Pergamon

*Automatica*, Vol 30 No 4 pp 679–690, 1994


0005-1098/94 $6 00 + 0 00

# Approximate Identification with Closed-loop Performance Criterion and Application to LQG Feedback Design*

RICHARD G HAKVOORT,† RUUD J P SCHRAMA‡ and PAUL M J VAN DEN HOF†

*An iterative approach of identification with properly filtered signals and control design appears to yield a nominal model that is better suited for feedback design than a model resulting from an unweighted open-loop identification.*

**Key Words**—System identification, closed-loop systems, feedback control, control system design, iterative methods

**Abstract**—A model-based controller generally works better with the model than with the modelled plant due to the modelling error. This difference between the performances can be made small by selecting a model that is accurate at the closed-loop relevant frequencies. In this paper it is shown that an iterative approach of identification and control design can lead to a model that is much better suited for feedback design than a model resulting from an unweighted open-loop identification. In this iteration each identification is performed such that a certain closed-loop criterion function is minimized. This is accomplished by closed-loop identification with persistent set-point excitation and a proper signal filtering. Each control design step employs the latest identified model to construct an LQG compensator. The performance requirements are gradually increased during the iteration.

## 1 INTRODUCTION

THE DESIGN OF a linear control system is frequently based on a model of the plant under consideration. A model is very unlikely to be an exact description of the system. Due to the model error the performance of the controller, designed for the model, will not be obtained when the controller is applied to the real system. Obviously, in order to have a controlled system performance that is close to the designed performance for the model, the model error should be tuned towards the control objective. The need for a high accuracy near the cross-over frequency is well recognized. This is however only a partial answer to the problem. In its generality the question is how to properly define a closed-loop relevant way of evaluating the model error.

When the model is obtained from identification experiments, the problem above boils down to the problem of finding a closed-loop relevant identification criterion. Several authors have paid attention to this problem. An *ad hoc* solution is obtained in Balas and Doyle (1990), which addresses a control problem with a prespecified bandwidth. In Gevers and Ljung (1986) experiment design for minimum variance control is studied in the prediction error identification framework. The closed-loop of the model-based controller is compared with the closed-loop of the optimal, system-based, controller. The difference of the resulting two closed-loops is shown to be minimized by an optimal identification experiment. In Rivera *et al* (1990) the prediction error method is applied as well, but there a desired sensitivity is used as a weighting function for open-loop identification.

In this paper we will not use desired feedback transfer function matrices, nor will we use knowledge of a plant-based controller. Instead we take the following starting point. In order to identify a model that is suitable for control design, we should be able to identify a model that accurately describes the closed-loop relevant system properties in the presence of a given compensator. The main problem considered in

* Received 6 April 1992, revised 11 November 1992, revised 24 May 1993, received in final form 22 June 1993. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor B Wahlberg under the direction of Editor T Söderström. Corresponding author R G Hakvoort. E-mail hakvoort@tudw03 tudelft nl

† Mechanical Engineering Systems and Control Group, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

‡ Now with the Royal Dutch/Shell Group

this paper is how to perform an identification experiment such that the resulting model in closed-loop optimally resembles the system in closed-loop for a give compensator If such a model has been identified, it can subsequently be used to design a new compensator with slightly increased performance requirements, which in this paper correspond to a higher bandwidth of a servo-controller The rationale is that the model will still be a good representation of the plant for a new compensator, provided that this new compensator differs not too much from the previous one Therefore the performance improvement that is achieved for the model is expected to be achieved for the modelled system as well Next a new identification is carried out in order to obtain a model that accurately describes the system for the new compensator, and the entire procedure is repeated until a satisfactory controller performance is achieved A similar iterative approach has also been suggested in Zang *et al* (1991) for LQ control design In Schrama (1992a, b) it is shown that such an iterative scheme of identification and controller design is actually necessary for high performance control design

In the light of this iterative scheme, we will analyze the identification problem mentioned above A solution will not only be shown to exist, but also to be simply applicable using standard identification tools As a control design method we will use LQG feedback design We utilize the prediction error identification method, Ljung (1987), and the concept of a performance criterion as introduced in Gevers and Ljung (1986) We will only consider the asymptotic bias contribution to this performance criterion, variance aspects will not be considered A closed-loop performance criterion is defined and the design variables of the prediction error method are chosen such that this criterion function is actually minimized by the identification procedure This makes the identification criterion compatible with the LQG control objective Using an iterative procedure of closed-loop relevant identification and control design, a model for high-performance control design is constructed, that could not have been obtained from open-loop experiments alone

Preliminary results on this problem have been published in Hakvoort (1990) and more recently in Hakvoort *et al* (1992) The LQG objective has also been addressed in Bitmead *et al* (1990), but there the identification procedure minimizes a model error that pertains to robust stability rather than to robust performance

The outline of the paper is as follows In the next section the prediction error identification procedure is summarized In Section 3 we define the closed-loop performance criterion of concern In Section 4 we adjust the prediction error method such that this criterion function is actually minimized Then in Section 5 we consider an example in which the iterative scheme is put into practice for a particular LQG control objective In Section 6 we discuss the results and we make some general observations concerning the interplay between identification and control design The paper ends with a summary and conclusions

## 2 PREDICTION ERROR IDENTIFICATION

In this section we adopt the relevant aspects of prediction error identification from Ljung (1987)

Consider a discrete-time representation of a linear, time-invariant SISO system with additive stochastic disturbances

$$\mathcal{S}\quad y(t) = G_0(q)u(t) + v(t) = G_0(q)u(t) + H_0(q)e(t), \tag{1}$$

where $G_0(q)$ is the deterministic and $H_0(q)$ the stochastic part of the plant, $u(t)$ and $y(t)$ are respectively the input and output at time $t$ and $e(t)$ is discrete white noise with zero mean value, $q$ is the shift operator $qu(t) = u(t+1)$ The leading coefficient of $H_0(q)$ is one

We choose a model set with a fixed noise model,

$$\mathcal{M}\quad y(t) = G(q,\theta)u(t) + \hat{H}_f(q)\varepsilon(t), \tag{2}$$

where $\varepsilon(t)$ is the one step ahead prediction error $G(q,\theta)$ and $\hat{H}_f(q)$ are defined analogously to $G_0(q)$ and $H_0(q)$ The leading coefficient of $\hat{H}_f(q)$ is one We do not assume that the true system $\mathcal{S}$ is in the model set $\mathcal{M}$ For notational convenience we introduce

$$T_0(q) = [G_0(q)\quad H_0(q)], \qquad T(q,\theta) = [G(q,\theta)\quad \hat{H}_f(q)] \tag{3}$$

An estimate of $\theta$ is obtained by minimizing the quadratic norm of the prediction error with respect to $\theta$

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N}\sum_{t=0}^{N-1} \varepsilon^2(t,\theta), \tag{4}$$

with $N$ the number of samples This yields the estimate $T(q,\hat{\theta})$

In Ljung (1987) it is shown that under weak conditions the asymptotic parameter estimate is given by

$$\lim_{N\to\infty} \hat{\theta} = \theta^* = \arg\min_{\theta} E\varepsilon^2(t,\theta) \quad \text{w p 1} \tag{5}$$

According to Janssen (1988) this can be given the frequency domain interpretation

$$\theta^* = \arg\min_{\theta} \int_{-\pi}^{\pi} \frac{\tilde{T}(e^{i\omega}, \theta)\Phi(\omega)\tilde{T}^T(e^{-i\omega}, \theta)}{|\hat{H}_f(e^{i\omega})|^2} d\omega, \tag{6}$$

where $\tilde{T}$ is the model error defined as

$$\tilde{T}(q, \theta) = T(q, \theta) - T_0(q) \tag{7}$$

and $\Phi(\omega)$ is the spectrum

$$\Phi(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{ue}(\omega) \\ \Phi_{eu}(\omega) & \Phi_e(\omega) \end{bmatrix}, \tag{8}$$

with $\Phi_u(\omega)$ the spectrum of $u(t)$ and $\Phi_{ue}(\omega)$ the cross-spectrum of $u(t)$ and $e(t)$ This result (6) also holds under closed-loop conditions In that case a necessary requirement is that either the controller or the model $G(q, \theta)$ and the system $G_0(q)$ have one delay, see Janssen (1988)

