for applications. Its usefulness for applications has been greatly enhanced by the simultaneous production by Ljung in 1987 of the Matlab identification toolbox. Ljung's book was complemented by that of Stoica and Söderström [30], 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.

Breakthroughs for MIMO and Closed-Loop Systems

During 1975–1985, theoretical breakthroughs were made in two directions. The first was the elucidation of the manifold structure of multi-input, multi-output (MIMO) systems. The second consisted of identifiability results for linear systems under closed-loop conditions. On the basis of earlier observations made in the context of stochastic realization theory and of identification from spectral or covariance data [9], [10], the view prevailed in the identification community that data collected in closed-loop operation would lead to a biased estimate of the plant model. Finding conditions under which the model can be consistently estimated from closed-loop data was therefore a problem of great theoretical and practical interest.

Many authors contributed to the solution of both problems. The manifold structure of MIMO systems, based on Kronecker (or structure) indices, was elucidated in [31] and [32]. It was shown that there is no unique state-space or ARMA parameterization that can represent all linear MIMO systems with m inputs and p outputs. However, each $p \times m$-dimensional rational system can be uniquely represented by a canonical parameterization whose structure depends on a finite set of integers, the Kronecker indices. In addition, the set of all $p \times m$-dimensional rational systems can be covered by a finite number of pseudocanonical or overlapping parameterizations. Subsequently, many authors worked on methods for estimating the Kronecker indices while others studied the relationship between the canonical (or pseudocanonical) forms in state-space and ARMA form [33]–[37]. The thinking was that for the parameter estimation problem to be well conditioned, it was important to choose a parameterization in which the true system could be represented by a single parameter vector.

As for the identifiability of closed-loop systems, one of the earliest solutions was provided by the Swedish trio made up of Gustavsson, Ljung, and Söderström, all Ph.D. students of K.J. Åström at the time [38]. Over a four-year period the trio jointly published no less than six important papers on various aspects of system identification. They showed that, in many situations of practical interest, direct application of prediction-error identification 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, as well as the joint input-output method

in which a model is first estimated for the joint vector process made up of the input and output vectors [39]–[42].

System Identification Viewed as Approximation

For most of the 1960s and 1970s, the prevailing assumption was that the system $\mathcal{S}$ was in the model set $\mathcal{M}$, that is, $\mathcal{S} \in \mathcal{M}$. 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-1970s, the first attempts were made to view system identification in the context of approximation [11]–[13]. This step marked the beginning of a 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.

The Birth of θ^*

In statistics, the natural way to analyze estimation errors is through the concepts of bias and variance errors. 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 object of interest is the transfer function, not the parameters that are used to represent it. By defining θ^* as

$$\theta^* = \arg\min_{\theta \in \mathcal{D}_\theta} \lim_{N\to\infty} E\{V_N(\theta)\}$$

and observing that, under reasonable conditions, the parameter estimate $\hat{\theta}_N$ converges to θ^*, that is, $\lim_{N\to\infty} \hat{\theta}_N = \theta^*$ [11], Ljung introduced the following decomposition of the total transfer function error at the frequency ω:

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

Within this framework, approximate asymptotic expressions for the transfer function variance, as well as integral expressions for the transfer function bias for the case in which a least squares prediction-error identification criterion is used, were subsequently derived [14], [15]. By assuming that the model order n increases with the number N of data in a specific way, an approximation was obtained in [14] for the asymptotic variance of the estimated transfer functions at a frequency ω. This approximation is given by

$$\begin{aligned} &E\begin{pmatrix} G(e^{j\omega}, \hat{\theta}_N) - G_0(e^{j\omega}) \\ H(e^{j\omega}, \hat{\theta}_N) - H_0(e^{j\omega}) \end{pmatrix} \\ &\quad \times \begin{pmatrix} G(e^{-j\omega}, \hat{\theta}_N) - G_0(e^{-j\omega}) \\ H(e^{-j\omega}, \hat{\theta}_N) - H_0(e^{-j\omega}) \end{pmatrix}^T \\ &\quad \cong \frac{n}{N}\Phi_v(\omega) \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{eu}(\omega) & \sigma^2 \end{bmatrix}^{-1}, \end{aligned} \tag{18}$$

where G and H are the parameterized transfer function models of order n defined in the generic model (15), $\hat{\theta}_N$ is the parameter estimate, N is the number of data, $\Phi_u(\omega)$ is the input spectrum, $\Phi_v(\omega)$ is the disturbance spectrum, and $\Phi_{ue}(\omega)$ is the cross-spectrum between the input and the white noise e driving the noise model.

As for the asymptotic bias, the expression [15]

$$\begin{aligned} \theta^* = \arg\min_\theta \int_{-\pi}^{\pi} &[|G_0(e^{j\omega}) \\ &- G(e^{j\omega}, \theta)|^2 \Phi_u(\omega) \\ &+ \Phi_v(\omega)] \left| \frac{D(e^{j\omega})}{H(e^{j\omega}, \theta)} \right|^2 d\omega, \end{aligned} \tag{19}$$

valid for open-loop identification, provides insight into the way the choice of the model parameterization $G(z, \theta)$ and $H(z, \theta)$ influences the bias error of the identified model. Here $D(z)$ is a data filter applied to all input and output data. These variance and bias formulas (18), (19), which have since been improved on, provide insight into the way the model structure, the data filter, the input spectrum, the number of data, the feedback controller (if any) influence the bias and variance error of the estimated model. These expressions were to become the cornerstone of the next major phase of the development of system identification: the design phase.

