PERGAMON

Annual Reviews in Control 22 (1998) 173–186

Annual Reviews in Control

CLOSED-LOOP ISSUES IN SYSTEM IDENTIFICATION

Paul Van den Hof

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

Abstract. The identification of dynamical systems on the basis of data, measured under closed-loop experimental conditions, is a problem which is highly relevant in many (industrial) applications. Initiated by an emerging interest in the area called 'identification for control', classical prediction error identification methods have been extended to also handle the problem of identifying approximate models from closed-loop observations. In this paper the several procedures that have resulted from this research are reviewed and their characteristic properties are compared. Additionally it is discussed which role closed-loop identification can play in the identification of (optimal) models for (robust) control design.

Keywords. Closed-loop identification; system identification; linear systems; control systems; asymptotic analysis.

1. INTRODUCTION

Many industrial processes operate under feedback control. Due to unstable behaviour of the plant, required safety and/or efficiency of operation, experimental data can only be obtained under so-called closed-loop conditions. This applies not only to many industrial production processes like e.g. paper production, glass production, and chemical separation processes like crystallization, but also to mechanical servo systems as robotic manipulators and high precision motion control systems in e.g. audio and CD-ROM disc drives. Besides, many processes in non-technical areas, as e.g. biological and economic systems, involve inherent feedback loops that can not be manipulated and/or removed.

Identification methods for dealing with closed-loop experimental data have been developed in the seventies and eighties, see e.g. Gustavsson et al. (1977) and Söderström and Stoica (1989) for an overview. These "classical" methods are typically directed towards solving the consistency problem, considering the situation that plant and disturbance model can be modeled exactly (system is in the model set).

Initiated by an emerging interest in the identification of models that are particularly suitable for model-based (robust) control design, renewed attention has been given lately to the problem of closed-loop identification. There is a number of arguments to prefer closed-loop experiments over open-loop ones, in case one is interested in model-based control design. These arguments comprise aspects of bias and variance, control of (input and/or output) signal power during experiments, input shaping, and the fact that a controller can linearize the (possibly nonlinear) plant behaviour in a relevant working point, thus enabling accurate linear modelling.

Unlike the classical situation, attention is now also given to properties of identified approximate models, handling the -more realistic - situation that plant and noise dynamics are not exactly present in the model set considered. Accounts of this area are given in the survey papers Gevers (1993) and Van den Hof and Schrama (1995).

In this paper recent results on closed-loop identification

PII: S1367-5788(98)00016-9

methods are discussed, leading to a "consumer's guide" that shows which method to prefer in which situation. Aspects of bias and variance are considered, as well as more algorithmic type aspects like: can attention be restricted to fixed order model classes, can unstable (but stabilized) plants be handled, and what kind of knowledge is presumed to be available a priori? It will be indicated in which way the "modern" methods are generalizations of "classical" ones.

Additionally the relevance of a closed-loop experimental setup for identifying control-relevant models will be discussed, by considering aspects of asymptotic bias and variance of nominal models, as well as aspects related to the construction of uncertainty models for (robust) control design.

Attention will be restricted to identification criteria originating from the main-stream prediction error framework (Ljung, 1987). However, many of the structural and parametrizational issues involved are directly applicable also in other identification frameworks, as e.g. set membership identification (Milanese and Vicino, 1991), and worst-case identification (Mäkilä, *et al.*, 1995).

2. PRELIMINARIES

The experimental setup to be considered in this paper is depicted in Figure 1.

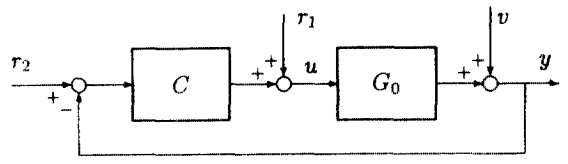

Fig. 1. Closed-loop configuration.

In this configuration the external signals r_1 and r_2 can be either a tracking signal, a setpoint or a noise disturbance on the regulator/process output. v is a (stochastic) noise disturbance that is modeled as a filtered white noise: $v(t) = H_0(q)e(t)$ with H_0 a monic stable and stably invertible filter, and e a sequence of independent identically distributed random variables (white noise) with bounded moments and variance λ_0. q is the forward shift operator: $qu(t) = u(t+1)$. The external signals r_1, r_2 are assumed to be uncorrelated with the noise disturbance v.

The two signals r_1, r_2 will often be collected into one signal:

$$r(t) := r_1(t) + C(q)r_2(t). \tag{1}$$

Then the feedback law can be written as:

$$u(t) = r(t) - C(q)y(t), \tag{2}$$

and the closed-loop system equations are: [1]

$$y(t) = G_0S_0r(t) + S_0H_0e(t) \tag{3}$$

$$u(t) = S_0r(t) - CS_0H_0e(t) \tag{4}$$

with the sensitivity function $S_0 = (1 + CG_0)^{-1}$.

It will be assumed that the closed-loop system is internally stable, meaning that the four transfer functions in (3)-(4) are stable, i.e. analytic in $|z| \geq 1$. In order to avoid technicalities, it will also be assumed that the product G_0C is strictly proper, which means that either C or G_0 contains a time delay.

The information that is typically available in a closed-loop identification problem can be characterized in several levels:

- Measurements of u and y;
- Knowledge about excitation properties of r_1, r_2;
- Measurements of r_1, r_2;
- Knowledge of C.

The identification objective will generally be the construction of a model of the transfer function $G_0(z)$ and possibly $H_0(z)$. Sometimes one may also wish to determine the controller $C(z)$ in the feedback path.

3. CLASSICAL SOLUTIONS

The typical problem in closed-loop identification is the fact that the plant input signal u is correlated with the output noise disturbance v. This is also the reason why a nonparametric (spectral) estimate of G_0, obtained from direct operation on u ad y, will deliver a plant estimate that is a weighted average between G_0 and $-1/C$ (Söderström and Stoica, 1989). Therefore, for nonparametric estimates of G_0, an external excitation signal r (either through or r_1 or through r_2) is required to provide an unbiased estimate of G_0 through

$$\hat{G}(e^{i\omega}) := \frac{\hat{\Phi}_{yr}}{\hat{\Phi}_{ur}}$$

with $\hat{\Phi}_{yr}$ and $\hat{\Phi}_{ur}$ spectral estimates of the corresponding cross-spectra.

In parametric identification three approaches have been followed (Söderström and Stoica, 1989):

- Direct identification
- Indirect identification
- Joint input/output identification

of which the direct method is most popular. Here one simply applies the standard (prediction error) identifica-

[1] When there is no risk of confusion, the arguments q will be discarded.

tion procedure without taking account of the presence of a feedback controller. A parameter estimate is obtained by

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

with $\varepsilon(t,\theta) = H(q,\theta)^{-1}[y(t) - G(q,\theta)u(t)]$.

This direct identification approach has a number of interesting properties. If the system $\mathcal{S} := (G_0, H_0)$ is present in the model set $\mathcal{M} := \{(G(q,\theta), H(q,\theta)), \theta \in \Theta\}$, then a consistent estimate is obtained in each of the following situations

- a sufficiently exciting signal r_1 or r_2 is present;
- C is a controller of sufficiently high order;
- C is a controller that switches between several settings during the experiment.

In this situation it is allowed that C is a nonlinear and/or time-varying controller.

However in the situation that $\mathcal{S} \notin \mathcal{M}$ the properties of the direct method collapse, and even in the situation that the plant G_0 can be modeled exactly within $\mathcal{M}$, consistency of $G(q,\hat{\theta}_N)$ is lost if the noise model is misspecified.