The input spectrum $\Phi_u(\omega)$ and the cross-spectrum $\Phi_{ue}(\omega)$ dictate the frequency distribution of the model error (see Wahlberg and Ljung, 1986, for details) As these design variables are at our disposal, we can choose them such that the model is optimal in view of the intended use We signify the design variables as

$$\mathcal{D} = \{\Phi_u(\omega), \Phi_{ue}(\omega)\} \tag{9}$$

The spectrum $\Phi_u(\omega)$ can be specified by an open-loop input design, but a nonzero $\Phi_{ue}(\omega)$ can be realized only by introducing feedback in the identification and specifying some external reference signal Of course one cannot assign $\Phi_e(\omega)$ as this would be in contradiction with the nature of a noise

## 3 A CLOSED-LOOP PERFORMANCE CRITERION

In this section we will define a closed-loop performance criterion to measure the model's capacity to describe the controlled operation of the plant We will define the criterion function for a given controller, irrespective of the applied control design technique At a later stage this controller will be determined by means of LQG feedback design

We consider the closed-loop configuration of Fig. 1, in which the plant is controlled by the fixed two-component controller $(C_1, C_2)$ which is assumed to be known The feedback system is driven by an external disturbance $\bar{v}$ and an external reference signal $\bar{r}$, which are assumed to be mutually uncorrelated The bars represent the operational conditions under which the model must appropriately describe the plant Hence $\bar{r}$ and $\bar{e}$ are signals with fixed spectra determined by the operational conditions

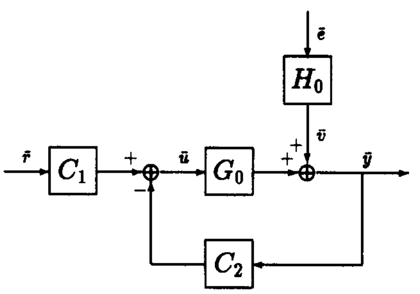


FIG 1 System in closed-loop

The output $\bar{y}$ satisfies

$$\bar{y}(t) = \frac{G_0(q)C_1(q)}{1 + G_0(q)C_2(q)}\bar{r}(t) + \frac{H_0(q)}{1 + G_0(q)C_2(q)}\bar{e}(t) \tag{10}$$

A similar equation can be written down for the model by replacing $G_0(q)$ and $H_0(q)$ with $G(q, \hat{\theta})$ and $\hat{H}_f(q)$ The model is a good closed-loop description of the system if the error terms

$$\frac{G(q, \hat{\theta})C_1(q)}{1 + G(q, \hat{\theta})C_2(q)} - \frac{G_0(q)C_1(q)}{1 + G_0(q)C_2(q)}$$

and

$$\frac{\hat{H}_f(q)}{1 + G(q, \hat{\theta})C_2(q)} - \frac{H_0(q)}{1 + G_0(q)C_2(q)},$$

are small in an $H_2$-sense We define a closed-loop performance criterion $J_1(\mathcal{D})$ as the 2-norm of these error terms weighted with the signal spectra,

$$J_1(\mathcal{D}) = \int_{-\pi}^{\pi} \left( \left| \frac{G(e^{i\omega}, \hat{\theta}(\mathcal{D}))C_1(e^{i\omega})}{1 + G(e^{i\omega}, \hat{\theta}(\mathcal{D}))C_2(e^{i\omega})} - \frac{G_0(e^{i\omega})C_1(e^{i\omega})}{1 + G_0(e^{i\omega})C_2(e^{i\omega})} \right|^2 \Phi_{\bar{r}}(\omega) + \left| \frac{\hat{H}_f(e^{i\omega})}{1 + G(e^{i\omega}, \hat{\theta}(\mathcal{D}))C_2(e^{i\omega})} - \frac{H_0(e^{i\omega})}{1 + G_0(e^{i\omega})C_2(e^{i\omega})} \right|^2 \Phi_{\bar{e}}(\omega) \right) d\omega, \tag{11}$$

where the argument $\mathcal{D}$ has been added to emphasize the dependency of the identification result on the design variables Note that, due to the presence of the term $\hat{H}_f/(1 + \hat{G}C_2)$ in the criterion (11), this criterion function is different from the one used in Zang *et al* (1991) Also note that this criterion has been formulated for fixed $\Phi_{\bar{r}}$ and $\Phi_{\bar{e}}$ We will not directly address the problem of designing an optimal $\Phi_{\bar{r}}$, i e we will not optimize the criterion function (11) over $\Phi_{\bar{r}}$, but assume $\Phi_{\bar{r}}$ (and $\Phi_{\bar{e}}$) to be determined by the operational conditions of the plant The criterion function (11) is small if the closed-loop of the model is close to the closed-loop of the system in respect to the spectra of the signals $\bar{r}$ and $\bar{e}$ that drive the feedback system The following proposition gives a useful alternative expression for $J_1(\mathscr{D})$

*Proposition* 3 1 The performance criterion $J_1(\mathscr{D})$ satisfies

$$J_1(\mathscr{D}) = \int_{-\pi}^{\pi} \frac{1}{|1 + G(e^{i\omega}, \hat{\theta}(\mathscr{D}))C_2(e^{i\omega})|^2} \times \bar{T}(e^{i\omega}, \hat{\theta}(\mathscr{D}))\bar{\Phi}(\omega)\bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathscr{D}))\, d\omega, \tag{12}$$

where $\bar{T}(q, \theta)$ is given by equation (7), and according to Fig 1 the signal $\bar{u}(t)$ satisfies

$$\bar{u}(t) = \frac{C_1(q)}{1 + G_0(q)C_2(q)}\bar{r}(t) - \frac{H_0(q)C_2(q)}{1 + G_0(q)C_2(q)}\bar{e}(t) \tag{13}$$

and $\bar{\Phi}(\omega)$ satisfies

$$\bar{\Phi}(\omega) = \begin{bmatrix} \Phi_{\bar{u}}(\omega) & \Phi_{\bar{u}\bar{e}}(\omega) \\ \Phi_{\bar{e}\bar{u}}(\omega) & \Phi_{\bar{e}}(\omega) \end{bmatrix} \tag{14}$$

*Proof.* See Appendix B □

We want to formulate an identification procedure such that this performance criterion $J_1(\mathscr{D})$ is minimized More precisely the objective is to determine the optimal design variables

$$\mathscr{D}_{1\,\mathrm{opt}} = \arg \min_{\mathscr{D}} J_1(\mathscr{D}) \tag{15}$$

According to (9) these design variables consist of $\Phi_u(\omega)$ and $\Phi_{ue}(\omega)$, the signal spectra during identification Once more it is emphasized that these spectra correspond to the identification stage, and are free to choose, they are design variables This is in contrast with the spectra $\Phi_{\bar{u}}(\omega)$ and $\Phi_{\bar{u}\bar{e}}(\omega)$, which are fixed as they are determined by the operational conditions of the plant If the identification is carried out according to the optimal choice of design variables, then the resulting model is an optimal closed-loop description of the system. Note that this optimal design $\mathscr{D}_{1,\mathrm{opt}}$ possibly is a function of the chosen controller $(C_1, C_2)$ as the criterion function (11) is a function of this controller.