IDENTIFICATION AS A DESIGN PROBLEM

If identification is viewed as approximation and if the estimated model is to be used for a specific purpose (as is most often the case), then it makes sense to control the model error so that it does not excessively penalize the goal for which the model is being built. This idea is behind the concept of goal-oriented identification. To formulate the identification problem as a *goal-oriented design problem* requires a good understanding of the connection between identification design (experiment design, choice of model structure, choice of criterion) and model quality on the one hand, and the effect of model quality on the intended model application on the other hand.

The first few steps of this paradigm were laid in [15]–[17]. This 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*.

SYSTEM IDENTIFICATION IN THE 1990s

At the triennial IFAC Symposium on System Identification, held in Budapest in 1991, there was a feeling that most of the important problems in system identification had been solved and that the golden age of identification was over. That prediction proved to be wrong. The research on system identification was pulled all through the 1990s essentially by two catalysts, whose first feeble signs emerged around 1990: subspace-based identification and identification for control. In

addition, new research activity took place in frequency-domain identification, closed-loop identification, the use of orthogonal basis functions for identification, the development of new methods for quantifying model uncertainty, errors-in-variables identification, and the identification of nonlinear systems. Thus, dire predictions about the disappearance of identification as an active research topic proved to be completely wrong.

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 1980s. Even though the manifold structure of MIMO systems had been completely characterized in the late 1970s, the practical problem of identifying MIMO systems remained wide open. Indeed, estimating the structure indices that characterize the parameterizations of multivariable systems remained tricky and led to ill-conditioned numerical procedures. Thus, there was great incentive to develop simple but suboptimal procedures based on the numerically robust singular value decomposition and least squares techniques, which bypass the need for estimating 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 hurdle was that the projection methods developed by Akaike [27], which were based on canonical correlation analysis, were not easily extendable to output data that contained, besides the stochastic components, a contribution due to a measured input. In the early 1990s, several research teams managed to crack this nut, providing closely related solutions [43]–[47]. These first solutions paved the way for research on the properties of subspace-based identification, their connection with stochastic realization theory, their application to closed-loop systems, and on improved numerical procedures [28], [48]–[51].

Identification for Control

Identification for control has been the major application for the new paradigm of system identification as a design problem. The reasons for the heavy research investment in identification for control are many: i) in the systems and control community of system identification, control is often the main motivation for model building; ii) it has been observed that high performance control can often be achieved with simple models, provided some basic dynamical features of the system are accurately captured; iii) although robust control theory based on models and uncertainty sets had been developed during the 1980s, the models and uncertainty sets were not data based for lack of a proper theory; iv) research on identification for control provided iterative model and controller tuning tools that were intuitive, practical, and easy to implement by 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 [52] at the 1991 IFAC Symposium on System Identification addressed many of the key issues concerning the interplay between identifying a reduced-order model and designing a controller from such a model; however, [52] was more an agenda for research than a presentation of solutions. Indeed, in 1990 there was limited understanding of the interplay between system identification and robust control. The two theories had been developed by separate communities that had had little contact with one another.

As noted above, the robust control community had developed a robust analysis and design theory based on uncertainty descriptions that were not based on data, but rather on prior assumptions. The identification community had delivered bias and variance error descriptions that were not explicit, and thus not transferable 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 in terms of the qualities that a model must possess (or, conversely, the plant-model errors that are acceptable) if the model is to be used for designing a controller that must achieve a given level of performance on the plant. An important workshop, held in Santa Barbara in 1992, brought together members of these two communities and played a key role in pinpointing the gaps between the two theories and in establishing a dialogue [53].

An apparent cause for discrepancy between robust control theory and system identification was the great "hard-versus-soft bound debate," as it was called in [52]. Robust control theory, as it was available around 1990, was essentially based on hard-bound assumptions on the disturbances, whereas prediction-error identification had been built on stochastic assumptions on the disturbances. The desire to use existing robust control theory tools in the context of models obtained from data led to the development of identification techniques that would deliver hard bounds on model errors, such as set membership and worst-case identification techniques [54]. However, it was shown in [55] that these techniques amounted to assuming that the input to the system is maximally correlated with the disturbance. With such a hard-bound paradigm on the noise, convergence to the true parameter values could only be obtained under the assumption that this hard bound tends to zero, leading to conservative results. As it happens, the existence of a maximal correlation between the disturbance and the input signal contradicts the definition of a disturbance, which must have nothing to do with the

particular input that is applied to the system. It was shown in [55] that if the input is chosen such that the sample cross correlation between input and disturbance tends to zero, which is obtained by using a "mixing" input, then it is immaterial whether the disturbance is stochastic or deterministic; the conventional identification results, such as parameter convergence, remain the same.

The conservatism of the control designs achieved within the worst-case framework, as well as their computational complexity, have led to the development of a probabilistic robust design framework based on randomized algorithms [56]–[58]. These algorithms guarantee that a certain design specification is met with a given probability. One advantage of this approach is that the robustness margin can often be increased by a considerable amount at the expense of a small risk, thereby circumventing the conservatism of the worst-case approach.

In July 1992, *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, this 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 rather with estimation of uncertainty sets without considering control-oriented design issues. A few papers in that special issue did address the joint identification and control design paradigm [59], [60], producing one of the first key results in identification for control, namely, the necessity of an iterative scheme for the design of a control-oriented nominal model; this observation had in fact been first made in [61].

