

# Least costly identification experiment for control<sup>☆</sup>

X. Bombois<sup>a,\*</sup>, G. Scorletti<sup>b</sup>, M. Gevers<sup>c</sup>, P.M.J. Van den Hof<sup>a</sup>, R. Hildebrand<sup>d</sup>

<sup>a</sup>*Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands*

<sup>b</sup>*GREYC-Equipe AUTO, Caen, France*

<sup>c</sup>*CESAME, Louvain-la-Neuve, Belgium*

<sup>d</sup>*IMAG, Grenoble, France*

Received 22 April 2004; received in revised form 16 May 2006; accepted 22 May 2006

## Abstract

All approaches to optimal experiment design for control have so far focused on deriving an input signal (or input signal spectrum) that minimizes some control-oriented measure of plant/model mismatch between the nominal closed-loop system and the actual closed-loop system, typically under a constraint on the total input power. In practical terms, this amounts to finding the (constrained) input signal that minimizes a measure of a control-oriented model uncertainty set. Here we address the experiment design problem from a “dual” point of view and in a closed-loop setting: given a maximum allowable control-oriented model uncertainty measure compatible with our robust control specifications, what is the cheapest identification experiment that will give us an uncertainty set that is within the required bounds? The identification cost can be measured by either the experiment time, the performance degradation during experimentation due to the added excitation signal, or a combination of both. Our results are presented for the situation where the control objective is disturbance rejection only.

© 2006 Elsevier Ltd. All rights reserved.

*Keywords:* Identification for control; Experiment design

## 1. Introduction

The problem addressed and solved in this paper is a novel formulation of an optimal experiment design problem, in the context of “identification for robust control”. It is novel essentially in that it takes the dual viewpoint to the classical way of posing optimal experiment design problems. To state this in a nutshell, the classical way is to seek the optimal input signal (or the optimal input signal spectrum) that minimizes some control-oriented measure of the quality of the identified model, subject to constraints on the input signal power (and/or on the output signal power). Representative examples of such approach can be found in Cooley and Lee (2001), Forssell and Ljung (2000), Gevers and Ljung (1986), Hildebrand and Gevers (2003), Hjalmarsson, Gevers, and De Bruyne (1996),

Lindqvist (2001), Ljung (1999). From a practical point of view (Rivera, Lee, Braun, & Mittelmann, 2003), such approach is not always the most sensible choice: one should not spend more effort on the identification than what is needed to achieve the required robust control performance.

Thus, here we consider a “dual approach” to this optimal experiment design problem. Given some prior bounds on the uncertainty around an identified model that is compatible with the robust control performance specifications when such model is to be used for robust control design, we seek to minimize the identification cost such that the uncertainty set around the identified model is (just) within these prior bounds. Initial steps in the direction of minimizing the input effort given a required quality measure have been taken in Bombois, Scorletti, Van den Hof, and Gevers (2004), where open-loop identification was considered with a constraint on the total input power, together with an  $H_\infty$  robust control design criterion. The idea of minimizing the total input power for open-loop identification, subject to achieving a required quality measure on the estimated model, has been further developed in Jansson and Hjalmarsson (2004), where a variety of quality measures has been

<sup>☆</sup> This paper was presented at the IFAC World Congress, Prague, July 2005. This paper was recommended for publication in revised form by Associate Editor Antonio Vicino under the direction of Editor Torsten Söderström.

\* Corresponding author.

E-mail address: [x.j.a.bombois@dsc.tudelft.nl](mailto:x.j.a.bombois@dsc.tudelft.nl) (X. Bombois).

considered. Here we solve this optimal experiment design problem in a closed-loop setup.

The problem formulation addressed in this paper is as follows. We consider that an unknown “true” linear time-invariant system  $G_0$  is controlled with an existing controller  $C_{id}$  that is not fully satisfactory, and which one wants to replace by a new controller. The control objective is regulation only, i.e. the normal operation of the closed-loop system is with a constant reference signal, which, for simplicity, we assume to be zero. For the purpose of designing a new and robust controller, an identification step is performed in closed loop with the existing controller  $C_{id}$ , leading to a model  $\hat{G}$  of the unknown system  $G_0$  and a parametric uncertainty region  $\mathcal{D}$  centered around  $\hat{G}$  and containing the true system  $G_0$  with some pre-chosen probability, say  $\beta$  (Bombois, Gevers, Scorletti, & Anderson, 2001; Ljung, 1999).

A new controller  $\hat{C}(\hat{G})$  will be designed from this identified model  $\hat{G}$  using a pre-defined control design method (e.g. an  $H_\infty$  control design method with fixed weights). The requirement imposed on the new controller is that it achieves a sufficient  $H_\infty$  performance with the true system  $G_0$ . However,  $G_0$  is unknown and therefore  $\hat{C}(\hat{G})$  is designed in such a way that it achieves this sufficient level of  $H_\infty$  performance for a set of systems  $\mathcal{D}_{adm}$  around the estimated  $\hat{G}$  (the set  $\mathcal{D}_{adm}$  will be formally defined at the end of Section 3). Returning now to the identification part of the exercise, the whole idea of our optimal experiment design program is to perform a closed-loop identification experiment with the smallest possible cost so that the uncertainty set  $\mathcal{D}$  around the estimated  $\hat{G}$  lies within the admissible uncertainty set  $\mathcal{D}_{adm}$  imposed by the control performance specifications. This implies, in particular, that the true system  $G_0$ , which lies within  $\mathcal{D}$  with probability  $\beta$ , will also lie in  $\mathcal{D}_{adm}$  with probability at least equal to  $\beta$ , and will therefore achieve the required  $H_\infty$  performance level with  $\hat{C}(\hat{G})$  with the same probability. The size of the identified  $\mathcal{D}$  is a function of the covariance matrix  $P_{\hat{\theta}}$  of the identified parameter vector  $\hat{\theta}_N$  and, consequently, a function of length  $N$  of the identification experiment and of the power spectrum  $\Phi_r(\omega)$  of the excitation signal  $r(t)$  used for the identification (see Fig. 1).

It remains to specify what is meant by “identification with the smallest possible cost”. In the present context of a regulation problem, any external excitation signal  $r$  that is added for the purpose of identifying the system produces added perturbations  $y_r$  and  $u_r$  to the normal operating signals  $y$  and  $u$ . During the experiment time, these perturbations increase the variance of  $y$  and  $u$  beyond their normal value, and consequently cause

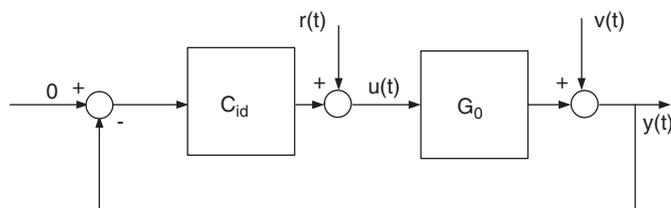


Fig. 1. Closed loop  $[C_{id} G_0]$  during an identification experiment with  $r(t)$  as excitation signal.

a degradation of the production quality. It is this degradation that we shall seek to minimize while achieving the required precision ( $\mathcal{D} \subseteq \mathcal{D}_{adm}$ ).

All results in this paper are derived under the assumption that the system is in the model set, which implies that we consider variance errors only. However, unlike (Forsell & Ljung, 2000; Gevers & Ljung, 1986; Hjalmarsson et al., 1996), our results are not based on the asymptotic (in model order) transfer function variance expressions (Ljung, 1999, p. 295) but on the more accurate parameter covariance expressions for finite model orders, that were also used in Hildebrand and Gevers (2003).

Our first contribution, based on a classical result (Söderström & Stoica, 1989, p. 392), is to show that, when the controller  $C_{id}$  is sufficiently complex, the model accuracy required for any robust performance specifications can be achieved at no cost at all, provided that one can wait long enough, i.e. no external excitation signal  $r$  is required provided the data length  $N$  can be made sufficiently large. The minimal data length required to achieve the desired model accuracy can be determined via an LMI optimization. Such a strategy is not always practically feasible. Indeed, the existing controller  $C_{id}$  may e.g. be of insufficient order or one might have to wait an unduly long time to get the required accuracy. Thus, we shall present a number of optimal experiment design scenarios, where the optimal design will be a trade-off between data length and a sufficient amount of excitation power  $\Phi_r(\omega)$ . This is done by providing an LMI-based solution to the following two experiment design problems:

- For a given excitation spectrum  $\Phi_r(\omega)$ , determine the minimal experiment time  $N$  that is necessary to meet the robust performance constraint, i.e.  $\mathcal{D} \subseteq \mathcal{D}_{adm}$ .
- For a given experiment time  $N$ , determine the power spectrum  $\Phi_r(\omega)$  which induces the smallest perturbation with respect to the normal operating conditions of the control loop  $[C_{id} G_0]$ , and which achieves the required robust performance constraints, i.e.  $\mathcal{D} \subseteq \mathcal{D}_{adm}$ .