## 4 OPTIMAL IDENTIFICATION STRATEGY

In this section we will derive the optimal choice of design variables such that the closed-loop performance criterion $J_1(\mathscr{D})$ defined in (11) is minimized First we recapitulate some theory presented in Gevers and Ljung (1986), where the general scalar criterion $J_G(\mathscr{D})$,

$$J_G(\mathscr{D}) = \int_{-\pi}^{\pi} \bar{T}(e^{i\omega}, \hat{\theta}(\mathscr{D}))\Gamma(\omega) \times \bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathscr{D}))\, d\omega, \tag{16}$$

has been introduced as a measure for the model quality In here $\Gamma(\omega)$ is a $2\times2$ Hermitian weighting matrix that describes the relative importance of a good fit at different frequencies depending on the intended use of the model

For the number of samples increasing to infinity, it is shown in Gevers and Ljung (1986) that

$$\lim_{N\to\infty} J_G(\mathscr{D}) = J_B(\mathscr{D}) = \int_{-\pi}^{\pi} \bar{T}(e^{i\omega}, \theta^*(\mathscr{D}))\Gamma(\omega) \times \bar{T}^T(e^{-i\omega}, \theta^*(\mathscr{D}))\, d\omega \quad \text{w p 1}, \tag{17}$$

where $J_B$ is a bias-contribution to the performance criterion $J_G$. We consider the optimization problem

$$\mathscr{D}_{\mathrm{opt}} = \arg \lim_{\mathscr{D}} J_B(\mathscr{D}) \tag{18}$$

In Gevers and Ljung (1986) this optimization problem has been solved by matching the criterion function (17) to the criterion that is minimized in the identification procedure (6). In this way it is achieved that the identification (6) actually performs the desired minimization (17) The formal result is given in the next Theorem

*Theorem* 4 1 (*Gevers and Ljung,* 1986) The optimal choice of design variables (18) is given by

$$\frac{\Phi_{u,\mathrm{opt}}(\omega)}{|\hat{H}_f(e^{i\omega})|^2} = c\Gamma_{11}(\omega), \quad \frac{\Phi_{ue\,\mathrm{opt}}(\omega)}{|\hat{H}_f(e^{i\omega})|^2} = c\Gamma_{12}(\omega), \tag{19}$$

where $\Gamma_{ij}$ is the $i$th row, $j$th column entry of $\Gamma$ and $c$ is an arbitrary positive constant.

We intend to apply this result to the situation of the performance criterion (12) This can however not be done straightforwardly The reason is that the criterion function $J_G$ in (16) is quadratic in the model error, while $J_1$ is not a quadratic criterion function as the corresponding weight $\Gamma(\omega)$ would depend on $G(q, \hat{\theta}(\mathscr{D}))$ We proceed by first introducing the auxiliary quadratic performance criterion $J_2$ as

$$J_2(\mathscr{D}) = \int_{-\pi}^{\pi} \frac{1}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} \times \bar{T}(e^{i\omega}, \hat{\theta}(\mathscr{D}))\bar{\Phi}(\omega)\bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathscr{D}))\, d\omega, \tag{20}$$

where $\hat{G}_f$ is some fixed model This criterion function is quadratic in the model error, with the (constant) weighting matrix $\Gamma(\omega)$ given by

$$\Gamma(\omega) = \frac{\bar{\Phi}(\omega)}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} \tag{21}$$

So Theorem 4.1 can be straightforwardly applied to find

$$\mathcal{D}_{2,\text{opt}} = \arg\min_{\mathcal{D}} J_2(\mathcal{D}), \tag{22}$$

which is for $c = 1$ given by

$$\mathcal{D}_{2,\text{opt}} = \begin{cases} \Phi_{u,\text{opt}}(\omega) = \dfrac{\Phi_{\bar{u}}(\omega)\,|\hat{H}_f(e^{i\omega})|^2}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} \\ \Phi_{ue,\text{opt}}(\omega) = \dfrac{\Phi_{\bar{u}\bar{e}}(\omega)\,|\hat{H}_f(e^{i\omega})|^2}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} \end{cases}. \tag{23}$$

Next we define the discrepancies $\delta_1(\omega)$ and $\delta_2(\omega)$ as

$$\delta_1(\omega) = \frac{1}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} - \frac{1}{|1 + G(e^{i\omega}, \hat{\theta}(\mathcal{D}_{1,\text{opt}}))C_2(e^{i\omega})|^2}, \tag{24}$$

$$\delta_2(\omega) = \frac{1}{|1 + G(e^{i\omega}, \hat{\theta}(\mathcal{D}_{2,\text{opt}}))C_2(e^{i\omega})|^2} - \frac{1}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2}, \tag{25}$$

which are of use in the next Theorem

*Theorem* 4.2 Consider the performance criterion defined by (11) If a fixed noise model $\hat{H}_f(q) = 1 + \hat{G}_f(q)C_2(q)$ is used in the prediction error identfication and if the number of samples tends to infinity then the choice of design variables

$$\mathcal{D}_{2,\text{opt}} = \begin{cases} \Phi_{u,\text{opt}}(\omega) = \Phi_{\bar{u}}(\omega) \\ \Phi_{ue\,\text{opt}}(\omega) = \Phi_{\bar{u}\bar{e}}(\omega) \end{cases}, \tag{26}$$

converges to the optimal solution $\mathcal{D}_{1\,\text{opt}}$ if $\delta_1(\omega)$ and $\delta_2(\omega)$ converge to zero, i e

$$\lim_{\delta_1\,\delta_2 \to 0} J_1(\mathcal{D}_{2,\text{opt}}) - J_1(\mathcal{D}_{1\,\text{opt}}) = 0 \tag{27}$$

*Proof* If the fixed noise model satisfies $\hat{H}_f(q) = 1 + \hat{G}_f(q)C_2(q)$ then the design (26) is identical to the optimal design (23). This optimal solution has of course the property that $J_2(\mathcal{D}_{2,\text{opt}}) \le J_2(\mathcal{D}_{1,\text{opt}})$ Using this and Proposition 3 1 we obtain

$$\begin{aligned}
&J_1(\mathcal{D}_{2,\text{opt}}) - J_1(\mathcal{D}_{1,\text{opt}}) \\
&\quad = (J_1(\mathcal{D}_{2,\text{opt}}) - J_2(\mathcal{D}_{2,\text{opt}})) \\
&\qquad + (J_2(\mathcal{D}_{2,\text{opt}}) - J_2(\mathcal{D}_{1,\text{opt}})) \\
&\qquad + (J_2(\mathcal{D}_{1,\text{opt}}) - J_1(\mathcal{D}_{1,\text{opt}})) \\
&\quad \le (J_1(\mathcal{D}_{2,\text{opt}}) - J_2(\mathcal{D}_{2,\text{opt}})) \\
&\qquad + (J_2(\mathcal{D}_{1,\text{opt}}) - J_1(\mathcal{D}_{1,\text{opt}})) \\
&\quad = \int_{-\pi}^{\pi} \left( \frac{1}{|1 + G(e^{i\omega}, \hat{\theta}(\mathcal{D}_{2,\text{opt}}))C_2(e^{i\omega})|^2} - \frac{1}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} \right) \\
&\qquad \times \bar{T}(e^{i\omega}, \hat{\theta}(\mathcal{D}_{2,\text{opt}}))\bar{\Phi}(\omega) \\
&\qquad \times \bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathcal{D}_{2,\text{opt}}))\,d\omega \\
&\qquad + \int_{-\pi}^{\pi} \left( \frac{1}{|1 + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})|^2} - \frac{1}{|1 + G(e^{i\omega}, \hat{\theta}(\mathcal{D}_{1,\text{opt}}))C_2(e^{i\omega})|^2} \right) \\
&\qquad \times \bar{T}(e^{i\omega}, \hat{\theta}(\mathcal{D}_{1\,\text{opt}}))\bar{\Phi}(\omega) \\
&\qquad \times \bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathcal{D}_{1,\text{opt}}))\,d\omega \\
&\quad = \int_{-\pi}^{\pi} \delta_2(\omega)\bar{T}(e^{i\omega}, \hat{\theta}(\mathcal{D}_{2,\text{opt}})) \\
&\qquad \times \bar{\Phi}(\omega)\bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathcal{D}_{2,\text{opt}}))\,d\omega \\
&\qquad + \int_{-\pi}^{\pi} \delta_1(\omega)\bar{T}(e^{i\omega}, \hat{\theta}(\mathcal{D}_{1,\text{opt}})) \\
&\qquad \times \bar{\Phi}(\omega)\bar{T}^T(e^{-i\omega}, \hat{\theta}(\mathcal{D}_{1,\text{opt}}))\,d\omega \to 0 \\
&\qquad\qquad \text{if } \delta_1(\omega) \to 0,\ \delta_2(\omega) \to 0.
\end{aligned} \tag{28}$$