The first half of the 1990s produced a series of results on the design of control-oriented nominal models. These results were produced by teams who used their favorite combinations of identification criteria and control design criteria [62]–[66]. This work confirmed the necessity of using an iterative scheme of model updates and controller updates, and produced significant evidence about the advantages of performing the identification in closed loop, rather than in open loop, when the model is to be used for designing a new controller. Of course, the closed-loop experimental conditions that produce the desired plant input signal spectrum can always be mimicked by an open-loop experiment whose input spectrum matches the required closed-loop input spectrum. This input spectrum contains the sensitivity function, which depends on the unknown true system. The advantage of the closed-loop experiment is that the frequency weighting by the sensitivity function is automatically present in the identification criterion, at least with the present controller, which should not be too different from the new controller. The requirement to keep successive experimental conditions close to one another during the iterations has led to the concept of cautious model and controller updates [67].

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

- 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 with successive controllers; this development was welcomed by 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 to design a better controller.

The early work on identification for control focused on control-oriented identification criteria. This objective amounts to constructing a nominal model whose bias error distribution is tuned for control design. Hence 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, namely, estimation of control-oriented uncertainty sets [68]–[71]. The idea is that since one can manipulate the shape of the model uncertainty set by the choice of 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 [72]. Another view is that the worst-case performance achieved by an 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 established and to reopen research questions that had been considered to be 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. A lively 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 a *near-optimal restricted-complexity model* introduced in [73], a must-read for any researcher in identification for control. Finally, identification for control has led to the recent rebirth of optimal experiment design, a subject that had been active in the 1970s but which had been almost completely abandoned since then. The overview paper [74] presents up-to-date results on the use of optimal experiment design in the context of identification for control, a topic that is also present in [73].

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

Quantification of Model Uncertainty

The demands of robust control theory for adequate uncertainty sets triggered interest in the estimation of uncertainty sets from data. It is fair to say that most of the robust control theory developed in the 1980s had been based on a priori assumed uncertainty bounds on model errors and on the noise. In the context of system identification, data-based estimation of model errors and noise properties from data is of interest in its own right; indeed, a reputable engineer should never deliver a model without a statement about its error margins. The activity on estimation of uncertainty sets was often erroneously put under the umbrella of identification for control, since in most of this work the control objective was not taken into account in the identification design.

A wide range of techniques and identification criteria were developed to provide error bounds for identified models using time-domain, frequency-domain, H_∞, l_1, probabilistic, worst-case, set-membership, and other methods [53], [54], [70], [75]–[80]. The rationale was to produce uncertainty sets from data that would be compatible with the available robust control analysis and design tools. The price to be paid for achieving this objective is that these techniques often led to conservative uncertainty sets, often caused by overbounding. An alternative route was taken in [72] and [81], where robust control theory was developed for uncertainty sets obtained by the prediction-error identification method. The relationship between identified uncertainty sets and robust control analysis and design is examined in [82].

Closed-Loop Identification Revisited

The work of the 1970s on closed-loop identification had focused entirely on the question of identifiability, which amounts to finding conditions under which the parameter estimates converge to the true parameters when the system is in the model set. Once that question had been resolved, research on closed-loop identification essentially stopped. In particular, there was no investigation of the influence of the experimental conditions on bias-error distribution in the case of restricted-complexity models as well as on asymptotic variance.

One of the important lessons that emerged from the study of the interplay between identification and control design is the benefit of closed-loop identification when the model is to be used for designing a new controller with better performance. Until the late 1980s, the common view was that it was preferable to avoid closed-loop identification. 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. In the ideal context of optimal experiment design with full-order models, the optimality of closed-loop identification was established based on variance formulas when the model is to be used for control design with a noise-rejection objective [16], [83], [84].

This observation triggered activity in the design of closed-loop identification methods, where the main goal is to obtain a better handle on the bias error in closed-loop identification [85]–[88].

Optimal Experiment Design for Identification and Control

In the 1970s, optimal input design for system identification was an active area of research, with various quality measures used to define optimality [89]–[91] The questions at that time addressed open-loop identification, and the quality measures were scalar measures of the parameter covariance matrix P_θ. After a hiatus of about 15 years, the paradigm of identification as a design problem gave optimal input design a new lease of life. The work on the connection between model uncertainty sets obtained by identification and corresponding sets of robust controllers put this subject in the limelight again [84], [92]–[96]. The emphasis in this research activity 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.

Frequency-Domain Identification

Another area of activity in the 1990s was frequency-domain identification. Frequency-domain identification, based on spectral analysis, had been the main tool for transfer function identification until the advent, in the 1960s, of the prediction-error method based on parametric models identified in the time domain. As a consequence, a gap developed between the two approaches, and there was no more than polite attention paid to frequency-domain identification by the prediction-error identification community. In 1981, Ljung and Glover [97] established some bridges over this gap, concluding that "the two approaches are complementary rather than rivaling."

Things changed drastically at the end of the 1980s, with a convergence of efforts arising from different directions. During the 1980s, the robust control community developed extensive analysis and design tools in the frequency

domain; thus, there was great demand for tools for obtaining frequency-domain models and, even more importantly, frequency-domain uncertainty descriptions from data. This need led, at the end of the 1980s, to the development of interpolation techniques that used noisy pointwise frequency-domain transfer-function measurements as their data [98]–[100].

At about the same time, identification methods based on frequency-domain data were developed for estimating flexible modes in mechanical structures, with applications mainly in the aerospace and automotive industry [101]. This work sparked the development of a range of methods, often based on state-space models [102]–[104]. In particular, [104] presents the optimal Kalman identification (OKID) method, where a Kalman filter innovations model is identified directly from data, which can be either in the time or frequency domain, while [102] compares the performance of four state-space-based methods for estimating the eigenmodes of flexible structures.

Independently, Pintelon and Schoukens developed frequency-domain identification techniques using transfer-function models, essentially based on the maximum likelihood principle [105]. 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 sinusoidal or multisine data sequences [106]. Their work later converged with that of the modal analysis community. Aware of the interest for frequency-domain identification emanating from the robust control community, Schoukens and Pintelon continued their work on frequency-domain identification through the 1990s using periodic excitation and maximally informative input signals. Their book [107] provides a comprehensive treatment of frequency-domain identification.