Our tools for cheap identification for robust control are developed for direct closed-loop prediction error identification (Söderström & Stoica, 1989, p. 389). The open-loop case was treated in Bombois, Scorletti, Gevers, Hildebrand, and Van den Hof (2004).

## 2. Prediction error identification aspects

We consider the identification of a stable linear time-invariant single input single output system with a model structure  $\mathcal{M} = \{G(z, \theta), H(z, \theta)\}$ ,  $\theta \in \mathbf{R}^k$ , that is able to represent the true system. Thus, the finite-dimensional true system is given by

$$\mathcal{S} : y(t) = \underbrace{G_0(z)}_{G(z, \theta_0)} u(t) + \underbrace{v(t)}_{H(z, \theta_0) e(t)} \quad (1)$$

for some unknown parameter vector  $\theta_0 \in \mathbf{R}^k$ . In (1),  $e(t)$  is a white noise with variance  $\sigma_e^2$  and  $G(z, \theta_0)$ ,  $H(z, \theta_0)$  are stable

discrete-time transfer functions. The transfer function  $H(z, \theta_0)$  is furthermore assumed monic and minimum-phase. We shall also assume throughout the paper that the model structure is globally identifiable at  $\theta_0$ . This means that  $\theta_0$  is the only value of  $\theta$  for which  $G(z, \theta)$  and  $H(z, \theta)$  represent the true system (see Theorem 4.2 of Ljung, 1999).

This true system is operated in closed loop with an initial (but unsatisfactory) controller  $C_{id} : u(t) = r(t) - C_{id}(z)y(t)$ . The signal  $r(t)$  is zero in normal operation but can be used to excite the system for a closed-loop identification experiment (see Fig. 1). With  $S_{id} \triangleq 1/(1 + C_{id}G_0)$ , the closed-loop system can be written as

$$y(t) = S_{id}v(t) + \overbrace{G_0 S_{id}r(t)}^{y_r(t)}, \tag{2}$$

$$u(t) = -C_{id}S_{id}v(t) + \underbrace{S_{id}r(t)}_{u_r(t)}. \tag{3}$$

Suppose it is desired to identify a model  $\hat{G}(z)=G(z, \hat{\theta}_N)$ ,  $\hat{H}(z)=H(z, \hat{\theta}_N)$  of the true system by applying an external signal  $r$  to the actual closed-loop system with the controller  $C_{id}$  in the loop and by collecting  $N$  input–output data, using a direct prediction error (PE) identification method. We shall assume throughout this paper that the signal  $r$  is taken as a part of length  $N$  of a realization of a quasi-stationary signal with power spectrum  $\Phi_r(\omega)$ , yielding measured signals  $\{u(t), y(t), t = 1, \dots, N\}$ .  $\Phi_r(\omega)$  is defined as the Fourier transform of  $R_r(\tau) = \lim_{N \rightarrow \infty} (1/N) \sum_{t=1}^N E r(t)r(t - \tau)$ . The estimated parameter vector  $\hat{\theta}_N$  is then defined by:  $\hat{\theta}_N \triangleq \arg \min_{\theta} (1/N) \sum_{t=1}^N \varepsilon^2(t, \theta)$  with  $\varepsilon(t, \theta) \triangleq H(z, \theta)^{-1}(y(t) - G(z, \theta)u(t))$ . Note that  $\varepsilon(t, \theta)$  depends on the chosen signal  $r(t)$  via (2)–(3).

In this paper, as is usual in PE identification, the experimental conditions will be chosen/designed such that the closed-loop identification experiment is “informative enough” (Ljung, 1999) i.e. such that the cost function

$$\bar{V}(\theta) = \bar{E} \varepsilon^2(t, \theta) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E \varepsilon^2(t, \theta) \tag{4}$$

has a unique minimum (to which  $\hat{\theta}_N$  tends w.p. 1 when  $N \rightarrow \infty$ ). By our standing assumption that the system is in the model set and that the model structure is globally identifiable at  $\theta_0$ , this unique minimum is  $\theta_0$ . The informative character of the experiment depends on the complexity of the controller  $C_{id}$  (see Section 5) and can always be guaranteed by proper choice of  $\Phi_r(\omega)$ . Indeed, a signal  $r(t)$  generated by filtering a white noise with a finite-dimension filter will always lead to an informative enough experiment. See e.g. (Ljung, 1999, p. 427) for details. Under these conditions, we have the following important result.

**Lemma 1.** *Consider a closed-loop identification experiment, as described above, that is informative enough. Then the identified parameter vector  $\hat{\theta}_N$  is asymptotically normally distributed around the true parameter vector  $\theta_0$ , i.e.  $\hat{\theta}_N \sim \mathcal{N}(\theta_0, P_\theta)$  with  $P_\theta$  a strictly positive definite matrix given by Ljung (1999,*

Chapter 9)

$$P_\theta = \frac{\sigma_e^2}{N} (\bar{E}(\psi(t, \theta_0)\psi(t, \theta_0)^T))^{-1} \quad \text{with} \\ \psi(t, \theta) = -\frac{\partial \varepsilon(t, \theta)}{\partial \theta}. \tag{5}$$

In this paper, the design variables with respect to which we shall optimize the identification experiment are the data length  $N$ , and the external signal  $r$  via its spectrum  $\Phi_r(\omega)$ . The following expression, which is easily deduced from (3) and (5), shows precisely how the covariance matrix  $P_\theta$  in (5) depends on these design variables:

$$P_\theta^{-1} = N \left( \overbrace{\frac{1}{\sigma_e^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} F_r(e^{j\omega}, \theta_0) F_r(e^{j\omega}, \theta_0)^* \Phi_r(\omega) d\omega}^{\mathcal{P}_r^{-1}(\Phi_r(\omega), \theta_0, \sigma_e^2)} \right) \\ + N \left( \underbrace{\frac{1}{2\pi} \int_{-\pi}^{\pi} F_e(e^{j\omega}, \theta_0) F_e(e^{j\omega}, \theta_0)^* d\omega}_{\mathcal{P}_v^{-1}(\theta_0)} \right). \tag{6}$$

Here,  $F_r(z, \theta_0) = S_{id}(A_G(z, \theta_0)/H(z, \theta_0))$ ,  $F_e(z, \theta_0) = A_H(z, \theta_0)/H(z, \theta_0) - C_{id}S_{id}A_G(z, \theta_0)$ ,  $A_G(z, \theta) = \partial G(z, \theta)/\partial \theta$  and  $A_H(z, \theta) = \partial H(z, \theta)/\partial \theta$ .

Using Lemma 1, it is possible to define an uncertainty region  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  around the identified model and containing the unknown true system  $G(z, \theta_0)$  at any desired probability level  $\beta$  (Bombois et al., 2001):

$$\mathcal{D}(\hat{\theta}_N, P_\theta) = \left\{ G(z, \theta) = \frac{Z_N(z)\theta}{1 + Z_D(z)\theta} \mid \theta \in U, \right. \\ \left. U = \{ \theta \mid (\theta - \hat{\theta}_N)^T P_\theta^{-1} (\theta - \hat{\theta}_N) < \chi \} \right\}, \tag{7}$$

where  $\chi$  is a real constant dependent on the chosen probability level  $\beta$  and  $Z_N, Z_D$  are row vectors containing delays and zeros. These vectors introduce a general parametrization of  $G(z, \theta)$  which will be important in the sequel. Note that the size of the uncertainty region  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  is a function of the covariance matrix  $P_\theta$  and thus, by (6), a function of the design variables  $N$  and  $\Phi_r(\omega)$ .

**3. Control design objectives and control design method**

As stated before, our aim is to replace the “unsatisfactory” controller  $C_{id}(z)$  by a “satisfactory” controller  $\hat{C}(z)$ . In order to define what we mean by satisfactory controller, we adopt the following performance measure for a stable loop  $[C \ G]$ :

$$J(G, C, W_l, W_r) = \sup_{\omega} \bar{J}(\omega, G, C, W_l, W_r) \tag{8}$$

with

$$\bar{J}(\omega, G, C, W_l, W_r) = \bar{\sigma}(W_l(e^{j\omega})F(G(e^{j\omega}), C(e^{j\omega}))W_r(e^{j\omega})), \tag{9}$$

$$F(G, C) \triangleq \begin{pmatrix} \frac{GC}{1 + GC} & \frac{G}{1 + GC} \\ \frac{C}{1 + GC} & \frac{1}{1 + GC} \end{pmatrix},$$

where  $\bar{\sigma}(A)$  denotes the largest singular value of  $A$  and  $W_l(z)$ ,  $W_r(z)$  are chosen diagonal performance filters. The performance measure (8) is quite general:  $J(G, C, W_l, W_r) \leq 1$  ensures that the four entries of  $W_l(z)F(G, C)W_r(z)$  have an  $H_\infty$  norm smaller than one. Simpler  $H_\infty$  criteria can be chosen as special cases; e.g., for  $W_l(z) = \text{diag}(0, W(z))$  and  $W_r = \text{diag}(0, 1)$ ,  $J(G, C, W_l, W_r) \leq 1$  corresponds to  $\|W/(1+CG)\|_\infty \leq 1$ . The performance filters  $W_l(z)$  and  $W_r(z)$  are chosen in such a way that they reflect the performance specifications that we want to achieve with the true system. Thus, a controller  $C$  will be deemed *satisfactory* for the system  $G_0$  if  $[C \ G_0]$  is stable and if  $J(G_0, C, W_l, W_r) \leq 1$ .