This can be visualized by expressing the asymptotic identification criterion in the frequency domain. Following the standard prediction error framework (Ljung, 1987), it reads that $\hat{\theta}_N \to \theta^*$ w.p. 1 for $N \to \infty$, where $\theta^* = \arg\min_\theta \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_\varepsilon(\omega,\theta)d\omega$. For the direct identification method it can be verified that

$$\Phi_\varepsilon = \frac{|S_0|^2|G_0 - G(\theta)|^2}{|H(\theta)|^2}\Phi_r + \frac{|H_0|^2|S_0|^2}{|H(\theta)|^2|S(\theta)|^2}\lambda_0$$

where the arguments $e^{i\omega}$ are suppressed for brevity, and $S(q,\theta) = (1 + CG(q,\theta))^{-1}$ is the sensitivity function of the parametrized model. If G_0 can be modeled exactly, i.e. $G_0 \in \mathcal{G}$, it can not be concluded that consistency of G will result. This is due to the fact that $G(q,\theta)$ appears in both terms of the integrand (in the second term through $S(q,\theta)$) and so any misfit in $H(q,\theta)$ will be compensated for by $G(q,\theta)$.

The above characterization of the asymptotic bias is rather implicit. A more explicit expression is presented by Ljung (1993) for the situation of a fixed noise model $H(q,\theta) = H_*(q)$:

$$\theta^* = \arg\min_{\theta} \frac{1}{2\pi}\int_{-\pi}^{\pi} |G_0 - B - G(\theta)|^2 \frac{\Phi_u}{|H_*|^2}d\omega \quad (5)$$

with $B = M_+(H_0 - H_*)\lambda_0$

and M_+ the causal stable part of CS_0H_0/Φ_u. This expression shows that the plant model will converge to a biased estimate $G_0 - B$, where the bias term is determined by the noise level in the loop, the input signal power, and the accuracy of the noise model. The bias will thus be small when

- The noise model is an accurate description of H_0, and/or
- The signal to noise ratio at the input u is large;

Summarizing, the direct method can provide good estimates when one is willing to identify full order models for both plant and noise dynamics. In case one is aiming at approximate models or when one refrains from modelling the full noise dynamics, the consequences are that

- G_0 is not identified consistently, and
- The bias expression that governs the identification of G_0 is not explicitly tunable by the user.

The latter aspect is particularly important when one is interested in identifying reduced-order models that approximate the original system in a predefined way, as e.g. present in an open-loop identification using an output error model structure, determined by

$$\varepsilon(t,\theta) = L(q)[y(t) - G(q,\theta)u(t)]$$

and L a stable prefilter that is applied to the input/output data. For this situation it can be shown that

$$\theta^* = \arg\min_{\theta} \frac{1}{2\pi}\int_{-\pi}^{\pi} |G_0 - G(\theta)|^2\Phi_u|L|^2d\omega \quad (6)$$
$$= \arg\min_{\theta} \|[G_0 - G(\theta)]H_uL\|_2$$

with H_u a stable spectral factor of Φ_u; reduced-order identification now involves an intrinsic model reduction step with a (frequency weighted) norm that can be directly tuned by the user through designing Φ_u and/or L.

The direct method of closed-loop identification is -in some cases- also able to identify unstable plants. The basic restriction here is that the predictor filters $1 - H^{-1}(q,\theta)$ and $H^{-1}(q,\theta)G(q,\theta)$ are uniformly stable. This situation is satisfied when the system (G_0, H_0) and the model structure are of type ARX or ARMAX; i.e. those situations where G and H have a common denominator. The lack of a consistency property in the situation $G_0 \in \mathcal{G}$ and the desire to achieving more explicit bias expressions has been the motivation to develop a number of alternative identification methods, that relate to and actually generalize the classical methods of indirect and joint i/o identification.

4. ASSESSMENT CRITERIA FOR CLOSED-LOOP IDENTIFICATION METHODS

The several closed-loop identification methods that will be discussed in this paper can be evaluated with respect to a number of different criteria. We will first present these assessment criteria.

- **Consistency of** $(\hat{G}, \hat{H})$. This is a basic requirement. Whenever our model set is rich enough to contain the data generating system ($\mathcal{S} \in \mathcal{M}$), the identification method should be able to consistently identify the plant, under additional conditions on excitation properties of the plant signals.
- **Consistency of** $\hat{G}$. The ability to identify G_0 consistently in the situation $G_0 \in \mathcal{G}$. This implies that the consistent modelling of G_0 is not dependent on possible undermodelling or misspecification of H_0. Particularly in situations where the disturbance process v contains complex dynamics, this property is favorable.
- **Tunable bias expression.** An explicit approximation criterion can be formulated that governs the asymptotic i/o model $G(q, \theta^*)$ in a way that is not dependent on Φ_v. This refers to an expression for the asymptotic bias distribution, as also formulated for open-loop experimental conditions in (6).
- **Fixed model order**. The ability of identification methods to consider model sets $\mathcal{G}$ of models with a fixed and prespecified model order. This property is important when the application of the identified model, e.g. in model-based control design, puts limitations on the acceptable complexity of the model.
- **Unstable plants**. The ability to (consistently) identify unstable plants.
- **Stabilized model** $(G(q, \theta^*), C)$. This refers to the situation that there is an a priori guarantee that the (asymptotically) identified model $G(q, \theta^*)$ is guaranteed to be stabilized by the present controller C.
- **Knowledge of controller** C. This concerns the question whether exact knowledge of the controller is required by the considered identification method.
- **Accuracy**. The (asymptotic) variance of the model estimates.

For the direct identification method, the corresponding properties are listed in the first column of Table 1. Accuracy properties of the several methods will be discussed separately in Section 6.

5. INDIRECT SOLUTIONS

5.1 *Introduction*

In this section a number of alternative methods will be discussed, that are introduced in the literature over the last couple of years. All methods are still considered within the standard prediction error identification framework (Ljung, 1987). They differ in the way that the plant model is parametrized and in the way that they deal with (removing) the noise contribution on the input of the plant. The main difference with the direct method, is that for indirect methods use is made of a measurable external excitation signal. It will generally be assumed that there is an external signal r_1 present in the experimental configuration, that is sufficiently exciting for identifying the plant dynamics. Alternative experimental conditions will be discussed in section 7.

5.2 *Two-stage method*

The two-stage method introduced in Van den Hof and Schrama (1993) is based on a rewriting of the system's equations in the following way:

$$\begin{aligned} y(t) &= G_0(q)u^r(t) + S_0(q)v(t) \\ u(t) &= u^r(t) - C(q)S_0(q)v(t) \\ u^r(t) &:= S_0(q)r_1(t) \end{aligned}$$

It is composed of the following two steps:

- Identify the transfer function between r_1 and u, using a model structure

$$u(t) = S(q, \beta)r_1(t) + W(q, \beta)\varepsilon_u(t)$$

  Then the LS-estimate $S(q, \hat{\beta}_N)$ is used to simulate a "noise-free" input signal, according to

$$\hat{u}^r(t) := S(q, \hat{\beta}_N)r_1(t).$$

- In the second step the plant model is identified by applying a LS criterion to the prediction error:

$$\varepsilon_y(t) = K(q, \theta)^{-1}[y(t) - G(q, \theta)\hat{u}^r(t)].$$

The two identification steps in this procedure are "standard" open-loop identifications for which no special tools are required; in both steps the noise contribution on the "output" signals (u, resp. y) is uncorrelated to the input signals (r_1, resp. $\hat{u}^r$).