The activity in frequency-domain identification during the 1990s has closed the gap that had existed between time- and frequency-domain methods. In [108], Ljung updates the comparative analysis of [97] between these two approaches to system identification on the basis of the new understanding gained about frequency-domain identification over the last 20 years.

Identification of Nonlinear Systems

Identification theory for nonlinear systems is almost as old as identification theory for linear systems. In many different fields of application, the structure of the nonlinear system is obtained from physical laws describing the various components as well as interconnection laws describing the interconnection structure. Identification then reduces to estimating unknown parameters appearing in the model structure on the basis of measured signals. The analysis and solution of a nonlinear identification problem with known model structure but unknown parameters parallels the analysis and solution of linear identification problems. For both problems, one first needs to check whether the parameters that one seeks to estimate are identifiable, which essentially amounts to checking whether the predicted outputs are sensitive to these parameters.

For a long time, attempts were made to go beyond the identification of nonlinear systems with known structure by introducing special classes of nonlinear black-box models such as Wiener, Hammerstein, and Wiener-Hammerstein models. Black-box models refer, as in the linear case, to model structures that have not been derived from physics laws and whose parameters therefore have a priori no physical significance. When some parts of the model structure are obtained by modeling using physics laws, the term grey-box modeling is used. A Wiener model is a linear dynamic model with a static nonlinearity at the output, while a Hammerstein model has the static nonlinearity at the input; the Wiener-Hammerstein model combines both nonlinearities. The search for universal classes of nonlinear black-box models gathered steam in the 1980s with the introduction of broader classes of basis functions, such as splines, neural networks, wavelets, and radial basis functions. The risk was great to repeat the chaotic development of model structures and identification methods that had characterized linear system identification during the prior decade. Before chaos had a chance to take over, a collective effort set a common framework for identifying nonlinear black-box models [109], [110]. In this common framework, the nonlinear structures are seen as a concatenation of a mapping from observed data to a regressor vector, followed by a nonlinear mapping from regressor vector to output space, where the latter mapping is typically expressed as an expansion of parameterized basis functions.

The last few years have seen a renewed interest in the identification of nonlinear systems, with special emphasis on the detection of nonlinearities, as well as the estimation and properties of the best linear approximation [111]. The area of nonlinear identification is vast and difficult, and it will undoubtedly keep the identification community busy for many years to come. To highlight the scope and complexity of the topic, let us just quote from [1]: "'identification of nonlinear systems' is like a statement about 'non-elephant zoology.'"

Other Areas of Recent Activity

Most of the research activity on system identification over the last 15 years has focused on subspace methods, on identification of nonlinear systems, or on identification for control, as well as on topics that are motivated by identification for control, such as estimation of model uncertainty, closed-loop identification, optimal experiment design for control, or frequency-domain identification.

But progress has also been made on other fronts. Alternative basis functions (other than the shift operator) have been considered for representing 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 about the system is available but also led to improved formulas for estimating the variance of black-box transfer function models [112]–[114]. The book [115] surveys this field.

Errors-in-variables identification has been another active topic during the last decade. In the errors-in-variables framework, it is assumed that not only the output of the dynamical system is measured with noise but also the input measurements are noisy. This framework makes the identification problem much harder and, in particular, raises nontrivial identifiability questions. Much progress has been made on this topic since the early paper [116]. A range of techniques have been applied to the problem, including maximum likelihood, total least squares, and instrumental variables. A rather recent survey can be found in [117], while [118] compares the performance of three different methods for errors-in-variables identification on a series of data sets.

CONCLUDING REMARKS

In this article I have attempted to present the way I have seen the evolution of system identification over the last few decades, both as a student and as an actor in the field. I have tried to illustrate both the role of a few individuals whose milestone contributions opened the way for new insights and developments, as well as the importance of the socio-technological environment created by the evolution of technology or by developments in neighboring sciences that create a demand for new scientific developments.

The evolution of system identification in our engineering field beautifully illustrates how some avenues of research remain in a dormant stage for long periods of time, only to reemerge years later as a result of new developments in a neighboring field. As shown in this article, examples include the phasing out of Hankel-based realization theory around 1975 and its reappearance in the late 1980s under the name subspace identification, the halt of any significant research on closed-loop identification around 1980 and its resurrection in the early 1990s in the context of approximate models for control, the disappearance of optimal experiment design research for about 20 years and its reemergence around 2003 as a tool for the definition of control-oriented identification design.

In conclusion, I believe that system identification is such a fundamental discipline, with such wide-ranging and cross-disciplinary applications, that it will remain an active and exciting research area for many years to come.

ACKNOWLEDGMENTS

This article presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with the author.

AUTHOR INFORMATION

Michel Gevers obtained an electrical engineering degree from the Université Catholique de Louvain, Belgium, in 1968 and the Ph.D. from Stanford University in 1972. He is currently a professor in the Department of Mathematical Engineering at the Université Catholique de Louvain, where he is the coordinator of the Belgian Interuniversity Network on Dynamical Systems and Control. He has spent long-term visits at the University of Newcastle, Australia, the Technical University of Vienna, and the Australian National University. He is an associate editor at large of the *European Journal of Control* and an associate editor of *Mathematics of Control, Signals, and Systems*. He has published over 230 archival and conference papers and is a coauthor of *Adaptive Optimal Control—The Thinking Man's GPC* and *Parametrizations in Control, Estimation and Filtering Problems: Accuracy Aspects*. He was president of the European Union Control Association from 1997 to 1999 and a vice president of the IEEE Control Systems Society in 2000 and 2001. He is an IFAC fellow, a Fellow of the IEEE, and a distinguished member of the IEEE Control Systems Society.