As mentioned in the Introduction, we want to design the new controller  $\hat{C}(z)$  using an identified model  $\hat{G}(z) = G(z, \hat{\theta}_N)$  of  $G_0(z)$ . For this purpose, we will use a pre-selected nominal control design method:

**Assumption 1.** We have pre-selected a fixed nominal control design method which will map the identified model  $\hat{G}$  to one controller  $\hat{C}$ . This control design method has been chosen in such a way that the controller  $\hat{C}$  stabilizes  $\hat{G}$  and achieves with this model a nominal performance level:

$$J(\hat{G}, \hat{C}, W_l, W_r) \leq \gamma < 1, \quad (10)$$

where  $\gamma$  is a fixed scalar.

One possible way to choose a control design method satisfying Assumption 1 is to choose for  $\hat{C}$  the central controller of the four-block  $H_\infty$  control design method with performance objective (10). When this controller  $\hat{C}$  is applied to the true system  $G_0$ , the achieved performance will generically be poorer than the designed performance. However, by choosing the design criterion (10) with  $\gamma < 1$ , we ensure that there is a whole set of systems  $G$  around  $\hat{G}$  that are also stabilized by  $\hat{C}$  and that achieve  $J(G, \hat{C}, W_l, W_r) \leq 1$ . In the sequel,  $\mathcal{D}_{\text{adm}}(\hat{\theta}_N)$  denotes the set containing all systems  $G(z)$  having these properties. This set  $\mathcal{D}_{\text{adm}}(\hat{\theta}_N)$  of course contains  $G(z, \hat{\theta}_N)$ .

#### 4. Identification for control at the lowest cost

Let us now proceed to our experiment design problem. As stated earlier, besides the standard requirement of an experiment which is informative enough, our objective is to determine the experimental conditions (i.e.  $N$  and  $\Phi_r(\omega)$ ) of the identification experiment on the loop  $[C_{\text{id}} \ G_0]$  in such a way that the model  $\hat{G}$ , identified through this experiment, delivers a controller  $\hat{C}$  which stabilizes the unknown  $G_0$  and achieves  $J(G_0, \hat{C}, W_l, W_r) \leq 1$ . Since  $G_0$  is unknown but lies (with probability  $\beta$ ) in the uncertainty region  $\mathcal{D}(\hat{\theta}_N, P_\theta)$ , this performance constraint will be replaced by the following checkable constraint.

**Constraint 1.** *The experimental conditions (i.e.  $N$  and  $\Phi_r(\omega)$ ) of the identification experiment on the loop  $[C_{\text{id}} \ G_0]$  (see Section 2) must be such that the identified model  $\hat{G} = G(z, \hat{\theta}_N)$  and the identified uncertainty region  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  have the*

*property that  $J(G, \hat{C}, W_l, W_r) \leq 1$  for all  $G(z) \in \mathcal{D}(\hat{\theta}_N, P_\theta)$ . In the previous expression,  $\hat{C}$  is the controller designed from  $\hat{G}$  using the control design method presented in Assumption 1.*

Since  $J(G, \hat{C}, W_l, W_r) \leq 1$  for all  $G(z)$  in the set  $\mathcal{D}_{\text{adm}}(\hat{\theta}_N)$  defined in the last paragraph of Section 3, Constraint 1 is ensured when  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  is a subset of this set  $\mathcal{D}_{\text{adm}}(\hat{\theta}_N)$ . Since  $P_\theta > 0$ ,  $\mathcal{D}(\hat{\theta}_N, P_\theta) \subseteq \mathcal{D}_{\text{adm}}(\hat{\theta}_N)$  is always achievable if  $N$  and/or  $\Phi_r(\omega)$  are chosen large enough: see (6). Moreover, we have the following trade-offs: the larger  $N$  is chosen, the *smaller*  $\Phi_r(\omega)$  can be while still verifying Constraint 1. Conversely, the *larger*  $\Phi_r(\omega)$  is chosen, the smaller  $N$  can be while still verifying this constraint.

Based on the reasoning above, it is clear that many possible choices of experimental conditions allow one to fulfill Constraint 1. Among these, we seek to determine those inducing the smallest possible economical cost. We now define this cost. In normal operation the signals  $u(t)$  (control signal) and  $y(t)$  (the controlled variable which is in many cases the product) are given by:  $y(t) = S_{\text{id}}v(t)$  and  $u(t) = -C_{\text{id}}S_{\text{id}}v(t)$ . By applying an external signal  $r(t)$  to the loop during the identification, we introduce *disturbances*  $y_r(t)$  and  $u_r(t)$  on top of the normal operation signals: see (2)–(3). Those disturbances induce a loss of production quality and thus entail an economical cost. Consequently, as far as the cost of the identification is concerned, the ideal closed-loop identification experiment would be, in many cases, one in which the normal operation signals  $u(t)$  and  $y(t)$  are used for a certain length  $N$  without any external excitation, i.e. with  $r(t) = 0$ . We show in Section 5 that such costless identification experiment can, in certain circumstances, fulfill Constraint 1. We also show how to compute the minimum number  $N_{\text{min}}$  of measurements that are necessary to reach this objective.

In the cases where Constraint 1 cannot be achieved with  $r(t) = 0$ , the application of a nonzero external signal  $r(t)$  for a certain amount of time is unavoidable, but we show how  $N$  and  $\Phi_r(\omega)$  can be chosen in order to achieve Constraint 1 with minimal economical cost. This economical cost will generally be a function of either the experiment time  $N$ , the power of the perturbations  $y_r$  and  $u_r$ , or a combination of both. In the sequel, we will therefore distinguish three different situations and determine for each of them how we can optimally choose the experimental conditions.

**Situation 1.** The cost of the identification is mainly determined by the power of the perturbations  $y_r(t)$  and  $u_r(t)$ . Based on the trade-off presented below Constraint 1, the experiment time  $N$  is in this case chosen as large as we are allowed to. For such fixed  $N$ , the optimal power spectrum  $\Phi_r(\omega)$  can subsequently be determined by minimizing the following cost function  $\mathcal{J}_r$ :

$$\begin{aligned} \mathcal{J}_r &= \alpha_y \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{y_r}(\omega) d\omega \right) + \alpha_u \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_r}(\omega) d\omega \right) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\alpha_y |G_0(e^{j\omega}) S_{\text{id}}(e^{j\omega})|^2 \\ &\quad + \alpha_u |S_{\text{id}}(e^{j\omega})|^2) \Phi_r(\omega) d\omega \end{aligned} \quad (11)$$

while guaranteeing Constraint 1. Here  $\alpha_y$  and  $\alpha_u$  are arbitrarily chosen scalars and  $\Phi_{y_r}(\omega)$  and  $\Phi_{u_r}(\omega)$  are the power spectra of the disturbance signals  $y_r(t)$  and  $u_r(t)$ , respectively.

**Situation 2.** Situation 2 is the converse situation: the cost of the identification is mainly determined by the duration of the identification. Based on the trade-off presented below Constraint 1, the power spectrum  $\Phi_r(\omega)$  of the to-be-applied signal  $r(t)$  is in this case chosen at each frequency as large as the constraints on the actuators allow. For such fixed  $\Phi_r(\omega)$ , the optimal experiment time is the smallest experiment time  $N_{\min}$  satisfying Constraint 1. Note that, in this situation, the identification with  $r = 0$  is generally not the optimal experiment.

**Situation 3.** Situation 3 is where  $N$  and  $\Phi_r(\omega)$  are both important in the cost of the identification. In this situation, we can determine the optimal spectrum  $\Phi_r(\omega)$  as in Situation 2 for different values of the length  $N$ . Since, for increasing values of  $N$ , the optimal cost function  $\mathcal{J}_r$  decreases, such approach allows one to find the “optimal” combination for the duration of the identification and the induced disturbance  $\mathcal{J}_r$ .