As the estimate $S(q, \hat{\beta}_N)$ from the first step is only used for simulation purposes of $\hat{u}^r$, one can adhere to a high-order accurate modelling; the order of the plant model will only be determined in the second step. The resulting model is obtained by

$$\begin{aligned} \hat{G}(q) &= G(q, \hat{\theta}_N) \quad \text{and} \\ \hat{H}(q) &= K(q, \hat{\theta}_N)S(q, \hat{\beta}_N)^{-1} \end{aligned}$$

where the latter equation stems from the fact that the noise model $K(q, \hat{\theta}_N)$ is a model for the disturbance filter

S_0H_0 rather than for H_0 only. If in the second step a fixed (non-parametrized) noise model K_* is chosen, then the asymptotic parameter estimate θ^* satisfies

$$\theta^* = \arg\min_{\theta} \int_{-\pi}^{\pi} |G_0S_0 - G(\theta)S(\beta^*)|^2 \frac{\Phi_{r_1}}{|K_*|^2}\, d\omega$$

where $S(\beta^*)$ the asymptotically identified model in the first step.
Both G_0 (situation $G_0 \in \mathcal{G}$) and (G_0, H_0) (situation $S \in \mathcal{M}$) can be identified consistently provided that in the first step of the procedure a sufficiently rich model set is chosen for modelling S_0. Under this condition, the approximation criterion reduces to

$$\theta^* = \arg\min_{\theta} \int_{-\pi}^{\pi} |G_0 - G(\theta)|^2 \frac{|S_0|^2\Phi_{r_1}}{|K_*|^2}\, d\omega. \quad (7)$$

This two-stage approach does not require knowledge of the controller, while the model order of the resulting model in the second step of the procedure is fully under control. The method has been successfully applied to several industrial processes as a compact disc servo mechanism (De Callafon *et al.*, 1993), a sugar cane crushing plant (Partanen and Bitmead, 1995), and a crystallization plant (Eek *et al.*, 1996).

Whenever r_2 is available as an external signal instead of r_1, the results remain similar. In that case the role of S_0 will be taken over by the product CS_0. The reconstruction of the noise model is then obtained by $H(q, \hat{\theta}_N) = K(q, \hat{\theta}_N)[1 - G(q, \hat{\theta}_N)S(q, \hat{\beta}_N)]^{-1}$ with $S(q, \hat{\beta}_N)$ the estimate of CS_0 obtained in the first step.

5.3 *Coprime factor identification*

In the coprime factor approach the two steps of the procedure discussed above, are actually performed simultaneously. Starting from the system's equations

$$y(t) = G_0S_0r_1(t) + S_0v(t) \quad (8)$$
$$u(t) = S_0r_1(t) - CS_0v(t) \quad (9)$$

one can consider the one-input two-output identification problem with r_1 as input signal and $[y\ u]^T$ as output. This is an open-loop type of identification problem as r_1 and v are uncorrelated. By denoting $N_0 = G_0S_0$ and $D_0 = S_0$, it appears that (N_0, D_0) is a (rational) factorization of G_0 since $G_0 = N_0/D_0$. The factorization is called coprime (over $\mathbb{R}\mathcal{H}_\infty$) if the two factors are stable and there are no canceling unstable zeros, i.e. zeros in $|z| \geq 1$ (Vidyasagar, 1985).

All open-loop identification results apply to this identification problem. This means that consistency for G_0 will be guaranteed -irrespective of the noise modelling- provided that sufficiently large model sets are chosen for the identification of G_0S_0 and S_0 and provided that independently parametrized noise models are used.

When N_0 and D_0 are parametrized independently, the estimated transfers $\hat{N}$ and $\hat{D}$ are not likely to contain the same redundant dynamics that are present by construction in the two factors N_0 and D_0. Consequently, when constructing the final plant model through $\hat{G} = \hat{N}/\hat{D}$ the order of the model will increase to a level that generally will be equal to twice the orders of $\hat{N}$ and $\hat{D}$.

This problem can be overcome by considering an approach presented in Van den Hof *et al.*(1995). By constructing an auxiliary signal

$$x(t) := F(q)r_1(t)$$

with F a stable filter to be specified later, the system's equations can be rewritten into the form

$$y(t) = G_0S_0F^{-1}x(t) + S_0v(t) \quad (10)$$
$$u(t) = S_0F^{-1}x(t) - CS_0v(t). \quad (11)$$

while F can be chosen such that the two factors $N_{0,F} := G_0S_0F^{-1}$ and $D_{0,F} := S_0F^{-1}$ no longer exhibit redundant dynamics. There are several options for doing this, as e.g. choosing F such that $N_{0,F}$ and $D_{0,F}$ become polynomial in q^{-1}. An alternative that has better robustness properties is by considering *normalized factorizations.*

Definition 1. A factorization (N, D) is called a normalized coprime factorization (ncf) it if satisfies $|N(e^{i\omega})|^2 + |D(e^{i\omega})|^2 = 1$ for all ω.

One of the properties of ncf's is that they form a decomposition of the system G_0 in minimal order (stable) factors. In other words, if G_0 has McMillan degree n, then the coprime factors of a ncf of G_0 will also have McMillan degree n.[2] The following result is taken from Van den Hof *et al.*(1995).

Proposition 1. Let

$$F(q) := D_x(q) + C(q)N_x(q) \quad (12)$$

with (N_x, D_x) a coprime factorization of a model $G_x = N_x/D_x$ that is stabilized by C. Then

(a) $(N_{0,F}, D_{0,F})$ is a coprime factorization of G_0;

[2] In the exceptional case that G_0 contains all-pass factors, (one of) the ncf's will have McMillan degree $< n$, see Tsai *et al.* (1992).

(b) If (N_x, D_x) is a ncf of G_0 then $(N_{0,F}, D_{0,F})$ is normalized.

Apparently, knowledge of G_0 is required to construct ncf's $(N_{0,F}, D_{0,F})$. If these ncf's are accessible, then a model parametrization of $N(q,\theta)$ and $D(q,\theta)$ with a common denominator is justified. This leads to the following algorithm:

(1) Construct a high order estimate of G_x of G_0 that is stabilized by C, and construct a data filter F according to (12) on the basis of a ncf (N_x, D_x) of G_x.
(2) Construct the auxiliary signal $x = F(q)r_1(t)$;
(3) Identify a one-input two-output model in the model structure determined by

$$\varepsilon(t,\theta) = \begin{bmatrix} K_y(q,\theta)^{-1}[y(t) - N(q,\theta)x(t)] \\ K_u(q,\theta)^{-1}[u(t) - D(q,\theta)x(t)] \end{bmatrix}$$

where $N(q,\theta)$ and $D(q,\theta)$ are parametrized in terms of polynomial fractions and with a common denominator:

$$N(q,\theta) = \frac{B(q^{-1},\theta)}{D(q^{-1},\theta)}; \quad D(q,\theta) = \frac{A(q^{-1},\theta)}{D(q^{-1},\theta)} \tag{13}$$

with A, B and D polynomials in q^{-1} with prespecified model orders n, n and m. The finally estimated plant model of order n is then given by

$$G(q,\hat{\theta}_N) = \frac{B(q^{-1},\hat{\theta}_N)}{A(q^{-1},\hat{\theta}_N)}. \tag{14}$$

$K_y(q,\theta)$ and $K_u(q,\theta)$ are the noise models in the two transfer functions.

The common denominator parametrization (13) can also be realized in a one-input, two-output state space model (A,B,C,D), with

$$A = \begin{bmatrix} -d_1 & -d_2 & \cdots & -d_m \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C = \begin{bmatrix} b_1 & \cdots & b_n & 0 \\ a_1 & \cdots & a_n & 0 \end{bmatrix} \quad D = \begin{bmatrix} b_0 \\ a_0 \end{bmatrix}.$$

In step (1) of the algorithm, exact knowledge of G_0 is actually replaced by a (high order) estimate of G_0. A possible inaccuracy in this step will have an effect on the final estimate $G(q,\hat{\theta}_N)$ that can be bounded in terms of the gap-metric. For more details one is referred to Van den Hof *et al.*(1995).