□

This means that the choice of the design variables (26) generally is a good choice, and it is even the best possible design (in a quadratic error sense) if both $\delta_1$ and $\delta_2$ vanish From equation (25) it follows that $\delta_2$ is small if $\hat{G}_f(q)$ is close to $G(q, \hat{\theta}(\mathcal{D}_{2,\text{opt}}))$, which is the result of the identification conducted according to Theorem 4 2; more specifically, the corresponding sensitivity functions have to be similar This discrepancy $\delta_2$ can be calculated afterwards Moreover it can be reduced to an arbitrarily small value by an iterative procedure In each step of this iteration $\hat{G}_f(q)$ is chosen as the identification result of the previous step This means that the fixed noise model $\hat{H}_f(q) = 1 + \hat{G}_f(q)C_2(q)$ is determined iteratively, in an inner-loop iteration that is independent of the iteration of identification and control design outlined in the introduction In this inner-loop iteration an optimal nominal model is identified for a fixed controller Based on the estimated $G(q, \hat{\theta})$ a new fixed noise model is constructed and a new model $G(q, \hat{\theta})$ is estimated with this new fixed noise model, based on one and the same data set Wahlberg and Ljung (1986) have shown that prefiltering the data $u(t)$ and $y(t)$ with a stable linear filter $L(q)$ is equivalent to changing the noise model $\hat{H}_f(q)$ to $\hat{H}_f(q)L^{-1}(q)$ Hence the choice of a fixed noise model $\hat{H}_f(q) =$

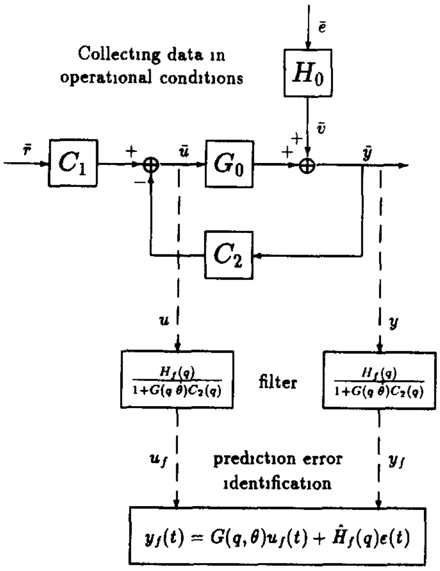


FIG 2 Optimal identification strategy

$1 + \hat{G}_f(q)C_2(q)$ may in practice be realized by applying a filter $L(q) = (1 + \hat{G}_f(q)C_2(q))^{-1}$ in combination with output error identification (noise model is fixed to one)

From equation (24) it follows that $\delta_1$ is small if $\hat{G}_f(q)$ is close to the (unknown) optimal identification result $G(q, \hat{\theta}(\mathscr{D}_{1,\mathrm{opt}}))$ This discrepancy $\delta_1$ cannot be determined precisely, but it is small if for example the modelling error is made sufficiently small, i e if both $G(q, \hat{\theta}(\mathscr{D}_{1,\mathrm{opt}}))$ and $G(q, \hat{\theta}(\mathscr{D}_{2,\mathrm{opt}}))$ (or equivalently $\hat{G}_f(q)$) are close to the real system $G_0(q)$ If $\delta_1$ or $\delta_2$ are not zero (which may often happen in practice) then $\mathscr{D}_{2\ \mathrm{opt}}$ is in general not equal to $\mathscr{D}_{1,\mathrm{opt}}$ In that case the design (26) is not optimal any more, but because of continuity considerations it is still expected to be a very good design

The optimal identification strategy derived is visualized in Fig 2 It says that the input spectrum (and the cross-spectrum of noise and input) in the identification experiment should be the same as those in the operational conditions (Fig 1), which means identification in closed-loop The data collected under operational conditions have to be properly filtered in order to obtain the optimal model The interpretation of the optimal identification is that it includes a weight at those frequencies where the closed-loop of the plant is close to the stability margin ($\bar{u}$ contains much energy) and/or where the closed-loop of the model is close to the stability margin ($L(q)$ has a large gain) We notice that in the identification procedure no perfect knowledge of the true system $T_0(q)$ is required, which is a very attractive property It is mentioned that in practice the identification procedure will only work if identifiability is ensured by using a persistent exciting external reference signal $\bar{r}$ In Appendix A the result of the optimal identification strategy is extended to the MIMO case

As clarified in the introduction this optimal identification strategy, derived for a given controller $(C_1, C_2)$, can be combined with an iterative scheme of identification and controller desgin in order to arrive at a high-performing controller This iterative scheme is visualized in Fig 3 As explained the iteration may contain subiterations at the moment that an optimal model is identified as the prefilter is dependent on the (unknown) identification result $G(q, \hat{\theta})$ So the inner-loop iteration in Fig 3 corresponds to the iterative prefilter (or fixed noise model) design, for which no new measurements are needed The outer-loop iteration involves the implementation of a new controller and collecting new data

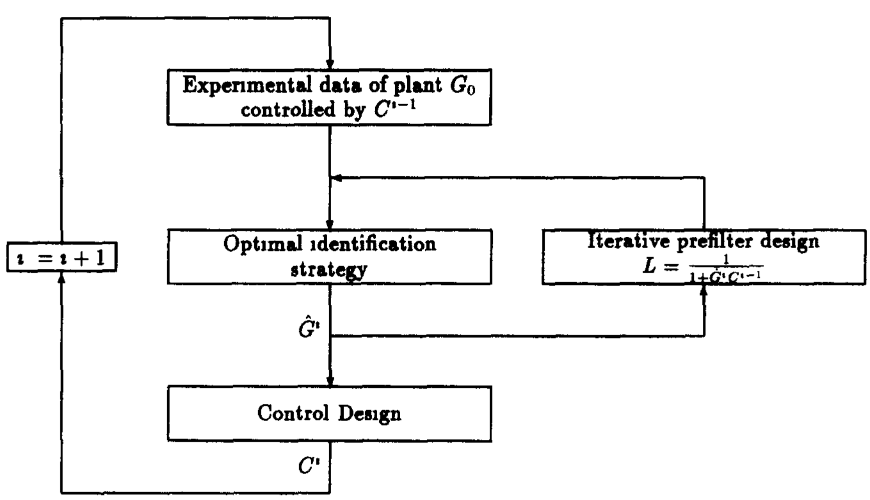


FIG 3 Iterative scheme of identification and control design, $i = 1, 2, 3, \ldots, C^0 = 0$

5 APPLICATION TO LQG FEEDBACK DESIGN

The theory of the previous sections has been developed without making assumptions about a specific controller design method. In the example of this section we will employ one particular controller design procedure, viz LQG feedback design, in order to illustrate the presented identification procedure We shortly summarize the relevant topics For a more detailed discussion the reader is referred to, for example, Maciejowski (1989, ch 5)

Consider the discrete time SISO model $\mathcal{M}$ for which a controller has to be designed

$$\mathcal{M} \quad \begin{cases} x(t+1) = Ax(t) + Bu_m(t) + Fw(t) \\ y_m(t) \quad = Cx(t) + v(t) \end{cases}, \quad (29)$$

where $w$ and $v$ are zero-mean white noises with covariance matrices

$$E\{ww^T\} = W \geq 0, \quad E\{vv^T\} = V > 0, \quad E\{wv^T\} = 0 \quad (30)$$

The signal $u_m$ is the control signal to the model and $y_m$ is the output of the model Now the LQG problem is to devise a feedback control law which minimizes the cost function

$$J_{\mathrm{LQG}} = \lim_{N \to \infty} E\left\{\frac{1}{N} \sum_{t=0}^{N-1} (x^T Q x + u_m^T R u_m)\right\}, \quad (31)$$

with $Q$ a positive semi-definite weighting matrix and $R$ a positive definite weighting matrix.