## 5. Costless identification experiment

In this section, we show that there are conditions under which Constraint 1 can be achieved using an identification experiment on the closed loop  $[C_{\text{id}} G_0]$  without any external excitation signal  $r(t)$ , i.e. using only the excitation due to the noise  $v(t)$ . We have called such experiments “costless identification” since  $r = 0$  also implies  $\mathcal{J}_r = 0$ . The main condition for this to be possible is that Lemma 1 holds in this situation or, in other words, that the closed-loop identification experiment with  $r(t) = 0$  is informative enough i.e. that the expected identification criterion  $\bar{V}(\theta)$  defined in (4) has a unique global minimum at  $\theta_0$ . A necessary and sufficient condition for this to hold is that  $\varepsilon(t, \theta) = \varepsilon(t, \theta_0) \implies \theta = \theta_0$ . In closed-loop identification with  $r = 0$ , this condition specializes to the following result (Söderström & Stoica, 1989, p. 392):  $\bar{V}(\theta)$  has a unique global minimum at  $\theta_0$  if and only if, for any  $\theta$ ,

$$\begin{aligned} H^{-1}(z, \theta)(1 + C_{\text{id}}G(z, \theta)) &= H^{-1}(z, \theta_0)(1 + C_{\text{id}}G(z, \theta_0)) \\ \implies \theta &= \theta_0 \end{aligned} \quad (12)$$

A sufficient condition for this identifiability condition to hold, in the case of a linear time-invariant regulator  $C_{\text{id}}$  considered here, is that the regulator be sufficiently complex (i.e. of sufficiently high order). One can make this statement more precise by considering specific model structures. This has been done in Söderström and Stoica (1989, p. 416) in the case where the model structure  $\mathcal{M}$  is ARMAX or ARX. Here we extend the results of Söderström and Stoica (1989) to OE and BJ model structures. The model structures we consider are all special cases of the general class  $\mathcal{M} = \{G(z, \theta); H(z, \theta)\}$  of model structures having the following parametrization:

$$G(z, \theta) = \frac{z^{-n_k} B(z, \theta)}{F(z, \theta)A(z, \theta)}, \quad H(z, \theta) = \frac{C(z, \theta)}{D(z, \theta)A(z, \theta)}$$

with  $A(z, \theta)$ ,  $B(z, \theta)$ ,  $C(z, \theta)$ ,  $D(z, \theta)$  and  $F(z, \theta)$  polynomials of degree  $n_a$ ,  $n_b$ ,  $n_c$ ,  $n_d$  and  $n_f$ , respectively. We also consider that the controller  $C_{\text{id}}$  is represented as  $C_{\text{id}} = X/Y$  where  $X$  and  $Y$  are polynomials of degree  $n_x$  and  $n_y$ , respectively.

We make the following assumption concerning the true closed-loop system and the parametric model structures.

**Assumption 2.** Consider the true system  $[G(z, \theta_0) H(z, \theta_0)]$  in closed loop with the controller  $C_{\text{id}} = X/Y$  where  $X$  and  $Y$  are polynomials of degree  $n_x$  and  $n_y$ , respectively. The identification is performed with a model structure that is contained in the general class of structures described above, with the following assumptions:

- (1) for the true system (i.e. with  $\theta = \theta_0$ ) there is no common factor in the polynomials  $\{A, B, C\}$ , the polynomials  $\{B, F\}$  are coprime, and the polynomials  $\{C, D\}$  are coprime;
- (2) the controller polynomials  $X$  and  $Y$  are coprime;
- (3) the closed-loop denominator polynomial does not cancel any root of  $C(z, \theta_0)$ ;
- (4) the degrees  $n_k, n_a, n_b, n_c, n_d, n_f$  used in the model structures are those of the true system (i.e. they are assumed known).

The coprimeness assumptions for the true system and for the controller parametrization are very weak. The assumption that the zeros of the noise model are not cancelled by the closed-loop poles is generically satisfied given that the controller  $C_{\text{id}}$  is fixed, while the noise model numerator  $C(z, \theta_0)$  is unknown. The only really restrictive assumption is the last one on exact knowledge of the degrees of the true system.

**Proposition 1.** Consider the closed-loop identification described in Section 2 with  $C_{\text{id}}$  a linear time-invariant controller, and assume that  $r(t) = 0 \forall t$ . Consider that the controller and the chosen model structure obey Assumption 2. Then,  $\bar{V}(\theta)$  has a unique minimum in  $\theta_0$ :

- in the case  $\mathcal{M} = \text{ARX}$  ( $C = D = F = 1$ ) or  $\text{ARMAX}$  ( $D = F = 1$ ) if and only if either  $n_x + n_k > n_a$  or  $n_y > n_b$ ;
- in the case  $\mathcal{M} = \text{BJ}$  ( $A = 1$ ) if and only if  $n_x + n_k > n_d + n_f$  or  $n_y > n_b + n_d$ ;
- in the case  $\mathcal{M} = \text{OE}$  if and only if the controller  $C_{\text{id}}(z)$  is not identically zero.

**Proof.** The proof consists of verifying (12) for each of these model structures. For the ARX and ARMAX model structures, the result has been established in Söderström and Stoica (1989, p. 416). In the Box–Jenkins case, the left-hand side of (12) becomes

$$\begin{aligned} & \frac{D(\theta)[F(\theta)Y + z^{-n_k}XB(\theta)]}{C(\theta)F(\theta)} \\ &= \frac{D(\theta_0)[F(\theta_0)Y + z^{-n_k}XB(\theta_0)]}{C(\theta_0)F(\theta_0)}. \end{aligned} \quad (13)$$

The polynomials  $C(\theta_0)$  and  $D(\theta_0)$  are coprime by condition 1 of Assumption 2, while the polynomials  $C(\theta_0)$  and

$F(\theta_0)Y + z^{-n_k}XB(\theta_0)$  are coprime by condition 3. Suppose now that  $F(\theta_0)$  and  $D(\theta_0)$  have a common polynomial factor  $H$ , i.e.

$$F(\theta_0) = \bar{F}_0H, \quad D(\theta_0) = \bar{D}_0H. \quad (14)$$

Since, by condition 4 of Assumption 2,  $\deg C(\theta) + \deg F(\theta) = \deg C(\theta_0) + \deg F(\theta_0)$ , it follows that the solution set of (13) is given by  $C(\theta)F(\theta) = C(\theta_0)\bar{F}_0M$  and

$$D(\theta)[F(\theta)Y + z^{-n_k}XB(\theta)] = [F(\theta_0)Y + z^{-n_k}XB(\theta_0)]\bar{D}_0M,$$

where, by condition 4 of Assumption 2,  $M$  is an arbitrary polynomial of the same degree as that of  $H$ . Equivalently, the solution set is described by

$$\begin{aligned} C(\theta)F(\theta) &= C(\theta_0)\bar{F}_0M, \\ D(\theta)F(\theta) &= F(\theta_0)\bar{D}_0M + \xi z^{-n_k}X, \\ D(\theta)B(\theta) &= B(\theta_0)\bar{D}_0M - \xi Y, \end{aligned} \quad (15)$$

where  $\xi$  is an arbitrary polynomial. The solution set reduces to

$$\begin{aligned} C(\theta)F(\theta) &= C(\theta_0)\bar{F}_0M, \\ D(\theta)F(\theta) &= F(\theta_0)\bar{D}_0M, \\ D(\theta)B(\theta) &= B(\theta_0)\bar{D}_0M, \end{aligned} \quad (16)$$

if and only if either  $n_x + n_k > n_d + n_f$  or  $n_y > n_b + n_d$ . It now follows from the first and second of these equations, using (14), that

$$\frac{C(\theta)}{D(\theta)} = \frac{C(\theta_0)\bar{F}_0}{\bar{D}_0F(\theta_0)} = \frac{C(\theta_0)\bar{F}_0}{\bar{D}_0H\bar{F}_0} = \frac{C(\theta_0)}{D(\theta_0)}.$$

It follows from the second and third equation of (16) that:

$$\frac{B(\theta)}{F(\theta)} = \frac{B(\theta_0)}{F(\theta_0)}.$$

From these last two expressions and condition 4 of Assumption 2 about the known degrees, we conclude that the only solution is  $\theta = \theta_0$ .

Finally, for the OE case, the left-hand side of (12) becomes  $C_{\text{id}}G(\theta) = C_{\text{id}}G(\theta_0)$ . As long as the linear time-invariant controller  $C_{\text{id}}(z)$  is not identically zero, the latter implies  $\theta = \theta_0$ .  $\square$

We have given conditions on the controller complexity under which an identification experiment with no external excitation delivers a cost function  $\bar{V}(\theta)$  that has  $\theta_0$  as its unique global minimum, i.e. the closed-loop identification with  $r(t) = 0$  is informative enough. It is then trivial to show that, under those conditions, Constraint 1 can be made to hold.