When using fixed noise models, i.e. $K_y(q,\theta) = K_{y*}(q)$ and $K_u(q,\theta) = K_{u*}(q)$, it can be verified that a least squares identification criterion will yield the asymptotic parameter estimate $\theta^* = \arg\min_\theta \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_\varepsilon(\omega)d\omega$ with

$$\Phi_\varepsilon = \left\{ \frac{|G_0S_0 - N(\theta)F|^2}{|K_{y*}|^2} + \frac{|S_0 - D(\theta)F|^2}{|K_{u*}|^2} \right\} \Phi_{r_1}.$$

When also noise models are estimated, a model of H_0 can be constructed by:

$$H(q,\hat{\theta}_N) = (1 + C\hat{G})\ \hat{K}_y(q),$$

while additionally an estimate of the controller can be obtained through: $\hat{C}(q) = [F^{-1} - \hat{D}]/\hat{N}$.
The coprime identification method has some relation with the "classical" joint i/o method; the plant signals u and y are modeled with external signals r_1, e as inputs, whereas in the joint i/o method the plant signals are modeled as functions of two unmeasurable (noise) signals. Due to the fact that plant and controller are represented in terms of stable rational factors, there are no additional problems in handling unstable plants and/or unstable controllers.

5.4 *Identification in dual Youla/Kucera parametrization*

There is yet another parametrization of the data generating system, that is closely related to the coprime factor framework described above. When the controller C is known, then the set of all linear plant models that are stabilized by C can be parametrized by

$$G = \frac{N_x + D_cR}{D_x - N_cR} \tag{15}$$

where (N_x, D_x) is a coprime factorization of just any model stabilized by $C = N_c/D_c$, and R varies over the class of stable transfer functions. For the parametrization of all stabilizing controllers, this parametrization is known as the parametrization of Youla or Kucera (Desoer et al., 1980). Its use in identification has been introduced by Hansen and Franklin (1988) and further exploited in Schrama (1992) and Lee et al. (1993).

By solving R from (15) for a given system G_0, it follows that the corresponding (unique) R satisfies:

$$R_0 = \frac{(G_0 - G_x)D_x}{(1 + CG_0)D_c}, \tag{16}$$

and additionally, that the two rational factors in (15) satisfy

$$N_x + D_cR_0 = G_0S_0(D_x + CN_x) \tag{17}$$

$$D_x - N_cR_0 = S_0(D_x + CN_x). \tag{18}$$

Note that these are exactly the same expressions as the two coprime factors that represent G_0 in the coprime

factor approach. The difference is now, that the redundancy in the two separate factors is removed and a single stable transfer function R_0 remains. Substituting (17) and (18) in the system's equations (10), (11) and denoting

$$z(t) := (D_c + G_x N_c)^{-1}(y(t) - G_x u(t))$$

it follows after some manipulation that

$$z(t) = R_0(q)x(t) + K_0(q)e(t) \tag{19}$$

with $K_0 = D_c^{-1} S_0 H_0$.
It is important to note that $z(t)$ and $x(t)$ can simply be reconstructed from measured data using knowledge of the controller, and that $x(t)$ and $e(t)$ are uncorrelated. As a result, R_0 (and possibly K_0) can be identified by standard open-loop techniques, on the basis of available reconstructed signals x and z.
The transfer function R_0 is a particular closed-loop transfer function. For a specific choice of G_x the method simplifies to a classical indirect method as shown in the following example.

Example 1. If C is stable, then $G_x = 0$ is an (auxiliary) model that is stabilized by C. By choosing $N_x = 0$, $D_x = D_c = 1$ and $N_c = C$ the transfer function R_0 becomes $R_0 = G_0/(1 + CG_0)$, being the (closed-loop) transfer function from r_1 to y. In this case $z(t) = y(t)$ and $x(t) = r_1(t)$.

In Van den Hof and de Callafon (1996) it is shown how particular choices of factorizations of G_x and C lead to different closed-loop transfer functions to be identified, also applicable to the multivariable situation.

In the identification of R_0 and K_0, the related prediction error will be

$$\varepsilon(t,\theta) = K(q,\theta)^{-1}[z(t) - R(q,\theta)x(t)], \tag{20}$$

and identified models $\hat{R}$, $\hat{K}$ will be used to construct the identified plant model according to

$$\hat{G} = \frac{N_x + D_c\hat{R}}{D_x - N_c\hat{R}} \text{ and} \qquad \hat{H} = [I + \hat{G}C]D_c\hat{K}. \tag{21}$$

In the situation of a fixed noise model, $K(q,\theta) = K_*$, the asymptotic parameter estimate θ^* satisfies: $\theta^* = \arg\min_\theta \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_\varepsilon(\omega)d\omega$ with

$$\Phi_\varepsilon(\omega) = \frac{|R_0 - R(\theta)|^2}{|K_*|^2}\Phi_x,$$

which after some manipulations using the expressions for R_0, $R(\theta)$ and x can be shown to reduce to

$$\Phi_\varepsilon(\omega) = \left|\left[\frac{G_0}{1+CG_0} - \frac{G(\theta)}{1+CG(\theta)}\right]\frac{1}{D_cK_*}\right|^2 \Phi_r.$$

Note that this expression is independent of the chosen auxiliary model G_x. One of the particular advantages of this approach, is that every estimate $\hat{R}$ that is stable, will provide a plant model $\hat{G}$ that - by construction- is stabilized by the controller C. This is due to the particular Youla/Kucera parametrization.

The main problem of this approach is that the order of the identified model can become unnecessarily large. The step (21) from an estimated $\hat{R}$ (of specified order) to a plant model $\hat{G}$, will cause the model order to increase severely. This also holds for the classical indirect method.

5.5 *Tailor-made parametrization*

Indirect identification requires two separate steps: (1) identification of a "closed-loop" transfer function, and (2) recalculation of the open-loop plant model. The two steps can be combined into one, by using a tailor-made parametrization for the closed-loop system, using knowledge of the controller to parametrize the closed-loop system in terms of open-loop plant parameters.
This leads to the prediction error:

$$\varepsilon(t,\theta) = K(q,\theta)^{-1}[y(t) - \frac{G(q,\theta)}{1 + G(q,\theta)C(q)}r_1(t)]. \tag{22}$$

With this tailor-made parametrization, least squares estimation will also require a tailor-made optimization algorithm, as the model set is parametrized in a structure that is different from the standard (open-loop) model sets.

For a fixed noise model $K_*(q)$ it follows directly that the asymptotic bias distribution is governed by:

$$\theta^* = \arg\min_{\theta\in\Theta} \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_0G_0 - S(\theta)G(\theta)|^2 \frac{\Phi_{r_1}}{|K_*|^2}d\omega.$$

Consistency properties of this identification method can be shown by applying the standard (open-loop) prediction error framework. The price that has to be paid is that one has to deal with a complicated parametrized model structure (22), with its resulting computational burden. Additionally, a problem of lack of connectedness of the parameter space can occur. When $G(q,\theta)$ is parametrized as a quotient of two polynomials with free coefficients, the region of θ for which the transfer function $\frac{G(q,\theta)}{1+C(q)G(q,\theta)}$ is stable can be composed of two (or more) disconnected regions. This can complicate the

identification of an accurate model severely. The problem can be avoided by using a model order for $G(q,\theta)$ that is not smaller than the order of the controller (Van Donkelaar and Van den Hof, 1997).
This method has also been analyzed in a recursive implementation in Landau and Boumaïza (1996).

One can make further use of knowledge of the closed-loop structure by also parametrizing the noise model $K(q,\theta)$ in terms of its open-loop parameters:

$$K(q,\theta) = \frac{H(q,\theta)}{1 + C(q)G(q,\theta)}.$$

As indicated in van Donkelaar and Van den Hof (1997), substituting this expression into (22), leads to

$$\begin{aligned}\varepsilon(t,\theta) &= H^{-1}(q,\theta)[y - G(q,\theta)(r - C(q)y)] \\ &= H^{-1}(q,\theta)[y - G(q,\theta)u] \end{aligned} \quad (23)$$

and this expression is exactly the same as the related one for the direct identification method. As a result, the two identification methods become equivalent when using this specifically parametrized noise model.