There are several weighting matrices that we can freely choose We want to investigate the impact of the identification procedure on the quality of the resulting controller and not the impact of the design weight. Therefore we pragmatically fix the weighting matrices,

$$F = B, \quad W = 1, \quad V = c, \quad Q = C^T C, \quad R = c \quad (32)$$

Then the LQG criterion function becomes

$$J_{\mathrm{LQG}} = \lim_{N \to \infty} E\left\{\frac{1}{N} \sum_{t=0}^{N-1} (y_m^2(t) + cu_m^2(t))\right\} \quad (33)$$

and it implies that the white noise $w$ is assumed to be additive at the input $u_m$ The latter is equivalent to stating that a white noise external reference signal $r$ enters at the input This actually determines the operational conditions in the Figs 1 and 2, i e $C_1(q) = 1$ and $\bar{r}$ is white noise

The parameter $c$ is the only design variable that is left and we will use it to establish the performance requirements on the controller A relatively small value of $c$ gives less weight to $u_m$ in the criterion function, and the output is assumed to be disturbed less, which gives rise to a tighter feedback-loop This will generally also lead to less robustness, even though no LQG controller optimizes robustness at all.

Now we apply the optimal identification procedure derived in the previous sections in combination with this fixed controller design procedure, performing an iteration of identification and feedback design. We use low order models in order to emphasize the effects due to undermodelling and use 4000 samples in the identification such that the variance effects can be neglected We compare the outcome of the iteration with the result of a direct open-loop identification The simulation example is carried out in continuous time due to the availability of software to design continuous time LQG controllers This means that the discrete time models that result from the identification are transformed to continuous time, assuming zero order hold. The error introduced by this transformation is very small as the sampling rate has been chosen high

We consider the fifth-order system shown in Fig. 4 Open-loop measurements are carried out with a white noise input signal and about 7% coloured noise being added to the output. Also in Fig. 4 the result of the open-loop identification of a strictly proper third-order output error model is given The low-frequency fit appears to be very good Next we design an LQG controller for the model, choosing $c = 0\,0002$ In Fig 5 the Bode diagram of this controller is shown In

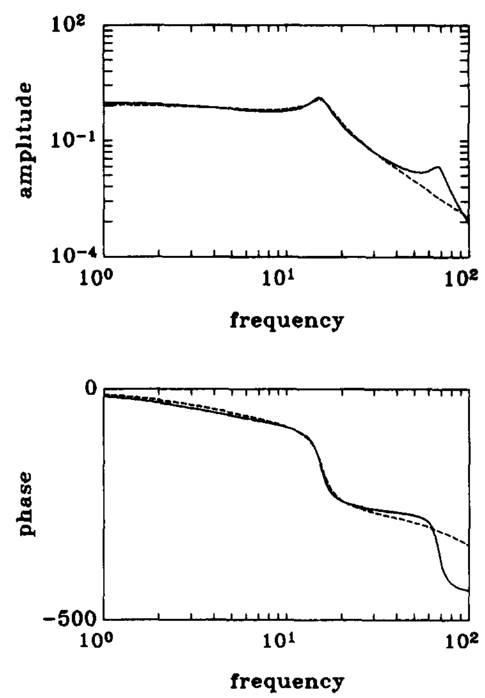


FIG 4 Bode diagram fifth-order system (solid) and third-order output-error model (dashed) obtained from open-loop experiments

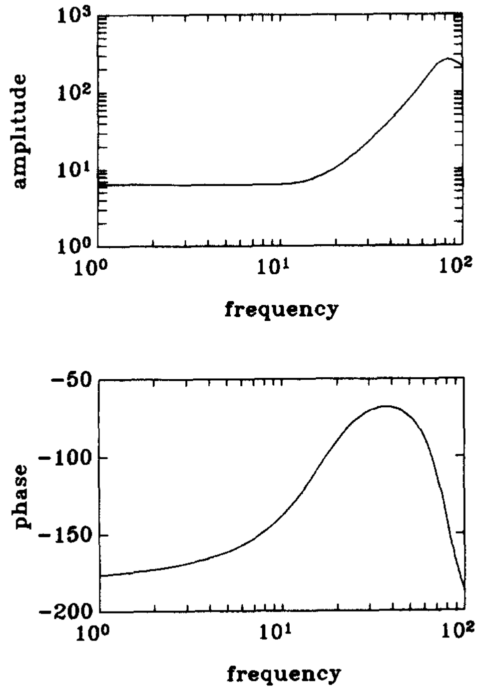


FIG 5 Bode diagram controller designed for the model identified in open-loop with $c = 0\ 0002$

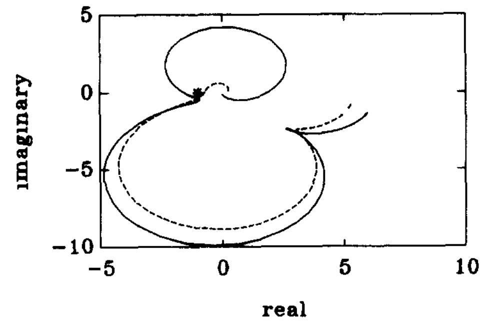


FIG 6 Nyquist diagram of $\hat{G}C$ (dashed) and $G_0C$ (solid), where $\hat{G}$ is the model identified in open-loop and $C$ is the LQG controler designed for this open-loop model with $c = 0\ 0002$, the * denotes the point $-1$

Fig 6 the Nyquist diagram of controller times model and controller times system is given, clearly indicating the model error near the critical point $-1$ In Fig 7 the Bode plot is presented of the resulting closed-loops of the controller implemented on the model and on the system It turns out that the controller destabilizes the system! Apparently the model identified in open-loop does not describe the relevant closed-loop properties of the system sufficiently well

We now want to identify a third-order model that gives an optimal closed-loop description of the system, using the identification scheme of Fig 2 We do this in an iteration of identification and feedback design as shown in Fig. 3 First we design a low-performance controller ($c = 0.0008$) for the model identified in open-loop Then

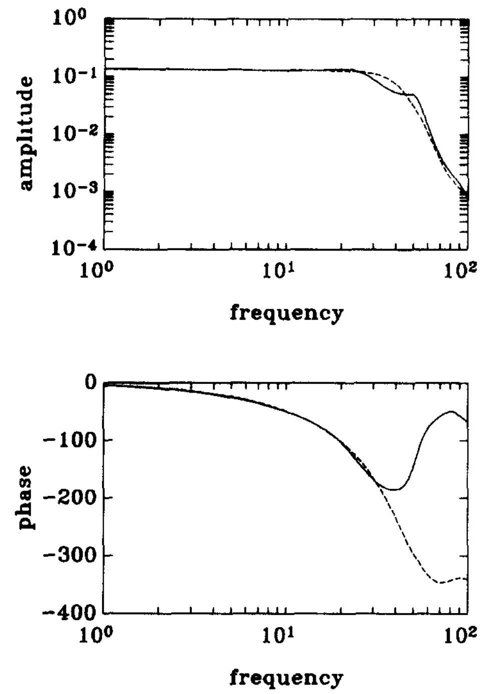


FIG 7 Bode diagram closed-loops $\hat{G}/(1+\hat{G}C)$ (dashed) and $G_0/(1+G_0C)$, where $\hat{G}$ is the model identified in open-loop and $C$ is the LQG controller designed for this open-loop model with $c = 0\ 0002$

closed-loop measurements are performed with a white noise external reference signal and again about 7% coloured additive output noise Using these measurements an output error model is identified applying a proper prefilter which is calculated using the designed controller and the open-loop identification result Next we design a new controller for the resulting model with increasing performance requirement ($c = 0\ 0004$) Then we conduct a new identification and we design a controller with $c = 0\ 0002$ We repeat this last step till there is no significant change in controller or model