**Theorem 1.** *Consider a closed-loop identification experiment as presented in Section 2 with  $r(t) = 0$  and with  $C_{\text{id}}$  satisfying the conditions in Proposition 1. Then, Constraint 1 can always be achieved by using for the identification a set of (normal operation) input–output data of sufficient length  $N$ .*

**Proof.** When  $r = 0$ , expression (6) of  $P_\theta^{-1}$  becomes  $P_{\theta_0}^{-1} = N\mathcal{P}_v^{-1}(\theta_0)$ . Since the experiment is informative enough,

we have, via Lemma 1,  $\mathcal{P}_v^{-1}(\theta_0) > 0$ . Therefore, with  $N$  sufficiently large,  $P_\theta^{-1} = N\mathcal{P}_v^{-1}(\theta_0)$  can be made such that  $\mathcal{D}(\hat{\theta}_N, P_\theta) \subseteq \mathcal{D}_{\text{adm}}(\hat{\theta}_N)$  for any set  $\mathcal{D}_{\text{adm}}(\hat{\theta}_N)$  around  $G(z, \hat{\theta}_N)$ , which implies that Constraint 1 holds (see Section 4).  $\square$

Theorem 1 shows that, if the controller complexity is chosen such that the closed-loop experiment with  $r = 0$  is informative enough, the identification leading to a new and satisfactory controller  $\hat{C}$  for  $G_0$  can be achieved without identification cost: we just need to measure the input and output signal in normal operation (i.e. with  $r = 0$ ) for a sufficient amount of time. The computation of the smallest data length  $N_{\text{min}}$  for which the robustness Constraint 1 is satisfied with  $r = 0$  is the special case of Situation 2 when the given reference excitation spectrum  $\Phi_r(\omega)$  is equal to 0; a method to compute this minimal number of data will be presented in Section 6.

## 6. Least costly identification when costless identification is impossible

An identification with  $r = 0$  may be impossible or non desirable for three reasons: (i) the initial controller  $C_{\text{id}}$ , which can often not be chosen by the user, is of lower complexity than is required by the conditions of Proposition 1, (ii) the data length required to satisfy Constraint 1 (see Theorem 1) is unrealistic (e.g. too long) or (iii) the controller  $C_{\text{id}}$  performs so badly that it is more economical to seek for the shortest identification duration (see Situation 2, Section 4). When identification with  $r = 0$  is ruled out, the least costly identification experiment for control is defined by the problems presented at the end of Section 4. Those problems involve the computation under Constraint 1 of either the power spectrum  $\Phi_r(\omega)$  minimizing  $\mathcal{J}_r$  for a given  $N$  (Situations 1 and 3) or the smallest data length for a given  $\Phi_r(\omega)$  (Situation 2).

We first examine Situation 1 in which the data length  $N$  is fixed and one seeks to satisfy Constraint 1 with a power spectrum  $\Phi_r(\omega)$  that minimizes the identification cost  $\mathcal{J}_r$  defined in (11). We restrict our search to signals whose power spectrum  $\Phi_r(\omega)$  can be written as (Lindqvist, 2001)

$$\Phi_r(\omega) = R_r(0) + 2 \sum_{i=1}^m R_r(i) \cos(i\omega) \geq 0 \quad \forall \omega, \quad (17)$$

where  $m$  is a positive integer selected by the user. By designing a spectrum as (17), we always ensure that the corresponding identification experiment is informative enough and thus that Lemma 1 holds. Indeed, the parameters  $R_r(i)$  ( $i = 0 \dots m$ ) can be interpreted as the auto-correlation sequence of a signal that has been generated by a white noise passing through an FIR filter of length  $m + 1$ .

Another important property of the parametrization (17) is that  $P_\theta^{-1}$  and  $\mathcal{J}_r$  (see (6) and (11)) are affine functions of the design variables  $R_r(i)$  ( $i = 0 \dots m$ ), as we show in the following proposition. Note that there are other parametrizations of  $\Phi_r(\omega)$  that have the same property and could therefore also be considered here: e.g.  $\Phi_r(\omega) = \sum_{i=1}^m R_r(i) \delta(\omega - \omega_i)$  corresponding to

a multisine signal  $r(t)$  (Van den Eijnde & Schoukens, 1991), or  $\Phi_r(\omega) = \sum_{i=0}^m R_r(i) (\mathcal{B}_i(e^{j\omega}) + \mathcal{B}_i^*(e^{j\omega}))$  where  $\mathcal{B}_i(e^{j\omega})$  are preselected basis functions (Jansson & Hjalmarsson, 2004).

**Proposition 2.** Consider expression (6) of  $P_\theta^{-1}$  and let  $\Phi_r(\omega)$  be parametrized by (17). Let  $\tilde{M}_k(\theta_0)$  be the sequence of Markov parameters of  $F_r F_r^*$  i.e.  $F_r(e^{j\omega}, \theta_0) F_r(e^{j\omega}, \theta_0)^* = \sum_{k=-\infty}^{\infty} \tilde{M}_k(\theta_0) e^{-jk\omega}$  with  $F_r(z, \theta_0)$  as defined in (6). Then,  $P_\theta^{-1} \in \mathbf{R}^{k \times k}$  can be written as

$$P_\theta^{-1} = \bar{M}(\theta_0) + \sum_{i=0}^m M_i(\theta_0, \sigma_e^2) R_r(i),$$

where  $\bar{M}(\theta_0) = N \mathcal{P}_v^{-1}(\theta_0)$ ,  $M_0(\theta_0, \sigma_e^2) = (N/\sigma_e^2) \tilde{M}_0(\theta_0)$ , and  $M_i(\theta_0, \sigma_e^2) = (N/\sigma_e^2) (\tilde{M}_i(\theta_0) + \tilde{M}_i^T(\theta_0))$  for  $i = 1 \dots m$ . Moreover, the cost function  $\mathcal{J}_r$  defined in (11) can be written as

$$\begin{aligned} \mathcal{J}_r &= [\alpha_y c_0(\theta_0) + \alpha_u d_0(\theta_0)] R_r(0) \\ &+ 2 \sum_{i=1}^m [\alpha_y c_i(\theta_0) + \alpha_u d_i(\theta_0)] R_r(i), \end{aligned}$$

where the coefficients  $c_i(\theta_0)$  and  $d_i(\theta_0)$  are the Markov parameters of  $G_0 G_0^* S_{id} S_{id}^*$  and  $S_{id} S_{id}^*$ , respectively, i.e.  $G_0(e^{j\omega}) G_0(e^{j\omega})^* S_{id}(e^{j\omega}) S_{id}(e^{j\omega})^* = \sum_{k=-\infty}^{\infty} c_k(\theta_0) e^{-jk\omega}$  and  $S_{id}(e^{j\omega}) S_{id}(e^{j\omega})^* = \sum_{k=-\infty}^{\infty} d_k(\theta_0) e^{-jk\omega}$ .

**Proof.** The first part of the proposition is a direct consequence of Result 5.6 of Lindqvist (2001) applied to the closed-loop expression for  $P_\theta^{-1}$  as given in (6). The second part is a direct consequence of Result 5.4 of Lindqvist (2001) applied to  $y_r(t) = G_0 S_{id} r(t)$  and  $u_r(t) = S_{id} r(t)$ .  $\square$

With the parametrization (17) for  $\Phi_r(\omega)$ , the experiment design problem corresponding to Situation 1 can then be formulated as follows:

**Experiment Design Problem 1.** Consider the closed-loop identification experiment of Section 2 with a fixed number  $N$  of data. Determine the parameters  $R_r(i)$  ( $i = 0 \dots m$ ) of the spectrum  $\Phi_r(\omega)$  in (17) which minimize  $\mathcal{J}_r$ , subject to satisfaction of Constraint 1.

We show in the sequel that this problem can be expressed as an LMI-based optimization problem (Boyd, El Ghaoui, Feron, & Balakrishnan, 1994). For this purpose, we express the robust performance constraint  $\bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\hat{\theta}_N, P_\theta)$  at one particular frequency  $\omega$  as an LMI, linear in  $P_\theta^{-1}$ . Note that, according to (8),  $J(G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\hat{\theta}_N, P_\theta) \iff \bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1 \forall \omega$  and  $\forall G \in \mathcal{D}(\hat{\theta}_N, P_\theta)$ .