REFERENCES

[1] M. Deistler, "System identification and time series analysis: Past, present, and future," in *Proc. Stochastic Theory and Control, Festschrift for Tyrone Duncan,* Kansas, 2002, pp. 97–108.
[2] B.L. Ho and R.E. Kalman, "Ective construction of linear state-variable models from input-output functions," *Regelungstechnik*, vol. 12, pp. 545–548, 1965.
[3] K.J. Åström and T. Bohlin, "Numerical identification of linear dynamic systems from normal operating records," in *Proc. IFAC Symp. Self-Adaptive Systems*, Teddington, U.K., 1965, pp. 96–111.
[4] P. Faurre, "Réalisations markoviennes de processus stationnaires," Rapport La-boria No. 13, IRIA, Rocquencourt, France, Tech. Rep. 1973.
[5] H. Akaike, "Stochastic theory of minimal realization," *IEEE Trans. Automat. Control*, vol. 26, pp. 667–673, Dec. 1974.
[6] T.C. Koopmans, H. Rubin, and R.B. Leipnik, *Measuring the Equation Systems of Dynamic Economics* (Cowles Commission Monograph, vol. 10, T.C. Koopmans, Ed.). New York: Wiley, 1950.
[7] E.J. Hannan, *Time Series Analysis*. New York: Methuen, 1960.
[8] G.E.P. Box and G.M. Jenkins, *Time Series Analysis, Forecasting and Control*. Oakland, CA: Holden-Day, 1970.
[9] K.J. Åström and P. Eykho, "System identification—A survey," *Automatica*, vol. 7, pp. 123–162, 1971.
[10] H. Akaike, "Some problems in the application of the cross-spectral method," in *Spectral Analysis of Time Series,* B. Harris, Ed. New York: Wiley, 1967, pp. 81–107.
[11] L. Ljung, "On consistency and identifiability," *Math. Program. Study*, vol. 5, pp. 169–190, 1976.
[12] B.D.O. Anderson, J.B. Moore, and R.M. Hawkes, "Model approximation via prediction error identification," *Automatica*, vol. 14, pp. 615–622, 1978.
[13] L. Ljung and P.E. Caines, "Asymptotic normality of prediction error estimators for approximative system models," *Stochastics*, vol. 3, pp. 29–46, 1979.
[14] L. Ljung, "Asymptotic variance expressions for identified black-box transfer func-tion models," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 834–844, 1985.
[15] B. Wahlberg and L. Ljung, "Design variables for bias distribution in transfer function estimation," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 134–144, 1986.
[16] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application," *Automatica*, vol. 22, pp. 543–554, 1986.
[17] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[18] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
[19] L.S. de Jong, "Numerical aspects of realization algorithms in linear system theory," Ph.D. dissertation, Eindhoven Univ. Technology, Eindhoven, The Netherlands, Tech. Rep., 1975.
[20] L.S. de Jong, "Numerical aspects of recursive realization algorithms," *SIAM J. Contr. Optim.*, vol. 16, pp. 646–659, 1978.
[21] K.J. Åström, "Maximum likelihood and prediction error methods," *Automatica*, vol. 16, pp. 551–574, 1980.