6. BIAS AND VARIANCE ASPECTS

6.1 *Bias*

The indirect methods presented in section 5 show expressions for the asymptotic bias distribution that are very much alike. When using fixed (non-parametrized) noise models during identification, consistent plant models of G_0 can be obtained, and the bias distribution has the form[3]

$$\theta^* = \arg\min_{\theta\in\Theta} \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_0G_0 - S(\theta)G(\theta)|^2 \frac{\Phi_{r_1}}{|K_*|^2} d\omega. \quad (24)$$

By designing the (fixed) noise model K_* (or the signal spectrum Φ_{r_1}), this bias expression can explicitly be tuned to the designer's needs. However the expression is different from the related open-loop expression (6). Instead of a weighted additive error on G_0, the integrand contains an additive error on G_0S_0. Straightforward calculations show that

$$G_0S_0 - G(\theta)S(\theta) = S_0[G_0 - G(\theta)]S(\theta),$$

so that the asymptotic bias distribution can be characterized by

$$\theta^* = \arg\min \| \frac{[G_0 - G(\theta)]\Phi_{r_1}^{\frac{1}{2}}}{(1 + CG_0)(1 + CG(\theta))K_*} \|_2. \quad (25)$$

[3] Details vary slightly over the several identification methods.

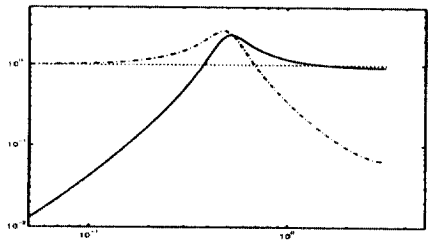

Fig. 2. Typical curve for Bode magnitude plot of sensitivity function S_0 (solid) and related complementary sensitivity $G_0C/(1 + CG_0)$ (dashed).

This implies that in the (indirect) closed-loop situation, the additive error on G_0 is always weighted with S_0. Thus emphasis will be given to an accurate model fit in the frequency region where S_0 is large and the identified model will be less accurate in the frequency region where S_0 is small. In Figure 2 a typical characteristic of S_0 and closed-loop transfer $G_0C/(1 + CG_0)$ is sketched. This illustrates that emphasis will be given to an accurate model fit in the frequency region that particularly determines the bandwidth of the control system. In this area (where $|S_0| \geq 1$), the noise contribution of v in the output signal y is amplified by the controller.
According to Bode's sensitivity integral (Sung and Hara, 1988) for a stable controller:

$$\int_0^{\pi} log\ |S(e^{i\omega})|\ d\omega = c \ \text{(constant)}$$

with c determined by the unstable poles of the plant, and $c = 0$ for G_0 stable. This implies that the attenuation of signal power in the low frequency range, will always be "compensated" for by an amplification of signal power in the higher frequency range.
The aspect that closed-loop identification stresses the closed-loop relevance of identified (approximate) models, has been given strong attention in the research on "identification for control".

Whereas direct identification needs consistent estimation of noise models in order to consistently identify G_0, indirect methods can do without noise models. Incorporation of noise models in indirect methods is very well possible, but this will result in bias distributions that become dependent on the identified noise models as well as on (the unknown) Φ_v.

6.2 *Variance*

For analyzing the asymptotic variance of the transfer function estimates we consider again the prediction error framework (Ljung, 1987) that provides variance ex-

pressions that are asymptotic in both n (model order) and N (number of data). For the direct identification approach, and in the situation that $\mathcal{S} \in \mathcal{M}$ this delivers:

$$cov \begin{pmatrix} \hat{G}(e^{i\omega}) \\ \hat{H}(e^{i\omega}) \end{pmatrix} \sim \frac{n}{N} \Phi_v(\omega) \cdot \begin{bmatrix} \Phi_u(\omega) & \Phi_{eu}(\omega) \\ \Phi_{ue}(\omega) & \lambda_0 \end{bmatrix}^{-1} \quad (26)$$

The following notation will be introduced:

$$u(t) = u^r(t) + u^e(t)$$

with $u^r := S_0(q)r_1$ and $u^e := -CS_0(q)v$ and the related spectra $\Phi_u^r = |S_0|^2 \Phi_{r_1}$ and $\Phi_u^e = |CS_0|^2 \Phi_v$. Using the expression $\Phi_{ue} = -CS_0H_0\lambda_0$, (26) leads to (Ljung, 1993; Gevers et al., 1997):

$$cov \begin{pmatrix} \hat{G} \\ \hat{H} \end{pmatrix} \sim \frac{n}{N} \frac{\Phi_v}{\Phi_u^r} \cdot \begin{bmatrix} 1 & (CS_0H_0)^* \\ CS_0H_0 & \frac{\Phi_u}{\lambda_0} \end{bmatrix},$$

and consequently

$$cov(\hat{G}) \sim \frac{n}{N} \frac{\Phi_v}{\Phi_u^r} \qquad cov(\hat{H}) \sim \frac{n}{N} \frac{\Phi_v}{\lambda_0} \frac{\Phi_u}{\Phi_u^r}. \quad (27)$$

This shows that only the noise-free part u^r of the input signal u contributes to variance reduction of the transfer functions. Note that for $u^r = u$ the corresponding open-loop results appear.

In Gevers et al. (1997) it is shown that for *all* indirect methods presented in section 5, these expressions remain the same. However, again there is one point of difference between the direct and indirect approach. The indirect methods arrive at the expression for $cov(\hat{G})$ also without estimating a noise model (situation $G_0 \in \mathcal{G}$), whereas the direct method requires a consistent estimation of H_0 for the validity of (27).

The asymptotic variance analysis tool gives an appealing indication of the mechanisms that contribute to variance reduction. It also illustrates one of the basic mechanisms in closed-loop identification, i.e. that noise in the feedback loop does not contribute to variance reduction. Particularly in the situation that the input power of the process is limited, it is relevant to note that only part of this input power can be used for variance reduction. This has led to the following results (Gevers and Ljung, 1986):

- If the input power is constrained then minimum variance of the transfer function estimates is achieved by an open-loop experiment;
- If the output power is constrained then the optimal experiment is a closed-loop experiment.

Because of the "doubly asymptotic" nature of the results ($N, n \to \infty$), this asymptotic variance analysis tool is also quite crude.

For finite model orders, the variance results will likely become different over the several methods. The direct method will reach the Cramer-Rao lower bound for the variance in the situation $\mathcal{S} \in \mathcal{M}$. Similar to the open-loop situation, the variance will typically increase when no noise models are estimated in the indirect methods. This will also be true for the situation that two - independent - identification steps are performed on one and the same data set, without taking account of the relation between the disturbance terms in the two steps. Without adjustment of the identification criteria, the two stage method and the coprime factor method are likely to exhibit an increased variance because of this.

For finite model orders and the situation $\mathcal{S} \in \mathcal{M}$, it is claimed in Gustavsson et al. (1977) that all methods (direct and indirect) lead to the same variance; however for indirect methods this result seems to hold true only for particular (ARMAX) model structures (Ljung and Forssel, 1997).

7. OVERVIEW AND EVALUATION OF PROPERTIES

The assessment criteria as discussed in section 4 have been evaluated for the several identification methods, and the results are listed in Table 1.

The methods that are most simply applicable are the direct method and the two-stage method. When considerable bias is expected from correlation between u and e, then the two-stage method should be preferred. For the identification of unstable plants the coprime factor, dual-Youla/Kucera and tailor-made parametrization method are suitable, of which the latter one seems to be most complex from an optimization point of view. When approximate -limited complexity- models are required, the coprime factor method is attractive. When additionally the controller is not accurately known, the two-stage method has advantages.