**Proposition 3.** Consider the controller  $\hat{C}$  designed from the model  $G(z, \hat{\theta}_N)$  using the control design method presented in Assumption 1. Consider the set  $\mathcal{D}(\hat{\theta}_N, P_\theta)$  defined in (7). Then  $\hat{C}$  achieves  $\bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1$  with all  $G$  in  $\mathcal{D}(\hat{\theta}_N, P_\theta)$

if and only if  $\exists \tau(\omega) > 0$ ,  $\tau(\omega) \in \mathbf{R}$  and  $\mathcal{L}(\omega) = -\mathcal{L}(\omega)^T \in \mathbf{R}^{(k+1) \times (k+1)}$  such that

$$\tau(\omega) E(\omega, \hat{\theta}_N) - \mathcal{R}(\hat{\theta}_N) + j\mathcal{L}(\omega) \leq 0, \quad (18)$$

where  $E(\omega, \hat{\theta}_N) = \Omega^*(e^{j\omega}) \begin{pmatrix} I_4 & 0 \\ 0 & -1 \end{pmatrix} \Omega(e^{j\omega})$  with

$$\Omega = \left( \begin{array}{c|c} \left( I_2 \otimes \left( W_r \begin{pmatrix} \hat{C} \\ 1 \end{pmatrix} \right) \right) & W_l \cdot 0 \\ \hline 0 & 1 \end{array} \right) \begin{pmatrix} Z_N & 0 \\ Z_D & 1 \\ Z_D + \hat{C} Z_N & 1 \end{pmatrix}$$

$\Omega(z)$  has the dimension  $5 \times (k+1)$  and  $\otimes$  denotes the Kronecker product. Finally

$$\begin{aligned} \mathcal{R}(\hat{\theta}_N) &= \begin{pmatrix} I_k \\ -\hat{\theta}_N^T \end{pmatrix} P_\theta^{-1} \begin{pmatrix} I_k \\ -\hat{\theta}_N^T \end{pmatrix}^T \\ &+ \begin{pmatrix} 0 & 0 \\ 0 & -\chi \end{pmatrix} \in \mathbf{R}^{(k+1) \times (k+1)}. \end{aligned}$$

**Proof.** See Appendix A.

The previous proposition shows that Constraint 1 can be replaced by LMIs at each frequency, linear in  $P_\theta^{-1}$  and thus, via Proposition 2, linear in the decision variables  $R_r(i)$  of Experiment Design Problem 1. Combining this fact with the fact that its cost function  $\mathcal{J}_r$  is also linear in  $R_r(i)$ , Experiment Design Problem 1 would be solvable exactly if the parametrizations of  $P_\theta^{-1}$  and  $\mathcal{J}_r$  with respect to the design variables  $R_r(i)$  were not functions of the unknown  $\theta_0$  and  $\sigma_e^2$ , and if condition (18) was not a function of the to-be-identified  $\hat{\theta}_N$ . This difficulty can be circumvented by using a priori estimates for those quantities:  $\theta_{o,\text{est}}$ ,  $\sigma_{e,\text{est}}^2$  and  $\hat{\theta}_{N,\text{est}}$ . The solution is then obtained by solving the LMI optimization problem described in Theorem 2 below. In Section 7 we shall present an attractive alternative to the reliance on a priori estimates; it is based on a small adaptation of the LMI problem of Theorem 2.

**Theorem 2.** Consider the approximations  $\theta_0 \approx \theta_{o,\text{est}}$ ,  $\hat{\theta}_N \approx \hat{\theta}_{N,\text{est}}$  and  $\sigma_e^2 \approx \sigma_{e,\text{est}}^2$  and the shorthand notations:  $c_i = c_i(\theta_{o,\text{est}})$ ,  $d_i = d_i(\theta_{o,\text{est}})$ ,  $\bar{M} = \bar{M}(\theta_{o,\text{est}})$ ,  $M_i = M_i(\theta_{o,\text{est}}, \sigma_{e,\text{est}}^2)$  and  $E(\omega) = E(\omega, \hat{\theta}_{N,\text{est}})$ . Then the auto-correlation sequence  $R_r(i)$  ( $i=0 \dots m$ ) which solves Experiment Design Problem 1 is the solution of the following LMI optimization problem:

$$\min_{R_r(i) (i=0 \dots m)} [\alpha_y c_0 + \alpha_u d_0] R_r(0) + 2 \sum_{i=1}^m [\alpha_y c_i + \alpha_u d_i] R_r(i)$$

under the constraint that there exist a matrix  $Q = Q^T$  of appropriate dimension,  $\tau(\omega) > 0$ ,  $\tau(\omega) \in \mathbf{R}$  and  $\mathcal{L}(\omega) = -\mathcal{L}(\omega)^T \in \mathbf{R}^{(k+1) \times (k+1)}$  such that

$$\begin{aligned} \tau(\omega) E(e^{j\omega}) - \begin{pmatrix} I_k \\ -\hat{\theta}_{N,\text{est}}^T \end{pmatrix} \left( \bar{M} + \sum_{i=0}^m M_i R_r(i) \right) \begin{pmatrix} I_k \\ -\hat{\theta}_{N,\text{est}}^T \end{pmatrix}^T \\ - \begin{pmatrix} 0 & 0 \\ 0 & -\chi \end{pmatrix} + j\mathcal{L}(\omega) \leq 0 \quad \forall \omega \end{aligned} \quad (19)$$

and that

$$\begin{pmatrix} Q - A^T Q A & C^T - A^T Q B \\ C - B^T Q A & D + D^T - B^T Q B \end{pmatrix} \geq 0 \quad (20)$$

with the following definitions of  $A, B, C, D$ :

$$A = \begin{pmatrix} 0 & 0 \\ I_{m-1} & 0 \end{pmatrix}, \quad B = (1 \ 0 \ \dots \ 0),$$

$$C = (R_r(1) \ R_r(2) \ \dots \ R_r(m)), \quad D = \frac{R_r(0)}{2}.$$

The optimal spectrum  $\Phi_r(\omega)$  can then be computed using (17).

**Proof.** As shown in Lindqvist (2001) via the positive real lemma, the constraint that there exists a symmetric matrix  $Q$  such that (20) holds is a necessary and sufficient condition for  $R_r(0) + 2\sum_{i=1}^m R_r(i) \cos(i\omega)$  to be positive at each  $\omega$  and thus for (17) to represent a spectrum. Consequently, the result in this theorem is a direct consequence of Propositions 2 and 3.  $\square$

**Comment 1.** Condition (19) must be considered at every frequency. This is impossible in practice. The optimal spectrum can nevertheless be approximated by using a finite frequency grid. An exact but more cumbersome solution consists of using the Kalman–Yakovovitch–Popov (KYP) Lemma (Popov, 1973): see Appendix B.

**Comment 2.** Additional constraints such as  $\Phi_u(\omega) < \beta_u(\omega)$  and  $\Phi_y(\omega) < \beta_y(\omega) \ \forall \omega$  can also be treated (see Lindqvist, 2001).

We now consider the experiment design problem corresponding to Situation 2.

**Experiment Design Problem 2.** Consider the closed-loop identification experiment of Section 2. Consider also that the power spectrum  $\Phi_r(\omega)$  of the excitation signal  $r(t)$  is given and leads to an informative enough identification experiment. Determine then the smallest length  $N$  of a part of  $r(t)$  that must be applied to  $[C_{id} \ G_0]$  in order to fulfill Constraint 1.

Since (6) shows that  $P_\theta^{-1}$  is linear in the decision variable  $N$  of Experiment Design Problem 2, it can be solved in a similar way and with the same approximations as Experiment Design Problem 1.

**Theorem 3.** Consider the approximations  $\theta_0 \approx \theta_{o,est}$ ,  $\hat{\theta}_N \approx \hat{\theta}_{N,est}$  and  $\sigma_e^2 \approx \sigma_{e,est}^2$  and the shorthand notations:  $\mathcal{P}_r^{-1}(\Phi_r(\omega)) = \mathcal{P}_r^{-1}(\Phi_r(\omega), \theta_{o,est}, \sigma_{e,est}^2)$ ,  $\mathcal{P}_v^{-1} = \mathcal{P}_v^{-1}(\theta_{o,est})$  (see (6)) and  $E(\omega) = E(\omega, \hat{\theta}_{N,est})$  (see (18)). Then, the minimum duration  $N$  which solves Experiment Design Problem 2 is the solution (rounded up to the nearest integer) of the LMI optimization problem minimizing  $N$  under the constraint that there exist  $\tau(\omega) > 0$ ,  $\tau(\omega) \in \mathbf{R}$  and  $\mathcal{L}(\omega) = -\mathcal{L}(\omega)^T \in \mathbf{R}^{(k+1) \times (k+1)}$  such that the constraint (19) with  $\bar{M} + \sum_{i=0}^m M_i R_r(i)$  replaced by  $N(\mathcal{P}_r^{-1}(\Phi_r(\omega)) + \mathcal{P}_v^{-1})$  holds at each frequency  $\omega$ .