[22] C. Gauss, *Teoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, ser. Reprinted translation: *Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections.* New York: Dover, 1809.
[23] M. Gevers and T. Kailath, "An innovations approach to least-squares estimation, part vi: Discrete-time innovations representations and recursive estimation," *IEEE Trans. Automat. Contr.*, vol. 18, pp. 588–600, 1973.
[24] P. Faurre, "Stochastic realization algorithms," in *System Identification: Advances and Case Studies* R.K. Mehra and D.G. Lainiotis, Eds. New York: Academic, 1976.
[25] P.E. Caines, *Linear Stochastic Systems*. New York: Wiley, 1988.
[26] M. Gevers and W.R. Wouters, "An innovations approach to the discrete-time stochastic realization problem," *J. A*, vol. 19, no. 2, pp. 90–110, 1978.
[27] H. Akaike, "Markovian representation of stochastic processes by canonical variables," *SIAM J. Contr. Optimin.* vol. 13, pp. 162–173, Dec. 1975.
[28] T. Katayama and G. Picci, "Realization of stochastic systems with exogenous inputs and subspace identification methods," *Automatica*, vol. 35, pp. 1635–1652, Oct. 1999.
[29] L. Ljung, "Convergence analysis of parametric identification methods," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 770–783, Oct. 1978.
[30] T. Söderström and P. Stoica, *System Identification* Hemel Hempstead, Hertfordshire, U.K.: Prentice-Hall Int., 1989.
[31] M. Hazewinkel and R.E. Kalman, *On Invariants, Canonical Forms and Moduli for Linear Constant, Finite-Dimensional Dynamical Systems* (Lecture Notes in Economics and Mathematical Systems, vol. 131). Berlin, Germany: Springer-Verlag, 1976, pp. 48–60.
[32] J.M.C. Clark, "The consistent selection of parametrizations in systems identification," in *Proc. Joint Automatic Control Conf.*, Purdue Univ., West Lafayette, IN, 1976, pp. 576–580.
[33] M. Deistler and E. Hannan, "Some properties of the parametrization of ARMA systems with unknown order," *Multivariate Anal.*, vol. 11, pp. 474–484, 1981.
[34] A.J.M. Van Overbeek and L. Ljung, "On-line structure selection for multivari-able state-space models," *Automatica*, vol. 18, pp. 529–544, 1982.
[35] G. Picci, "Some numerical aspects of multivariable systems identification," *Math. Program. Studies*, vol. 18, pp. 76–101, 1982.
[36] J. Rissanen, "Estimation of structure by minimal description length" *Circuits, Syst. Sig. Processing*, vol. 1, pp. 395–406, 1982.
[37] V. Wertz, M. Gevers, and E.J. Hannan, "The determination of optimum structures for the state-space representation of multivariable stochastic processes," *IEEE Trans. Automat. Contr.*, vol. 27, no. 6, pp. 1200–1211, 1982.
[38] T. Söderström, L. Ljung, and I. Gustavsson, "Identifiability conditions for linear multivariable systems operating under feedback," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 837–840, 1976.
[39] T.S. Ng, G.C. Goodwin, and B.D.O. Anderson, "Identifiability of MIMO linear dynamic systems operating in closed loop," *Automatica*, vol. 13, pp. 477–485, 1977.
[40] I. Gustavsson, L. Ljung, and T. Söderström, "Identification of processes in closed loop-identifiability and accuracy aspects," *Automatica*, vol. 13, pp. 59–75, 1977.
[41] K.S. Sin and G.C. Goodwin, "Checkable conditions for identifiability of linear systems operating in closed loop," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 722–729, 1980.
[42] B.D.O. Anderson and M. Gevers, "Identifiability of linear stochastic systems operating under linear feedback," *Automatica*, vol. 18, no. 2, pp. 195–213, 1982.
[43] W.E. Larimore, "Canonical variate analysis in identification, filtering, and adaptive control," in *Proc. 29th IEEE Conf. Decision and Control*, Honolulu, HI, 1990, pp. 596–604.
[44] J.N. Juang, M. Phan, L.G. Horta, and R.W. Longman, "Identification of observer Kalman filter Markov parameters—Ttheory and experiments," *J. Guidance, Contr. Dynam.*, vol. 16, no. 2, pp. 320–329, 1993.
[45] P. Van Overschee and B. De Moor, "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.
[46] M. Verhaegen, "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61–74, 1994.
[47] M. Viberg, "Subspace-based methods for the identification of linear time-invariant systems," *Automatica*, vol. 31, no. 12, pp. 1835–1851, 1995.
[48] N. Chui and J. Maciejowski, "Realization of stable models with subspace methods," *Automatica*, vol. 32, no. 11, pp. 1587–1595, Nov. 1996.
[49] A. Chiuso and G. Picci, "Geometry of oblique splitting subspaces, minimality and Hankel operators," in *Directions in Mathematical Systems Theory and Optimization*, A. Rantzer and C. Byrnes, Eds. New York: Springer-Verlag, 2002.
[50] A. Chiuso and G. Picci, "Consistency analysis of some closed-loop subspace identification methods," *Automatica*, vol. 41, no. 3, pp. 377–391, Mar. 2005.