All methods are presented in a one-input, one-output configuration. The basic ideas as well as the main properties are simply extendable to MIMO systems.

The basic choice between direct and indirect approaches should be found in the evaluation of the following questions:

(a) Is there confidence in the fact that (G_0, H_0) and e satisfy the basic linear, time-invariant and limited order assumptions in the prediction error framework?
(b) Is there confidence in the fact that C operates as a linear time-invariant controller?

	Direct	Two-stage	Copr.fact.	Indir/Dual-Y.K.	Tailor-m
Consistency $(\hat{G}, \hat{H})$	+	+	+	+	□[5]
Consistency $\hat{G}$	-	+	+	+	□[5]
Tunable bias	-	+	+	+	+
Fixed model order	+	+	□[3]	-	+
Unstable plants	□[1]	-[2]	+	+	+
$(G(\theta^*), C)$ stable	-	-	-	□/+[4]	+
C assumed known	no	no	no	yes	yes

Table 1. Main properties of the different closed-loop identification methods.

The direct method takes an affirmative answer to (a) as a starting point. Its results are not dependent on controller linearity; however the method requires exact modelling in terms of question (a). The indirect methods are essentially dependent on an affirmative answer to (b), and might be more suitable to handle departures from aspect (a). So far the experimental setup has been considered where a single external signal r_1 is available from measurements. In all methods the situation of an available signal r_2 (in stead of r_1) can be treated similarly without loss of generality. A choice of a more principal nature is reflected by the assumption that the controller output is measured disturbance free. This leads to the (exact) equality

$$r = u + C(q)y.$$

The above equality displays that whenever u and y are available from measurements, knowledge of r and C is completely interchangeable. I.e. when r is measured, this generates full knowledge of C, through a noise-free identification of C on the basis of a short data sequence r, u, y. Consequently, for the indirect methods that are listed in Table 1, the requirement of having exact knowledge of C is not a limitation.

This situation is different when considering an experimental setup where the controller output (like the plant output) is disturbed by noise. Such a configuration is depicted in Figure 3, where d is an additional (unmeasurable) disturbance signal, uncorrelated with the other external signals r and v.

The appropriate relation now becomes

$$r + d = u + C(q)y$$

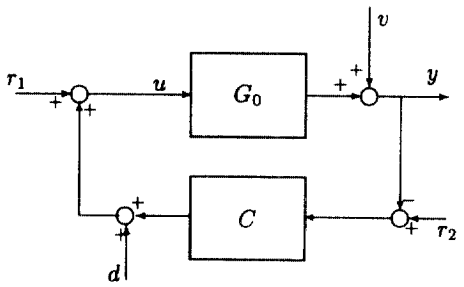

Fig. 3. Closed-loop configuration with disturbance on controller output.

and apparently now there does exist a principal difference between the information content in r and in knowledge of C. Two situations can be distinguished:

- r is available and C is unknown. In this case the indirect identification methods have no other option than to use the measured r as the external signal in the several methods. In this way the disturbance d will act as an additional disturbance signal in the loop that will lead to an increased variance of the model estimates.
- C is exactly known. In this case the signal $u + Cy$ can be exactly reconstructed and subsequently be used as the "external" signal in the several indirect methods. In this way the disturbance signal d is effectively used as an external input, leading to an improved signal to noise ratio in the estimation schemes, and thus to a reduced variance of the model estimates.

When measured signals u and y are given, one can argue what is the extra information content of having knowledge of r and/or C. From the comparative results of direct and indirect identification methods, one can conclude that this extra information allows the consistent identification of G_0, irrespective of the noise model H_0.

An additional aspect that may favour closed-loop experiments over open-loop ones, is the fact that a controller can have a linearizing effect on nonlinear plant dynamics; the presence of the controller can cause the plant to behave linearly in an appropriate working point.

[1] Only in those situations where the real plant (G_0, H_0) has an ARX or ARMAX structure.
[2] Not possible to identify unstable plants if in the second step attention is restricted to independently parametrized G and K.
[3] An accurate (high order) estimate of G_0 as well as knowledge of C is required; this information can be obtained from data.
[4] For the indirect method, stability is guaranteed only if C is stable.
[5] Consistency holds when the parameter set is restricted to a connected subset containing the exact plant vector θ_0.

For all identification methods discussed, the final estimation step comes down to the application of a standard (open-loop) prediction error algorithm. This implies that also the standard tools can be applied when it comes down to model validation (Ljung, 1987).

Attention has been restricted to prediction error methods related to a least-squares type criterion; instrumental variable type estimators (Söderström and Stoica, 1989; Zheng and Feng, 1995) can lead to related results, but are less suitable for specifying the asymptotic bias distribution.

8. IDENTIFICATION FOR CONTROL

8.1 *Introduction*

In many situations models are identified for the purpose of using them as a basis for subsequent model-based (robust) control design. In that case the evaluation of models has to be undertaken in the scope of the control design. In other words: the best identified model (within a specific class) is that model that leads to a controller that controls the plant best. In the area of "identification for control" this issue has been addressed from several perspectives. A detailed analysis is outside the scope of this paper. Here a brief discussion will be incorporated, focusing on those aspects in which closed-loop experiments are involved. For more extensive discussions the reader is referred to Gevers (1993), Ninness and Goodwin (1995), Van den Hof and Schrama (1995) and the references therein.

8.2 *Asymptotic bias of nominal models*

In many situations an identified model can only be an approximation of exact plant dynamics, due to the fact that the model set is restricted to models of limited order. Next to the fact that "physical" plants will seldom exhibit exact limited-order behaviour, this reduced-order modelling mechanism can also be motivated by the fact that in many (industrial) control situations, highly complex plants can be controlled satisfactorily by controllers that are based on fairly simple models. In constructing reduced order models one has to deal with the issue of unmodelled dynamics, and the model has to be tuned to accurately fit those plant dynamics that are most essential for the subsequent control design. Several schemes have been developed, directed towards the identification of reduced order (nominal) models, most of them based on the following mechanism.

Let $\|J(G_0, C_{\hat{G}})\|$ be a control performance cost function related to a closed-loop system with plant G_0 and controller $C_{\hat{G}}$ being designed on the basis of a plant model $\hat{G}$. One can think of J as e.g. a weighted sensitivity function:

$$J(G_0, C_{\hat{G}}) = \frac{V}{1 + C_{\hat{G}} G_0} \tag{28}$$

It is the aim to achieve a minimum value of $\|J(G_0, C_{\hat{G}})\|$, through an appropriate choice of $\hat{G}$ and $C_{\hat{G}}$. Employing the triangle inequality:

$$\begin{aligned} \Big| \|J(\hat{G}, C_{\hat{G}})\| - \|J(G_0, C_{\hat{G}}) - J(\hat{G}, C_{\hat{G}})\| \Big| \leq \\ \leq \|J(G_0, C_{\hat{G}})\| \leq \\ \leq \|J(\hat{G}, C_{\hat{G}})\| + \|J(G_0, C_{\hat{G}}) - J(\hat{G}, C_{\hat{G}})\|, \end{aligned} \tag{29}$$

shows that the achieved performance cost $\|J(G_0, C_{\hat{G}})\|$ can be minimized by minimizing each of the two separate terms on the right hand side of (29). Since such a minimization over $\hat{G}$ involves the control design $C_{\hat{G}}$, this will generally be intractable. In addition, iterative schemes have been proposed to minimize both terms separately: minimizing $\|J(\hat{G}, C)\|$ over C for a fixed model $\hat{G}$, and minimizing the performance degradation term $\|J(G_0, C) - J(G, C)\|$ over G for a fixed controller C. In that case this degradation term can be given the interpretation of a control-performance induced identification criterion:

$$\hat{G} = \arg\min_{G} \|J(G_0, C) - J(G, C)\|$$

which for the choice of J as given above takes the form:

$$\hat{G} = \arg\min_{G} \left\| \frac{V(G_0 - G)C}{(1 + CG_0)(1 + CG)} \right\|.$$

Note that for a 2-norm this criterion has the same structure as the bias expression for indirect closed-loop methods as shown in (25). By appropriate choice of Φ_{r_1} and/or K_* in (25) the criteria can be made the same. This situation also extends to other choices of performance functions J (Van den Hof and Schrama, 1995).