## 7. A robust optimal design procedure

Theorems 2 and 3 require that initial estimates of  $\theta_0, \hat{\theta}_N$  and  $\sigma_e^2$  be used. If those approximations are not accurate, this could lead to poor results. In this section, we present a procedure that renders the computed optimal design solution more robust w.r.t. these unknown quantities. Very often, estimates of  $\theta_0$  and  $\sigma_e^2$  are available, because the initial controller  $C_{id}$  has been computed from an identified model. This initial identification delivers estimates not only of  $\theta_0$  and  $\sigma_e^2$ , but also of uncertainty regions for those quantities. It is then possible to deduce from this initial identification a (truncated) Gaussian probability density function which defines the likelihood of each element of these uncertainty regions. If this is not possible, the density functions are chosen uniform over the uncertainty regions. The estimate, the uncertainty region and the probability density function of the to-be-identified  $\hat{\theta}_N$  are typically chosen equal to those of  $\theta_0$ . To summarize, from an initial identification one can assume that  $q_0 = (\theta_0^T \ \hat{\theta}_N^T \ \sigma_e^2)^T$  lies in a set  $\mathcal{Q}$  and that the likelihood of the event  $q = q_0$  is given by a probability function  $p(q)$ . Based on this information, one can robustify the procedure that consists in adopting a unique and possibly poor estimate of  $q_0$  for the design of the experimental conditions. This is achieved by the use of randomized algorithms (see e.g. Tempo, Bai, & Dabbene, 1997; Tempo, Calafiore, & Dabbene, 2004). In the case of Experiment Design Problem 2, we want to determine the smallest duration  $N$  for which Constraint 1 is verified for all possible values of  $q_0$  in  $\mathcal{Q}$ . This is equivalent to computing an estimate  $\hat{N}$  of  $\sup_{q \in \mathcal{Q}} N_q$  where  $N_q$  is the solution obtained by Theorem 3 with the approximation  $q$ . Considering  $N_q$  as a function of  $q$ , this can be done (Tempo et al., 1997) with accuracy  $\varepsilon$  and confidence<sup>1</sup>  $\delta$  by generating  $n \geq \ln(\delta^{-1}) / \ln((1 - \varepsilon)^{-1})$  estimates  $q_j$  of  $q_0$  according to the probability density function  $p(q)$ , and by determining the corresponding optimal  $N_{q_j}$  for each of these estimates  $q_j$  using Theorem 3. An estimate  $\hat{N}$  of  $\sup_{q \in \mathcal{Q}} N_q$  is then given by  $\sup_{j=1, \dots, n} N_{q_j}$ . In the case of Experiment Design Problem 1, the approach above cannot be considered since we determine the parameters  $R_r(i)$  ( $i = 0 \dots m$ ) of  $\Phi_r(\omega)$  rather than  $\Phi_r(\omega)$  itself. Instead, the so called scenario-approach can be considered (see e.g. Tempo et al., 2004, Chapter 12). This approach is also based on a randomized algorithm which uses the probability density function  $p(q)$ . However, whereas for Experiment Design Problem 2 we apply Theorem 3 a fixed number of times for estimates of  $q_0$  randomly generated from  $p(q)$ , for Experiment Design Problem 1 we solve the optimization problem only once, but with several robust performance constraints (19), each evaluated at a different estimate of  $q_0$  randomly generated from  $p(q)$ . The scenario approach is thus similar to the idea presented in Jansson and Hjalmarsson (2004) (and Bombois et al., 2004) which was based on a gridding of  $\mathcal{Q}$ .

## 8. Simulation results

In order to illustrate our results, we consider as true system the following ARX system (Landau, Rey, Karimi, Voda,

<sup>1</sup> This means that  $Pr(Pr(N_q > \hat{N}) \leq \varepsilon) \geq 1 - \delta$ .

& Franco, 1995):  $y(t) = (z^{-3}B_0(z))/(A_0(z))u(t) + (1)/(A_0(z))e(t)$  with  $B_0(z) = 0.10276 + 0.18123z^{-1}$ ,  $A_0(z) = 1 - 1.99185z^{-1} + 2.20265z^{-2} - 1.84083z^{-3} + 0.89413z^{-4}$ , and  $e(t)$  a realization of a white noise process of variance  $\sigma_e^2 = 0.5$ . The control performance criterion  $J(G, C, W_l, W_r)$  focuses on the sensitivity function. It is defined as in (8) with the filters:  $W_l(z) = \text{diag}(0, W(z))$ ,  $W_r(z) = \text{diag}(0, 1)$  with  $W(z) = (0.5165 - 0.4632z^{-1})/(1 - 0.999455z^{-1})$ .

The true system initially operates in closed loop with a controller  $C_{id}$  which has been designed using an initial estimate of the true system  $\theta_{o,est} = (-1.9755, 2.1965, -1.8495, 0.8881, 0.0817, 0.172)^T$  and the 4-block  $H_\infty$  control design method of Ferreres and Fromion (1997) that satisfies Assumption 1:  $C_{id} = (0.036249(z + 0.9244)(z^2 - 1.951z + 1.101)(z^2 - 0.5109z + 0.8248)(z^2 - 0.1828z + 0.9416))/((z - 0.9995)(z^2 - 1.002z + 0.3641)(z^2 - 1.279z + 0.835)(z^2 - 0.1746z + 0.9229))$ .

*Costless identification:* The order of the initial controller  $C_{id}$  is sufficient for the experiment with  $r = 0$  to be informative enough (see Proposition 1). Thus, Constraint 1 can be verified with an identification experiment using normal operation signals  $y$  and  $u$ . The minimal length for this data set can then be determined using Theorem 3 with  $\Phi_r(\omega) = 0 \forall \omega$ . This theorem is applied here using the initial estimates  $\theta_0 \approx \theta_{o,est}$  and  $\hat{\theta}_N \approx \theta_{o,est}$  (an estimate of  $\sigma_e^2$  is not necessary since  $\mathcal{P}_v^{-1}(\theta_0)$  is not a function of  $\sigma_e^2$ ). This delivers a minimal length  $N_{min} = 4901$ .

In order to verify the validity of this result, we have measured 4901 samples of the signals  $y(t)$  and  $u(t)$  obtained in normal operation on the loop  $[C_{id} G_0]$  and we have identified a model  $\hat{G} = G(z, \hat{\theta}_N)$  along with its uncertainty region  $\mathcal{D}(\hat{\theta}_N, P_\theta)$ . From  $\hat{G}$ , we have then designed a controller  $\hat{C}$  using the method of Ferreres and Fromion (1997) and we have verified whether  $\hat{C}$  achieves  $J(G, \hat{C}, W_l, W_r) \leq 1$  with all  $G$  in  $\mathcal{D}(\hat{\theta}_N, P_\theta)$ , or equivalently  $|1/1 + \hat{C}(e^{j\omega})G(e^{j\omega})| \leq |W(e^{j\omega})|^{-1}$  for all  $G$  in  $\mathcal{D}(\hat{\theta}_N, P_\theta)$ . This is indeed the case as can be seen in Fig. 2. Moreover, we also observe in Fig. 2 that  $\sup_{G \in \mathcal{D}(\hat{\theta}_N, P_\theta)} |1/1 + \hat{C}(e^{j\omega})G(e^{j\omega})| = |W(e^{j\omega})|^{-1}$  in low frequencies. Consequently,  $N = 4901$  is indeed the smallest  $N$  for which, in this example, Constraint 1 holds with  $\Phi_r(\omega) = 0 \forall \omega$ .

*Sensitivity to the initial estimates:* The initial estimate  $\theta_{o,est}$  chosen to approximate the unknown quantities  $\theta_0$  and  $\hat{\theta}_N$  has, in this example, delivered accurate results, as shown by Fig. 2. This may not always be the case. Thus, it is always safer to compute the minimal data length  $N$  using the method proposed in Section 7. We illustrate the application of this method to our example. In order to generate multiple estimates of  $\theta_0$  (which are then used to approximate both the true  $\theta_0$  and  $\hat{\theta}_N$ ), we have used the information delivered by the initial identification which had delivered  $\theta_{o,est}$ . This was an open-loop identification with  $\Phi_u(\omega) = 1 \forall \omega$  and  $N = 500$ . Using the covariance matrix of  $\theta_{o,est}$ , we have randomly generated 46 parameter vectors  $\theta_i$  ( $i = 1 \dots 46$ ) around  $\theta_{o,est}$ ; 46 samples correspond to an accuracy and a confidence of 80%. For each of these estimates, we have applied Theorem 3 and we have thus obtained 46 different lengths  $N_{\theta_i}$ . A more robust choice of the length  $N$  is

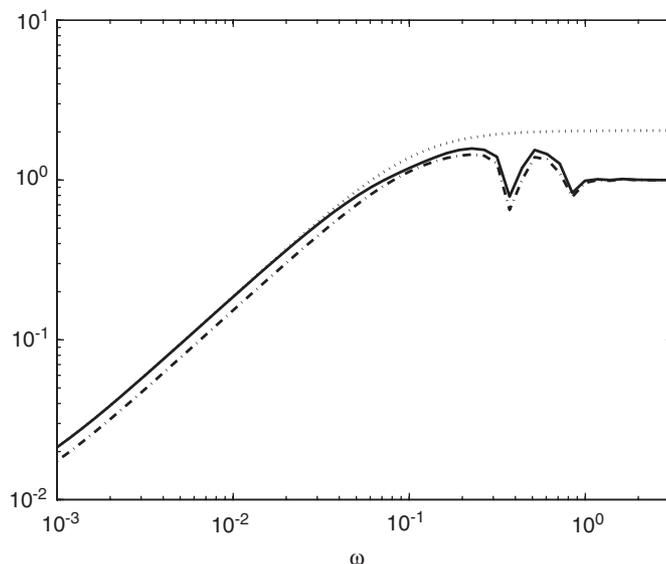


Fig. 2. Costless identification with  $N = 4901$ :  $\sup_{G \in \mathcal{D}(\hat{\theta}_N, P_\theta)} |1/1 + \hat{C}(e^{j\omega})G(e^{j\omega})|$  (solid),  $|1 + \hat{C}\hat{G}|^{-1}$  (dashdot) and  $|W|^{-1}$  (dotted).