Altogether four iterations were sufficient to reach the final result Bode plots of the resulting controllers are shown in Fig 8, which displays the increasing control action Figure 9 reveals that the resulting optimal model has a poor open-loop fit The closed-loops of the final controller implemented on the optimal model (designed loop) and on the system are depicted in Fig 10 The controller designed for the optimal model gives a satisfactory, stable performance for the system We remark that the optimal model has a bad open-loop behaviour, but it is nevertheless more suited for feedback design than the model identified in open-loop

## 6 DISCUSSION

In the example of the previous section it has been shown that for LQG controller design the

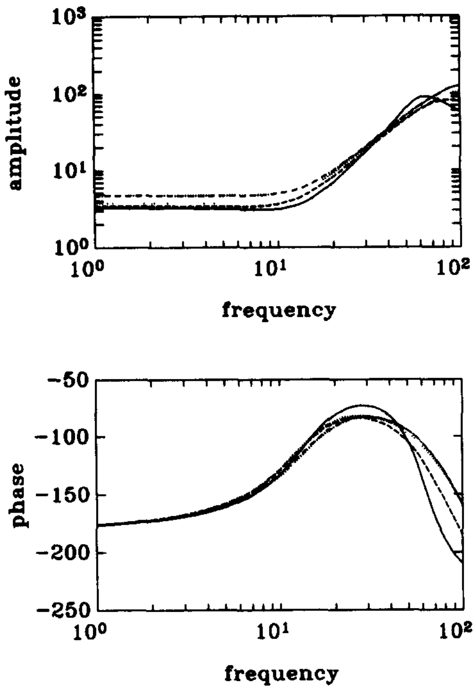


FIG 8 Bode diagram LQG controllers determined in an iterative way for $c = 0\ 0008$ (solid), $c = 0\ 004$ (dashed) and $c = 0\ 0002$ (dash–dotted, dotted)

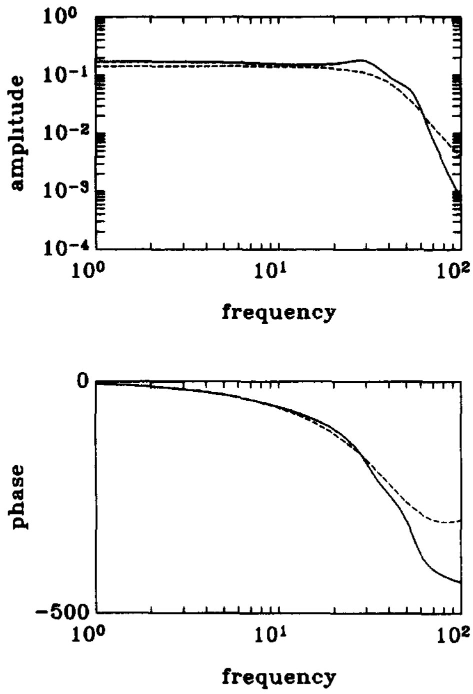


FIG 10 Bode diagram closed-loops $\hat{G}/(1 + \hat{G}C)$ (dashed) and $G_0/(1 + G_0C)$, where $\hat{G}$ is the optimal model and $C$ is the LQG controller designed for this optimal model with $c = 0\ 0002$

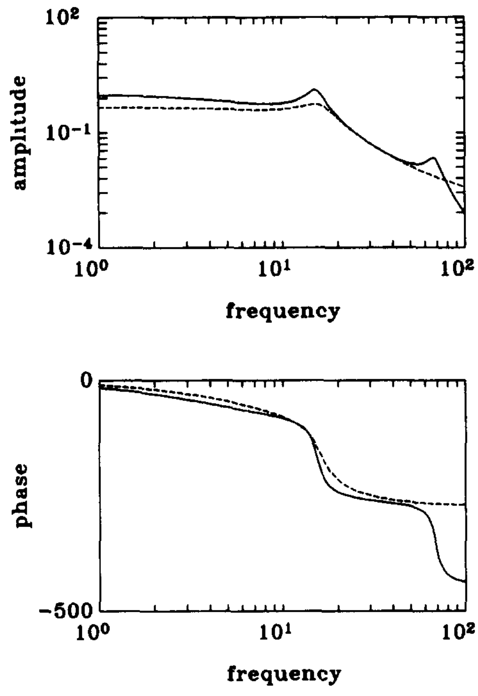


FIG 9 Bode diagram system (solid) and third-order optimal model (dashed)

optimal identification strategy of Fig. 2 in combination with the iterative scheme of Fig 3 yields a model that is superior to a model obtained by a simple open-loop identification This means that a combined iterative approach of identification and controller design can lead to results that are better than those obtained from open-loop considerations alone It is true that the open-loop identification was inappropriately weighted, but the point is that the optimal weighting is not known beforehand and needs to be determined iteratively The iterative aspect is essential, because a model is needed for controller design and knowledge of the controller is needed in order to identify a good model

The motivation for the applied iterative approach is, as already has been argued, that a model optimal for a certain controller will be close to optimality for a slightly different controller This explains why the procedure converged in the example of the previous section However it also means that the procedure might very well diverge if in each iteration the performance requirement is increased too much For in that case optimality is completely lost for the new controller Presently it is unknown under what conditions convergence can be guaranteed In the example of the previous section the controller update has simply been carried out by trial and error However also in the case that the performance requirements are increased slowly, there is a limit on the achievable performance This limit is determined by the required controller robustness The controller always has to be robust in the sense that it has to stabilize both the model and the system In the example in the previous section this means that the value of $c$ cannot be decreased arbitrarily, as at some moment the controller is not robust enough and it will destabilize the system

The results of this paper, based on an asymptotic bias analysis, can also be considered as a justification of other iterative schemes such as the one presented in Zang *et al* (1991) There a different closed-loop performance criterion is used, as the noise is treated differently, and moreover the controller design criterion is strongly connected to the identification result and always based on an LQ objective In our approach the controller design is basically completely free to choose, in the example in the previous section only a choice has been made for LQG design with a very simple choice for the weighting matrices

We now take a closer look at the criteria that are minimized in the identification and the controller design procedure In the LQG controller design procedure the quadratic criterion $J_{LQG}$ in (33) is minimized For a high performance controller ($c = 0\ 0002$ for example) the contribution of $\sum y_m^2(t)$ dominates this criterion function The external reference signal is white noise so that the LQG controller design procedure actually (approximately) minimizes

$$J_{LQG}(\hat{G}) = \|\hat{G}(I + C_2\hat{G})^{-1}\|_2 \qquad (34)$$

In the identification procedure the quadratic criterion $J_1$ in (11) is minimized As the external reference during identification is white noise, this means that the identification procedure minimizes

$$J_1 = \|G_0(I + C_2G_0)^{-1} - \hat{G}(I + C_2\hat{G})^{-1}\|_2, \qquad (35)$$

where we neglected the contribution of the noise Using the triangle inequality we obtain

$$J_{LQG}(G_0) = \|G_0(I + C_2G_0)^{-1}\|_2 \leq J_{LQG}(\hat{G}) + J_1, \qquad (36)$$

which means that the criterion value $J_{LQG}(G_0)$ is bounded Moreover, if the model is a good description of the system, $J_{LQG}(G_0)$ will be close to $J_{LQG}(\hat{G})$, which implies that in that case the controller $C_2$ is nearly optimal for the system This topic of matching criteria in identification and control design is further elaborated in Schrama (1992a, b)