The message here is that the (reduced-order) model for which the bias is optimally tuned to the control performance cost for a given controller, is obtained by doing closed-loop identification with an indirect identification method.

8.3 *Asymptotic variance of nominal models*

From a variance point of view, optimal experiment design in view of an intended model application has been analyzed in Gevers and Ljung (1986), Ljung (1987), Hjalmarsson *et al.* (1996) and Gevers *et al.* (1997). In all cases this concerns an analysis in the situation $S \in \mathcal{M}$,

and both number of data and model order tend to infinity.

Considering a variance-based identification design criterion

$$J(\mathcal{D}) = \int_{-\pi}^{\pi} tr[P(\omega, \mathcal{D})\Gamma(\omega)]d\omega$$

where $P(\omega, \mathcal{D}) = cov([\hat{G}(e^{i\omega})\ \hat{H}(e^{i\omega})]^T)$, and $\mathcal{D}$ denotes the design choices with respect to the experimental conditions, represented by $\{\Phi_u, \Phi_{ue}\}$, while $\Gamma(\omega)$ is a 2×2 Hermitian matrix reflecting the intended application of the model. For this design criterion the following results are available:

- If $\Gamma_{12} = 0$, (e.g. the control design is only based on $\hat{G}$ and not on $\hat{H}$), and if the power of the input signal is limited, then the optimal experiment that minimizes $J(\mathcal{D})$ is an open-loop experiment with

$$\Phi_u^{opt} = c \cdot \sqrt{\Gamma_{11}(\omega)\Phi_v(\omega)}.$$

  For the control design criterion (28) this reduces to

$$\Phi_u^{opt} = c \cdot \frac{|C_{\hat{G}}V|}{|1 + C_{\hat{G}}G_)|^2}\sqrt{\Phi_v(\omega)} \qquad (30)$$

  (Ljung, 1987).
- If the control design is only based on $\hat{G}$ and if the output power is limited, then the optimal experiment is a closed-loop experiment. If the plant is minimum-phase, the optimal controller to apply during identification is the minimum variance controller (Gevers and Ljung, 1986).
- If $\Gamma_{12}(\omega) \neq 0$, (the control design is based on both $\hat{G}$ and $\hat{H}$), then the optimal experiment is a closed-loop experiment with an optimal controller that can be characterized (Hjalmarsson *et al.*, 1996).

It has to be noted that in the first situation, where open-loop experiments are optimal, the required input spectrum (30) is proportional to the (unknown) sensitivity function of the closed-loop to be constructed. For the present controller, this input shaping is exactly achieved by performing the experiment in closed-loop. However thia also has the negative effect of feeding back the output noise to the input.

The third situation provides an optimal controller to apply during identification that generally will be dependent on (unknown) plant information. This has motivated the proposition of an iterative mechanism of identification and renewed (closed-loop) experiment design, being analyzed and illustrated in Hjalmarsson *et al.* (1996).

8.4 *Uncertainty structure of identified model sets*

When models are used as a basis for robust control design, both a nominal model and a model uncertainty characterization are required. In recent years attention has been given to the quantification of model uncertainty on the basis of experimental data, see e.g. Ninness and Goodwin (1995) and the references therein. An "optimal" choice of a model uncertainty set, would be to collect all models that are not invalidated by the data and the prior information on the system. However in most cases this set can not be simply characterized in a form that is manageable for a control design procedure. Therefore one restricts attention to uncertainty sets of a prechosen nature, such as e.g. additive or multiplicative uncertainty, situated around a nominal model.

A general characterization of such a set is:

$$\mathcal{P}_f(\hat{G}, \gamma) = \{G \mid G = f(\hat{G}, \Delta), |\Delta(e^{i\omega})| \leq \gamma(\omega),\ \forall\omega\}$$

with γ a positive real-valued function of ω and f a linear fractional transformation:

$$f(\hat{G}, \Delta) = \hat{G} + P_{21}\Delta(1 - P_{11}\Delta)^{-1}P_{12}. \qquad (31)$$

Although such an uncertainty set generally is denoted as "unstructured", the choice of f does provide the set with a particular uncertainty structure. Note that an additive uncertainty results through the choice $P_{21} = P_{12} = 1$, $P_{11} = 0$.

For a particular choice of f and $\hat{G}$, the "size" γ of the set has to be chosen as small as possible so as to contain the real plant G_0. This situation is depicted in an abstract way in Figure 4, where the shaded area reflects the set of all unfalsified models which is encapsulated in the uncertainty set $\mathcal{P}$. Note that the former set is principally implied by the measurement data, whereas the latter set is partly just chosen by the user. It is clear that there are many options for choosing $\mathcal{P}$ such that all unfalsified models are contained. However in choosing this $\mathcal{P}$ it is apparent that one should take account of the performance cost function J, by avoiding the incorporation of (falsified) models (in the white area of Figure 4) that lead to poor performance costs. Such incorporation would lead to a control design with considerable conservatism. For discussions on the role of the uncertainty structure in this respect, see also Schrama (1992) and Van den Hof *et al.* (1994).

For a given $\hat{G}$, a performance-cost relevant uncertainty structure is obtained, if for all $G \in \mathcal{P}_f(\hat{G}, \gamma)$, $J(G, C)$ can be written as an affine expression in Δ, i.e.

$$J(G, C) = M_1 + M_2\Delta \qquad (32)$$

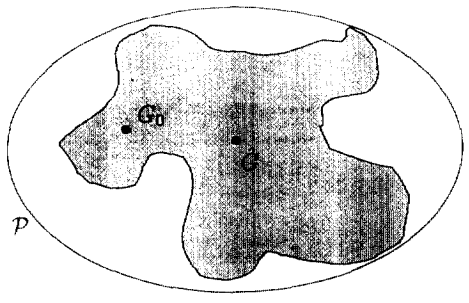

Fig. 4. Uncertainty set $\mathcal{P}_f(\hat{G},\gamma)$ (ellipsoid) and set of unfalsified models (shaded area).

With J chosen as a (weighted) closed-loop transfer function, as e.g. in (28), this implies that there is a direct linear relationship between the "size" γ of the uncertainty set and the (frequency-dependent) worst-case performance cost $|J(G,C)|$ over the set, in terms of

$$\sup_{G\in\mathcal{P}(\hat{G},\gamma)} |J(G,C)| = |M_1(e^{i\omega})| + \gamma(\omega)\cdot|M_2(e^{i\omega})| \quad \forall\omega.$$

In this way the "shape" of the uncertainty set is directly tuned towards the performance criterion. An affine relationship (32) can be achieved by using an uncertainty structure that is based on the dual Youla/Kucera parametrization:

$$\mathcal{P}(\hat{G},\gamma) = \{G \mid G = \frac{\hat{N} + D_c\Delta_R}{\hat{D} - N_c\Delta_R}, |\Delta_R(e^{i\omega})| \le \gamma(\omega)\}$$

with $C = N_c D_c^{-1}$ and $\hat{G} = \hat{N}\hat{D}^{-1}$. The expression for G above is a particular form of linear fractional transformation (31). For each $G \in \mathcal{P}(\hat{G},\gamma)$ it follows that

$$J(G,C) = \frac{V}{1+C\hat{G}} + M_2\Delta_R$$

with $M_2 = \dfrac{VN_c}{\hat{D}(1+C\hat{G})}$, being a filter dependent only on known elements. A data-based uncertainty modelling procedure, should provide the smallest bound $\gamma(\omega)$ that is required to guarantee that G_0 is an element of this set. This minimization of $\gamma(\omega)$ can be performed in a closed-loop experimental setup, by applying a model uncertainty estimation procedure to the Youla/Kucera parameter. In terms of the mechanism discussed in section 5.4 this refers to choosing the auxiliary model $G_x = \hat{G}$. This motivates the use of closed-loop experimental data not only for nominal model identification, but also for control-relevant uncertainty bounding. The indicated mechanism also extends to more general performance cost functions, as shown in de Callafon and Van den Hof (1997).