[51] S. Qin, W. Lin, and L. Ljung, "A novel subspace identification approach with enforced causal models," *Automatica*, vol. 41, no. 12, pp. 2043–2053, Dec. 2005.
[52] M. Gevers, "Connecting identification and robust control: A new challenge," in *Proc. IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, Hungary, 1991, pp. 1–10.
[53] R.S. Smith and M. Dahleh, *The Modeling of Uncertainty in Control Systems* (Lecture Notes in Control and Information Sciences, vol. 192). London, U.K.: Springer-Verlag, 1994.
[54] A.J. Helmicki, C.A. Jacobson, and C.N. Nett, "Control oriented system identification: A worst-case/deterministic approach in H_∞," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1163–1176, Oct. 1991.
[55] H. Hjalmarsson and L. Ljung, "A unifying view of disturbances in identification," in *Proc. 10th IFAC Symp. on System Identification,* 1994, vol. 2, pp. 73–78.
[56] G.C. Calafiore, F. Dabbene, and R. Tempo, "Randomized algorithms for probabilistic robustness with real and complex structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 45, no. 12, pp. 2218–2235, Dec. 2000.
[57] M. Vidyasagar, "Randomized algorithms for robust controller synthesis using statistical learning theory: A tutorial overview," *Eur. J. Contr.*, vol. 7, pp. 287–310, 2001.
[58] G.C. Calafiore and M.C. Campi, "The scenario approach to robust control design," *IEEE Trans. Automat. Contr.*, vol. 51, no. 5, pp. 742–753, May 2006.
[59] D.S. Bayard, Y. Yam, and E. Mettler, "A criterion for joint optimization of identification and robust control," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 986–991, 1992.
[60] R.J.P. Schrama, "Accurate identification for control: The necessity of an iterative scheme," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 991–994, 1992.
[61] K. Liu and R.E. Skelton, "Closed loop identification and iterative controller design," in *Proc. 29th IEEE Conf. Decision and Control*, 1990, pp. 482–487.
[62] R.J.P. Schrama and O.H. Bosgra, "Adaptive performance enhancement by iterative identification and control design," *Int. J. Adaptive Contr. Signal Processing*, vol. 7, no. 5, pp. 475–487, 1993.
[63] W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y. Mareels, "A new approach to adaptive robust control," *Int. J. Adaptive Contr. Signal Processing*, vol. 7, pp. 183–211, 1993.
[64] K.J. Åström and J. Nilsson, "Analysis of a scheme for iterated identification and control," in Proc. *IFAC Symp. Identification,* Copenhagen, Denmark, 1994, pp. 171–176.
[65] Z. Zang, R.R. Bitmead, and M. Gevers, "Iterative weighted least-squares identification and weighted LQG control design," *Automatica*, vol. 31, no. 11, pp. 1577–1594, 1995.
[66] R.A. de Callafon and P.M.J. Van den Hof, "Suboptimal feedback control by a scheme of iterative identification and control design," *Math. Model. Syst.*, vol. 3, no. 1, pp. 77–101, 1997.
[67] B.D.O. Anderson, X. Bombois, M. Gevers, and C. Kulcsár, "Caution in iterative modeling and control design," in *IProc. IFAC Workshop Adaptive Control and Signal Processing*, Glasgow, U.K., 1998, pp. 13–19.
[68] R.L. Kosut and B.D.O. Anderson, "Least-squares parameter set estima-tion for robust control design," in *Proc. American Control Conf.*, Baltimore, MD, 1994, pp. 3002–3006.
[69] P. Mäkilä, J. Partington, and T. Gustafsson, "Worst-case control-relevant identification," *Automatica*, vol. 31, no. 12, pp. 1799–1819, 1995.
[70] D.K. de Vries and P.M.J. Van den Hof, "Quantification of uncertainty in transfer function estimation: A mixed probabilistic-worst-case approach," *Automatica*, vol. 31, pp. 543–558, 1995.
[71] X. Bombois, M. Gevers, and G. Scorletti, "A measure of robust stability for an identified set of parametrized transfer functions," *IEEE Trans. Automat. Contr.*, vol. 45, no. 11, pp. 2141–2145, 2000.
[72] M. Gevers, X. Bombois, B. Codrons, G. Scorletti, and B.D.O. Anderson, "Model validation for control and controller validation in a prediction error identification framework—Part I: Theory," *Automatica*, vol. 39, no. 3, pp. 403–415, Mar. 2003.
[73] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, pp. 393–438, 2005.
[74] M. Gevers, "Identification for control: From the early achievements to the revival of experiment design," *Eur. J. Contr.*, vol. 11, pp. 1–18, 2005.
[75] G.C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with application to model order selection," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 913–928, 1992.
[76] K. Poolla, P.P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 951–959, May 1994.
[77] R.L. Kosut, "Uncertainty model unfalsification: A sytem identification paradigm compatible with robust control design," in *Proc. 34th IEEE Conf. Decision and Control*, New Orleans, LA, 1995, pp. 3492–3497.
[78] R.G. Hakvoort and P.M.J. Van den Hof, "Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors," *IEEE Trans. Automat. Contr.*, vol. 42, no. 11, pp. 1516–1528, 1997.