Finally we remark that identification in closed-loop may be troublesome if there is noise present in the loop, as is practically always the case If the noise model is too simple to represent the noise, then the deterministic part of the model cannot be estimated consistently, see Soderstrom and Stoica (1989) This problem can be circumvented by 'decoupling' the deterministic and noise contribution for instance by the two-step procedure proposed in Van den Hof *et al* (1992)

## 7 CONCLUSIONS

Based on asymptotic results for prediction error idetification a scheme has been developed to identify a model that gives an optimal closed-loop description of the controlled system under investigation The procedure consists of data collection in operational conditions and after that the data are filtered properly The identified model can be used for feedback design This is carried out in an iterative procedure of identification and controller design In each iteration step a new model is identified, which is then used to design a new controller for increased performance requirements In an example the procedure has successfully been applied to design a high-performance LQG feedback controller The identification procedure turns out to be superior to straightforward open-loop identification This arises from the fact that the identification minimizes a criterion that is compatible with the LQG objective

*Acknowledgement*—The authors gratefully acknowledge Michel Gevers for the fruitful discussions on the topics addressed in this paper

## REFERENCES

Balas, G J and J C Doyle (1990) Identification of flexible structures for robust control *IEEE Contr Syst Mag*, 51–58

Bitmead, R R , M Gevers and V Wertz (1990) *Adaptive Optimal Control, The Thinking Man's GPC* Prentice-Hall, Englewood Cliffs, NJ

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

Hakvoort, R G (1990) Optimal experiment design for prediction error identification in view of feedback design In O H Bosgra and P M J Van den Hof (Eds), *Sel Topics in Ident , Mod and Contr* , Vol 2, pp 71–78 Delft Univ Press, The Netherlands

Hakvoort, R G , R J P Schrama and P M J Van den Hof (1992) Approximate identification in view of LQG feedback design *Proc Am Contr Conf*, pp 2824–2828 Chicago, IL

Janssen, P H M (1988) On model parameterization and model structure selection for identification of MIMO systems Ph D Thesis, Eindhoven Univ of Tech , The Netherlands

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

Maciejowski, J M (1989) *Multivariable Feedback Design* Addison-Wesley, NY

Rivera, D E , J F Pollard, L E Sterman and C E García (1990) An industrial perspective on control relevant identification *Proc Am Contr Conf*, pp 2406–2411 San Diego, CA

Schrama, R J P (1992a) Approximate identification and control design with application to a mechanical system Ph D Thesis, Delft Univ of Tech , The Netherlands

Schrama, R J P (1992b) Accurate models for control design the necessity of an iterative scheme *IEEE Trans Autom Contr* , **AC-37**, 991–994

Soderstrom, T and P Stoica (1989) *System Identification* Prentice-Hall, U K

Van den Hof, P M J , R J P Schrama and O H Bosgra (1992) An indirect method for transfer function

estimation from closed-loop data *Proc 31st Conf Dec and Contr*, pp 1702–1706 Tucson, AZ
Wahlberg, B and L Ljung (1986) Design variables for bias distribution in transfer function estimation *IEEE Trans Autom Contr*, **AC-31,** 134–144
Zang, Z, R R Bitmead and M Gevers (1991) Iterative model refinement and control robustness enhancement, Technical Report 91 137, Dept of Syst Eng, Res School of Phys Sc and Eng, Australian National University

## APPENDIX

### A *The MIMO case*

In this section we will briefly indicate how the optimal choice of design variables can be extended to the MIMO case The definitions of $\mathcal{S}$, $\mathcal{M}$ and $T_0(q)$ in the equations (1) till (3) remain unchanged The parameter vector $\hat{\theta}$ is calculated as

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N} \sum_{t=0}^{N-1} \varepsilon^T(t, \theta) W_1 \varepsilon(t, \theta), \qquad \text{(A 1)}$$

where $W_1$ is a symmetric weighting matrix

The asymptotic result (5) holds with straightforward modifications, see Janssen (1988, ch 2) There we also find the frequency domain interpretation

$$\theta^* = \arg\min_{\theta} \int_{-\pi}^{\pi} \mathrm{tr}\,\{\hat{H}_f^{-T}(e^{-i\omega}) W_1 \hat{H}_f^{-1}(e^{i\omega}) \tilde{T}(e^{i\omega}, \theta)\Phi(\omega)\tilde{T}^T(e^{-i\omega}, \theta)\}\, d\omega, \qquad \text{(A 2)}$$

with $\tilde{T}(q, \theta)$ and $\Phi(\omega)$ defined by the equations (7) and (8) respectively The design variables are defined by

$$\mathcal{D} = \{\Phi_u(\omega), \Phi_{ue}(\omega), W_1\}, \qquad \text{(A 3)}$$

Next we consider Fig 1 representing the operational conditions of the controller The output $\bar{y}$ is given by

$$\bar{y}(t) = [I + G_0(q)C_2(q)]^{-1} G_0(q) C_1(q) \bar{r}(t) + [I + G_0(q)C_2(q)]^{-1} H_0(q) \bar{e}(t) \qquad \text{(A 4)}$$

Analogously to the SISO case we define the performance criterion as

$$J_1(\mathcal{D}) = \int_{-\pi}^{\pi} \mathrm{tr}\,\{W_2(\tilde{G}_{CL}(e^{i\omega})\Phi_{\bar{r}}(\omega)\tilde{G}_{CL}^T(e^{-i\omega}) + \tilde{H}_{Cl}(e^{i\omega})\Phi_{\bar{e}}(\omega)\tilde{H}_{CL}^T(e^{-i\omega}))\}\, d\omega, \qquad \text{(A 5)}$$

where

$$\tilde{G}_{CL}(q) = [I + G(q, \hat{\theta})C_2(q)]^{-1} G(q, \hat{\theta}) C_1(q) - [I + G_0(q)C_2(q)]^{-1} G_0(q) C_1(q), \qquad \text{(A 6)}$$

$$\tilde{H}_{Cl}(q) = [I + G(q, \hat{\theta})C_2(q)]^{-1} \hat{H}_f(q) - [I + G_0(q)C_2(q)]^{-1} H_0(q) \qquad \text{(A 7)}$$

$W_2$ is a symmetric weighting matrix, that should reflect the relative importance of each component in the performance criterion The following proposition gives an alternative expression for this performance criterion

*Proposition* A 1 The performance criterion $J_1(\mathcal{D})$ satisfies

$$J_1(\mathcal{D}) = \int_{-\pi}^{\pi} \mathrm{tr}\,\{[I + G(e^{-i\omega}, \hat{\theta}(\mathcal{D}))C_2(e^{-i\omega})]^{-T} \times W_2[I + G(e^{i\omega}, \hat{\theta}(\mathcal{D}))C_2(e^{i\omega})]^{-1} \times \tilde{T}(e^{i\omega}, \hat{\theta}(\mathcal{D}))\bar{\Phi}(\omega)\tilde{T}^T(e^{-i\omega}, \hat{\theta}(\mathcal{D}))\}\, d\omega, \qquad \text{(A 8)}$$

where the signal $\bar{u}(t)$ is according to Fig 1 given by

$$\bar{u}(t) = [I + C_2(q)G_0(q)]^{-1} C_1(q) \bar{r}(t) - [I + C_2(q)G_0(q)]^{-1} C_2(q) H_0(q) \bar{e}(t), \qquad \text{(A 9)}$$

and the matrix $\bar{\Phi}(\omega)$ is defined by equation (14)