Aspects of model (in)validation, also in relation to control design, are treated e.g. in Smith and Doyle (1992) and Ljung and Guo (1997).

8.5 *Model-free tuning of controllers*

Another area where closed-loop experimental data is involved in an identification-type problem is data-based controller tuning. In the recent work of Hjalmarsson *et al.*(1994) a (simply structured) controller is directly tuned on the basis of a number of specifically designed closed-loop experiments. This powerful approach has already delivered a number of interesting application results.

9. CONCLUSIONS

Closed-loop experimental conditions should not be considered as a degenerate or unfavourable situation for identifying dynamical systems. There are many good methods available, both for identifying consistent models, as well as for handling the situation of unmodelled dynamics (approximate modelling). In this paper the characteristic properties of both direct and indirect procedures have been evaluated on the basis of explicit assessment criteria, including aspects of bias and variance. Additionally it is shown that closed-loop experiments can be particularly suitable in relation with model-based control design. The identification procedures discussed in this paper have been implemented in the Matlab toolbox CLOSID (Van den Hof *et al.*, 1997).

ACKNOWLEDGMENTS

Many people have contributed to the results presented in this paper. I gratefully acknowledge the contributions of colleagues at Delft University of Technology, in particular Raymond de Callafon. I also would like to thank Michel Gevers and Lennart Ljung for sharing their views and results with me.

REFERENCES

de Callafon, R.A., P.M.J. Van den Hof and M. Steinbuch (1993). Control-relevant identification of a compact disc pick-up mechanism. *Proc. 32nd IEEE Conf. Decision and Control*, San Antonio, TX, pp. 2050-2055.

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

Desoer, C.A., R.W. Liu, J. Murray and R. Saeks (1980). Feedback system design: the fractional representation approach to analysis and synthesis. *IEEE Trans. Automat. Contr.*, **AC-25**, 399-412.

Eek, R.A., J.A. Both and P.M.J. Van den Hof (1996). Closed-loop identification of a continuous crystallization process. *AIChE Journal*, **42**, 767-776.

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

Gevers, M. (1993). Towards a joint design of identification and control? In: H.L. Trentelman and J.C. Willems (Eds.), *Essays on Control: Perspectives in the Theory and its Applications*. Birkhäuser, Boston, pp. 111-151.

Gevers, M., L. Ljung and P.M.J. Van den Hof (1997). Asymptotic variance expressions for closed-loop identification and their relevance in identification for control. *Prepr. 11th IFAC Symp. System Identification (SYSID97)*, Fukuoka, Japan, Vol. 3, pp. 1449-1454.

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

Hansen, F.R. and G.F. Franklin (1988). On a fractional representation approach to closed-loop experiment design. *Proc. American Control Conf.*, Atlanta, GA, pp. 1319-1320.

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

Hjalmarsson, H., S. Gunnarsson and M. Gevers (1994). A convergent iterative restricted complexity control design scheme. *Proc. 33rd Conf. Decision and Control*, Lake Buena Vista, FL, pp. 1735-1740.

Landau, I.D. and K. Boumaiza (1996). An output error recursive algorithm for unbiased identification in closed-loop. *Proc. 13th IFAC World Congress*, San Francisco, CA, Vol. I, pp. 215-220.

Lee, W.S., B.D.O. Anderson, R.L. Kosut and I.M.Y. Mareels (1993). A new approach to adaptive robust control. *Int. J. Adapt. Contr. Sign. Proc.*, **7**, 183-211.

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

Ljung, L. (1993). Information content in identification data from closed-loop operation. *Proc. 32nd IEEE Conf. Dec. Control*, San Antonio, TX, pp. 2248-2252.

Ljung, L. and U. Forssell (1997). Variance results for closed-loop identification methods. *Proc. 36th IEEE Conf. Dec. Control*, San Diego, CA, pp. 2435-2440.

Ljung, L. and L. Guo (1997). Classical model validation for control design purposes. *Mathematical Modelling of Systems*, **3**, 27-42.

Mäkilä, P.M., J.R. Partington and T.K. Gustavsson (1995). Worst-case control-relevant identification. *Automatica*, **31**, 1799-1819.

Milanese, M. and A. Vicino (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica*, **27**, 997-1009.

Ninness, B.M. and G.C. Goodwin (1995). Estimation of model quality. *Automatica*, **31**, 1771-1797.

Partanen, A.G. and R.R. Bitmead (1995). The application of an iterative identification and controller design to a sugar cane crushing mill. *Automatica*, **31**, 1547-1563.

Schrama, R.J.P. (1992). *Approximate Identification and Control Design*. Dr. Dissertation, Mechan. Engin. Systems and Control Group, Delft Univ. Technology.

Smith, R.S. and J.C. Doyle (1992). Model validation: a connection between robust control and identification. *IEEE Trans. Autom. Control*, **AC-37**, 942-952.

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

Sung, H.K. and S. Hara (1988). Properties of sensitivity and complementary sensitivity functions in single-input single-output digital control systems. *Int. J. Control*, **48**, 2429-2439.

Tsai, M.C., E.J.M. Geddes and I. Postlethwaite (1992). Pole-zero cancellations and closed-loop properties of an $\mathcal{H}_\infty$-mixed sensitivity design problem. *Automatica*, **28**, 519-530.

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

Van den Hof, P.M.J., R.J.P. Schrama and P.M.M. Bongers (1994). On nominal models, model uncertainty and iterative methods in identification and control design. *Lect. Notes in Control and Inform. Sciences*, **192**, Springer Verlag, Berlin, pp. 39-50.

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

Van den Hof, P.M.J. and R.J.P. Schrama (1995). Identification and control – closed-loop issues. *Automatica*, **31**, 1751-1770.

Van den Hof, P.M.J. and R.A. de Callafon (1996). Multivariable closed-loop identification: from indirect identification to dual-Youla parametrization. *Proc. 35th IEEE Conf. Dec. Contr.*, Kobe, Japan, pp. 1397-1402.

Van den Hof, P.M.J., R.A. de Callafon and E.T. van Donkelaar (1997). CLOSID - A closed-loop system identification toolbox for Matlab. *Prepr. 11th IFAC Symp. System Identification (SYSID97)*, Fukuoka, Japan, Vol. 2, pp. 983-989. Software available through anonymous ftp at `ftp-mesc.wbmt.tudelft.nl` directory `pub/matlab/closid`.

van Donkelaar, E.T. and P.M.J. Van den Hof (1997). Analysis of closed-loop identification with a tailor-made parametrization. *Proc. 4th European Control Conference*, Brussels, Belgium, Vol. 4, Paper TH-A-F4.

Vidyasagar, M. (1985). *Control System Synthesis - A Factorization Approach*. MIT Press, Cambridge, MA.

Zheng, W.X. and C.B. Feng (1995). A bias-correction method for indirect identification of closed-loop systems. *Automatica*, **31**, 1019-1024.