[79] L. Giarré, M. Milanese, and M. Taragna, "H 1 identification and model quality evaluation," *IEEE Trans. Automat. Contr.*, vol. 42, no. 2, pp. 188–199, 1997.
[80] B. Ninness and G.C. Goodwin, "Estimation of model quality," *Automatica*, vol. 31, no. 12, pp. 32–74, 1995.
[81] X. Bombois, M. Gevers, G. Scorletti, and B.D.O. Anderson, "Robustness analysis tools for an uncertainty set obtained by prediction error identification," *Automatica*, vol. 37, no. 10, pp. 1629–1636, Oct. 2001.
[82] S.G. Douma and P.M.J. Van den Hof, "Relation between uncertainty structures in identification for robust control," *Automatica*, vol. 41, no. 3, pp. 439–457, Mar. 2005.
[83] H. Hjalmarsson, M. Gevers, and F. De Bruyne, "For model-based control design, closed-loop identification gives better performance," *Automatica*, vol. 32, pp. 1659–1673, 1996.
[84] U. Forssell and L. Ljung, "Some results on optimal experiment design," *Automatica*, vol. 36, pp. 749–756, 2000.
[85] F. Hansen, G. Franklin, and R.L. Kosut, "Closed-loop identification via the fractional representation: Experiment design," in *Proc. American Control Conf.*, pp. 1422–1427, 1989.
[86] P.M.J. Van den Hof and R.J.P. Schrama, "An indirect method for transfer function estimation from closed loop data," *Automatica*, vol. 29, pp. 1523–1527, 1993.
[87] P.M.J. Van den Hof, R.J.P. Schrama, R.A. de Callafon, and O.H. Bosgra, "Identification of normalized coprime plant factors from closed-loop experimental data," *Eur. J. Contr.*, vol. 1, pp. 62–74, 1995.
[88] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, pp. 1215–1241, 1999.
[89] R.K. Mehra, "Optimal input signals for parameter estimation in dynamic systems—Survey and new results," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 6, pp. 753–768, Dec. 1974.
[90] M. Zarrop, *Optimal Experiment Design for Dynamic System Identification* (Lecture Notes in Control and Information Sciences, vol. 21); New York: Springer-Verlag, 1979.
[91] G.C. Goodwin and R.L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*. New York: Academic, 1977.
[92] M. Gevers, X. Bombois, B. Codrons, F. De Bruyne, and G. Scorletti, "The role of experimental conditions in model validation for control," in *Robustness in Identification and Control—Proc. of Siena Workshop, July 1998*, (Lecture Notes in Control and Information Sciences, vol. 245) A. Garulli, A. Tesi, and A. Vicino, Eds. New York: Springer-Verlag, 1999, pp. 72–86.
[93] R. Hildebrand and M. Gevers, "Identification for control: Optimal input design with respect to a worst-case ν-gap cost function," *SIAM J. Contr. Optimin.* vol. 41, no. 5, pp. 1586–1608, Mar. 2003.
[94] K. Lindqvist and H. Hjalmarsson, "Identification for control: Adaptive input design using convex optimization," in *Proc. 40th IEEE Conf. Decision and Control*, Orlando, FL, pp. 4326–4331, 2001.
[95] M. Barenthin, H. Jansson, and H. Hjalmarsson, "Applications of mixed H_2 and H_∞ input design in identification," in *Proc. 16th IFAC World Congr. Automatic Control*, Prague, Czechoslovakia, July 2005, paper 03882.
[96] X. Bombois, G. Scorletti, M. Gevers, R. Hildebrand, and P. Van den Hof, "Cheapest open-loop identification for control," in *Proc. 33rd IEEE Conf. Decision and Control*, The Bahamas, Dec. 2004, pp. 382–387.
[97] L. Ljung and K. Glover, "Frequency domain versus time domain methods in system identification," *Automatica*, vol. 17, no. 1, pp. 71–86, 1981.
[98] P. Parker and R.R. Bitmead, "Adaptive frequency response identification," in *Proc. Int. Symp. Signal Processing and Its Applications*, Brisbane, Australia, 1987, pp. 232–238.
[99] J.R. Partington, "Robust identification and interpolation in H_∞," *Int. J. Contr.*, vol. 54, pp. 1281–1290, 1991.
[100] G. Gu and P.P. Khargonekar, "A class of algorithms for identification in H_∞," *Automatica*, vol. 28, pp. 299–312, 1992.
[101] D.S. Bayard, F.Y. Hadaegh, Y. Yam, R.E. Scheid, E. Mettler, and M.H. Milman, "Automated on-orbit frequency-domain identification for large space structures," *Automatica*, vol. 27, no. 6, pp. 931–946, Nov. 1991.
[102] J.S. Lew, J.N. Juang, and R.W. Longman, "Comparison of several system identification methods for flexible structures," *J. Sound Vibration*, vol. 167, no. 3, pp. 461–480, 1993.
[103] D.S. Bayard, "An algorithm for state-space frequency domain identification without windowing distortions," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1880–1885, 1994.
[104] J.-N. Juang, *Applied System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[105] J. Schoukens, R. Pintelon, and J. Renneboog, "A maximum likelihood estimator for linear and nonlinear systems—A practical application of estimation techniques in measurement problems," *IEEE Trans. Instrumen. Meas.*, vol. 37, no. 1, pp. 10–17, Mar. 1988.
[106] R. Pintelon and J. Schoukens, "Robust identification of transfer functions in the s- and z-domains," *IEEE Trans. Instrumen. Meas.*, vol. 39, no. 4, pp. 565–573, 1990.
[107] R. Pintelon and J. Schoukens, *System Identification—A Frequency Domain Approach*. New York: IEEE Press, 2001.
[108] L. Ljung, "Frequency domain versus time domain methods in system identification," in *Control of Uncertain Systems: Modelling, Approximation and Design*, B. Francis and J. Willems, Eds. (Lecture Notes in Control and Information Sciences). New York: Springer-Verlag, 2006, pp. 277– 291.
[109] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: a unified overview," *Automatica*, vol. 31, no. 12, pp. 1691–1724, Dec. 1995.
[110] A. Juditsky, H. Hjalmarsson, A. Benveniste, B. Delyon, L. Ljung, J. Sjöberg, and Q. Zhang, "Nonlinear black-box models in system identification: Mathematical foundations," *Automatica*, vol. 31, pp. 1725–1750, 1995.
[111] J. Schoukens, R. Pintelon, T. Dobrowiecki, and Y. Rolain, "Identification of linear systems with nonlinear distortions," in *Proc. 13th IFAC Symp.on System Identification*, Rotterdam, The Netherlands, 2003, pp. 1761–1772.
[112] P.S.C. Heuberger, O.H. Bosgra, and P.M.J. Van den Hof, "A generalized orthonormal basis for linear dynamical systems," *IEEE Trans. Automat. Contr.*, vol. 40, no. 3, pp. 451–465, 1995.
[113] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "Asymptotic variance expressions for output error model structures," in *Proc. 14th IFAC World Congress*, Beijing, P.R. China, 1999, pp. 367–372.
[114] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "On the fundamental role of orthonormal bases in system identification," *IEEE Trans. Automat. Contr.*, vol. 44, no. 7, pp. 1384–1407, July 1999.
[115] P.S.C. Heuberger, P.M.J. Van den Hof, and B. Wahlberg, Eds., *Modelling and Identification with Rational Orthogonal Basis Functions*. London: Springer, 2005.
[116] B.D.O. Anderson, "Identification of scalar errors-in-variables models with dynamics," *Automatica*, vol. 21, no. 6, pp. 709–716, Nov. 1985.
[117] T. Söderström, U. Soverini, and K. Mahata, "Perspectives on errors-in-variables estimation for dynamic systems," *Signal Processing*, vol. 82, no. 8, pp. 1139–1154, Aug. 2002.
[118] I. Markovsky, J.C. Willems, S.V. Huel, B. De Moor, and R. Pintelon, "Application of structured total least squares for system identification and model reduction," *IEEE Trans. Automat. Contr.*, vol. 50, no. 10, pp. 1490–1500, Oct. 2005.

Force was the obvious means of persuasion and this was better used against women and children than against the tapper, who might then be unable to work efficiently. A procedure was soon established and documented in the official manual given to all agents. The soldiers would arrive at a settlement, loot it of animals and any other items of value, destroy the building, capture the women and children, and imprison them in stockades built close to each trading post for just this purpose. They would then be ransomed against an arbitrarily decided weight of rubber. On returning with the rubber, the tappers often found that their women had been raped by the "sentries" and/or had died from starvation or some disease.

—*John Loadman*, Tears of the Tree, The Story of Rubber—A Modern Marvel. *Oxford University Press, 2005, pp. 127–128.*