*Proof* See Appendix B □

Now we define the general scalar performance criterion $J_G$ for the MIMO case as

$$J_G(\mathcal{D}) = \int_{-\pi}^{\pi} \mathrm{tr}\,\{\bar{\Gamma}(\omega)\tilde{T}(e^{i\omega}, \hat{\theta}(\mathcal{D})) \times \Gamma(\omega)\tilde{T}^T(e^{-i\omega}, \hat{\theta}(\mathcal{D}))\, d\omega, \qquad \text{(A 10)}$$

where $\Gamma(\omega)$ and $\bar{\Gamma}(\omega)$ are Hermitian weighting matrices We partition $\Gamma(\omega)$ into four blocks matching the block structure in $\tilde{T}(q, \theta)$, see equation (3),

$$\Gamma(\omega) = \begin{bmatrix} \Gamma_{11}(\omega) & \Gamma_{12}(\omega) \\ \Gamma_{21}(\omega) & \Gamma_{22}(\omega) \end{bmatrix} \qquad \text{(A 11)}$$

We now formulate the MIMO analogon of Theorem 4 1

*Theorem* A 2 The optimal choice of design variables (18) is given by

$$\hat{H}_f^{-T}(e^{-i\omega}) W_{1\ \mathrm{opt}} \hat{H}_f^{-1}(e^{i\omega}) = c_1\bar{\Gamma}(\omega), \quad \Phi_{u\ \mathrm{opt}}(\omega) = c_2\Gamma_{11}(\omega), \quad \Phi_{ue\ \mathrm{opt}}(\omega) = c_2\Gamma_{12}(\omega), \qquad \text{(A 12)}$$

where $c_1$ and $c_2$ are arbitrary positive constants

*Proof* Follows from the application of Lemma 13 1 in Ljung (1987) to the equations (A 2) and (A 10) As both $\Gamma(\omega)$ and $\Phi(\omega)$ are Hermitian, a third constraint, namely on $\Phi_{eu}(\omega)$ is superfluous Also no constraint on $\Phi_e$ in relation to $\Gamma_{22}(\omega)$ is required because its contribution to the identification criterion (38) is independent of $\theta$ and does therefore not influence the minimizing value $\theta^*$ Finally the constant scaling factors $c_1$ and $c_2$ do not affect the optimality property □

Analogous to the SISO case an auxiliary quadratic performance criterion $J_2$ can be introduced, yielding the values for $\Gamma$ and $\bar{\Gamma}$

$$\Gamma(\omega) = \bar{\Phi}(\omega), \quad \bar{\Gamma}(\omega) = [I + \hat{G}_f(e^{-i\omega})C_2(e^{-i\omega})]^{-T} \times W_2[I + \hat{G}_f(e^{i\omega})C_2(e^{i\omega})]^{-1} \qquad \text{(A 13)}$$

The discrepancies $\delta_1(\omega)$ and $\delta_2(\omega)$ have an obvious MIMO analogon, that will not be given here explicitly Finally we are able to formulate the optimal choice of design variables in a Theorem

*Theorem* A 3 If a fixed noise model $\hat{H}_f(q) = I + \hat{G}_f(q)C_2(q)$ is used in the prediction error identification and if the number of samples tends to infinity the choice of design variables

$$\mathcal{D}_{2\ \mathrm{opt}} = \begin{cases} \Phi_{u\ \mathrm{opt}}(\omega) = \Phi_{\bar{u}}(\omega) \\ \Phi_{ue\ \mathrm{opt}}(\omega) = \Phi_{\bar{u}\bar{e}}(\omega), \\ W_{1\ \mathrm{opt}} = W_2 \end{cases} \qquad \text{(A 14)}$$

is arbitrarily close to the optimal solution $\mathcal{D}_{1\ \mathrm{opt}}$, provided $\delta_1(\omega)$ and $\delta_2(\omega)$ are sufficiently small, i e

$$\lim_{\delta_1\,\delta_2 \to 0} J_1(\mathcal{D}_{2\ \mathrm{opt}}) - J_1(\mathcal{D}_{1\ \mathrm{opt}}) = 0 \qquad \text{(A 15)}$$

*Proof* Follows from Theorem A 2, following the steps of the proof of Theorem 4 2 □

This means that again data collection should take place under operational conditions in order that a model is identified that gives an optimal closed-loop description of the system The fixed nosie model $\hat{H}_f(q) = [I + \hat{G}_f(q)C_2(q)]$ can be realized by prefiltering the prediction error $\varepsilon(t, \theta)$ with the filter $L(q) = [I + \hat{G}_f(q)C_2(q)]^{-1}$ in combination with an output error identification scheme Note that in the MIMO case this is not equivalent to filtering the input/output data with this filter

### B *Proof of Propositions* 3 1 *and* A 1

In this appendix we give a proof for Proposition A 1 As Proposition 3 1 is a special case of Proposition A 1 this is at the same time a proof for Proposition 3 1 For ease of notation we will not always explicitly mention dependency of a quantity on $q$ or $t$ and we will use the short-hand notation $\hat{G}$ and $\hat{H}$ for $G(q, \hat{\theta})$ and $\hat{H}_f(q)$ respectively

*Proof* Define the auxiliary signal $\bar{y}(t)$ as

$$\begin{aligned}\bar{y}(t) = &([I + \hat{G}(q)C_2(q)]^{-1}\hat{G}(q)C_1(q)\\ &- [I + G_0(q)C_2(q)]^{-1}G_0(q)C_1(q))\bar{r}(t)\\ &+ ([1 + \hat{G}(q)C_2(q)]^{-1}\hat{H}(q)\\ &- [1 + G_0(q)C_2(q)]^{-1}H_0(q))\bar{e}(t)\end{aligned} \quad \text{(A 16)}$$

Then as the signals $\bar{r}$ and $\bar{e}$ are uncorrelated the performance criterion $J_1(\mathscr{D})$ is obviously equal to

$$J_1(\mathscr{D}) = \int_{-\pi}^{\pi} \mathrm{tr}\,\{W_2\Phi_{\bar{y}}(\omega)\}\,\mathrm{d}\omega \quad \text{(A 17)}$$

Now the auxiliary signal $\bar{y}(t)$ can be written as

$$\begin{aligned}\bar{y}(t) &= ([I + \hat{G}C_2]^{-1}\hat{G} - [I + G_0C_2]^{-1}G_0)C_1\bar{r}\\ &\quad + ([I + \hat{G}C_2]^{-1}\hat{H} - [I + G_0C_2]^{-1}H_0)\bar{e}\\ &= [I + \hat{G}C_2]^{-1}\{(\hat{G} - [I + \hat{G}C_2]G_0[I + C_2G_0]^{-1})C_1\bar{r}\\ &\quad + (\hat{H} - [I + \hat{G}C_2][I + G_0C_2]^{-1}H_0)\bar{e}\}\\ &= [I + \hat{G}C_2]^{-1}\{(\hat{G}[I + C_2G_0] - [I + \hat{G}C_2]G_0)\\ &\quad \times [I + C_2G_0]^{-1}C_1\bar{r} + (\hat{H} - H_0)e\\ &\quad + ([I + G_0C_2] - [I + \hat{G}C_2])[I + G_0C_2]^{-1}H_0\bar{e}\}\\ &= [I + \hat{G}C_2]^{-1}\{(\hat{G} - G_0)[I + C_2G_0]^{-1}C_1\bar{r}\\ &\quad + (\hat{H} - H_0)\bar{e} - (\hat{G} - G_0)C_2[I + G_0C_2]^{-1}H_0\bar{e}\}\\ &= [I + \hat{G}C_2]^{-1}\{(\hat{G} - G_0)([I + C_2G_0]^{-1}C_1\bar{r}\\ &\quad - [I + C_2G_0]^{-1}C_2H_0\bar{e}) + (\hat{H} - H_0)\bar{e}\}\\ &= [I + \hat{G}C_2]^{-1}[(\hat{G} - G_0)\bar{u} + (\hat{H} - H_0)\bar{e}]\\ &= [I + G(q, \hat{\theta})C_2(q)]^{-1}\bar{T}(q, \hat{\theta})\begin{bmatrix}\bar{u}(t)\\ \bar{e}(t)\end{bmatrix},\end{aligned} \quad \text{(A 18)}$$

with the signal $\bar{u}(t)$ given by (A 9) Consequently (A 17) and (A 18) yield the desired result □