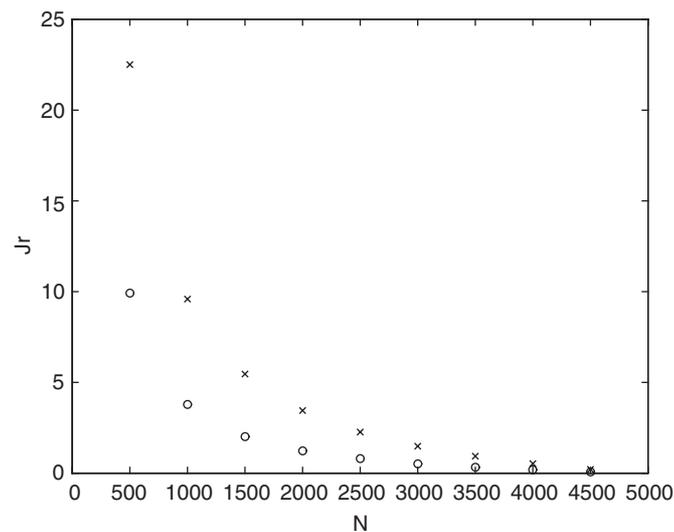


Fig. 3. Cost of the least costly experiment that satisfies Constraint 1 (circles) and of the white noise experiment that satisfies the same constraint (crosses), for different values of the data length.

then (see Section 7):  $\max_{\theta_i} N_{\theta_i} = 5897$ . Note that the standard deviation of the 46 found values of  $N_{\theta_i}$  was 383.

*Least costly identification with external excitation:* For the same example, we have also studied the effect of applying least costly identification signals for a range of data lengths that were too short to lead to identifiability using only the noise excitation. Thus, we have computed the optimal signal spectra  $\Phi_r(\omega)$  using Theorem 2 for data lengths  $N$  ranging from 500 to 4500 by steps of 500. For each of these data lengths, we have then compared the identification cost  $\mathcal{J}_r$  resulting from the application of the optimal excitation signal (when  $m = 10$  in (17)) with the cost that would result by applying a white

<sup>2</sup> This initial identification was too cheap to verify Constraint 1.

noise reference excitation (i.e.  $m = 0$  in (17)) with a variance that is sufficient to satisfy the robust performance Constraint 1. The comparison between the cost of the least costly experiment and the cost of a corresponding identification experiment with white noise excitation is shown in Fig. 3. As can be seen from this figure, the use of an optimally designed excitation signal  $r$  reduces the identification cost by a factor of 2–3 whatever the data length. Note that similar comparisons leading to similar conclusions can be found for the open-loop identification case in Jansson (2004).

**9. Conclusions**

We have proposed a new paradigm for optimal experiment design in an identification for robust control context, where the objective is to design an identification experiment at the smallest possible cost that just meets the demands of the robust control performance in terms of the quality constraints on the estimated model. In other words, the identification cost must be as small as possible while still delivering a model uncertainty set that just meets the robust control performance constraints. The cost is expressed either as the experiment time, or in terms of a measure of the deterioration of the closed-loop performance, during the identification experiment.

Even though these are classical assumptions for optimal experimental design in the PE framework, the assumptions that the true system can be described by the chosen model structure  $\mathcal{M}$  and the use of the covariance matrix expression asymptotic in the data length are the restrictive aspects of the methodology. The first assumption can be relaxed. In the recent contribution (Bombois & Gilson, 2006), the “least costly identification for control” framework is indeed extended to the case of model structures  $\mathcal{M}$  of reduced order. Further research will also focus on the relaxation of the second assumption. See e.g. the results in Jansson (2004, Chapter 4) and Bombois and Gilson (2006). Note that the counterpart for both relaxations is that only model structures linear in the parameter vector can be considered.

**Appendix A. Proof of Proposition 3**

**Lemma 2.** Let  $A = A^* \in \mathbb{C}^{n \times n}$ . Then

$$y^T A y \leq 0 \quad \text{for all } y \in \mathbb{R}^n \tag{A.1}$$

if and only if

$$\text{there exists } \mathcal{L} = -\mathcal{L}^T \in \mathbb{R}^{n \times n} \text{ such that } A + j\mathcal{L} \leq 0. \tag{A.2}$$

**Proof of Lemma 2.** (A.2)  $\Rightarrow$  (A.1). Note that, for any  $y \in \mathbb{R}^n$  and for any  $\mathcal{L} = -\mathcal{L}^T \in \mathbb{R}^{n \times n}$ ,  $y^T j\mathcal{L}y = 0$ . Condition (A.2) implies then that, for all  $y \in \mathbb{R}^n$ ,  $y^T A y \leq 0$ .

(A.1)  $\Rightarrow$  (A.2) is proved by contradiction. Assume thus that (A.1) holds and that:

$$\text{there is no } \mathcal{L} = -\mathcal{L}^T \in \mathbb{R}^{n \times n} \text{ such that } A + j\mathcal{L} \leq 0. \tag{A.3}$$

Since Condition (A.1) holds, we have that

$$\text{for all } S = S^T \geq 0 \in \mathbb{R}^{n \times n}, \quad \text{Trace}(SA) \leq 0. \tag{A.4}$$

On the other hand, based on Lemma III.1 in Meinsma, Shrivastava, and Fu (1997), condition (A.3) is equivalent to the fact that there exists  $W = W^* \geq 0$  ( $W \in \mathbb{C}^{n \times n}$ ) such that:

$$\text{Trace}(W(A + j\mathcal{L})) > 0 \quad \text{for all } \mathcal{L} = -\mathcal{L}^T \in \mathbb{R}^{n \times n}. \tag{A.5}$$

Note now that

$$\text{Trace}(Wj\mathcal{L}) = -\text{Trace}(\text{Im}(W)\mathcal{L}) = 2 \sum_{i=1}^{n(n-1)/2} w_i l_i,$$

where  $w_i$  ( $i = 1 \dots n(n-1)/2$ ) are the  $n(n-1)/2$  nonzeros entries of the upper-triangular part of the skew symmetric matrix  $\text{Im}(W)$  and  $l_i$  are the corresponding entries of  $\mathcal{L}$  in the same location. Thus, (A.5) is equivalent to

$$\text{Trace}(WA) + 2 \sum_{i=1}^{n(n-1)/2} w_i l_i > 0 \quad \text{for all } l_i \in \mathbb{R}$$

which is equivalent to  $w_i = 0$  (i.e. the matrix  $W$  is real) and  $\text{Trace}(WA) > 0$ . The latter is in contradiction with (A.4) and the proof is thus completed.  $\square$

**Proof of Proposition 3.** From Bombois et al. (2001), for  $G \in \mathcal{D}(\hat{\theta}_N, P_\theta)$ ,  $\bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1$  can be written as

$$\mathcal{G}^*(e^{j\omega}) \begin{pmatrix} x^* x I_2 & 0 \\ 0 & -1 \end{pmatrix} \mathcal{G}(e^{j\omega}) \leq 0 \tag{A.6}$$

with  $x = W_r(e^{j\omega})(\hat{C}(e^{j\omega}) - 1)^T$  and

$$\mathcal{G} = \begin{pmatrix} W_l & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Z_N \theta \\ 1 + Z_D \theta \\ 1 + (Z_D + \hat{C} Z_N) \theta \end{pmatrix}.$$

If we define  $\bar{\theta} = (\theta^T \ 1)^T$ , expression (A.6) can be rewritten as  $\bar{\theta}^T E(\omega, \hat{\theta}_N) \bar{\theta} \leq 0$ . On the other hand, the constraint  $\theta \in U$  of (7) can be rewritten as  $\bar{\theta}^T \mathcal{R}(\hat{\theta}_N) \bar{\theta} \leq 0$ . Consequently, proving that  $\bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1$  for all  $G \in \mathcal{D}(\hat{\theta}_N, P_\theta)$  is equivalent to proving that  $\bar{\theta}^T E(\omega, \hat{\theta}_N) \bar{\theta} \leq 0$  for all  $\bar{\theta} \in \mathbb{R}^{(k+1) \times (k+1)}$  such that  $\bar{\theta}^T \mathcal{R}(\hat{\theta}_N) \bar{\theta} \leq 0$ . Using the  $\mathcal{S}$  procedure (Boyd et al., 1994), the latter is equivalent to proving that  $\exists \tau(\omega) > 0$  such that, for all  $\bar{\theta} \in \mathbb{R}^{(k+1) \times (k+1)}$ ,  $\bar{\theta}^T (\tau(\omega) E(\omega, \hat{\theta}_N) - \mathcal{R}(\hat{\theta}_N)) \bar{\theta} \leq 0$ . The proposition then follows by applying Lemma 2.  $\square$

**Appendix B. Kalman Yakubovitch Popov (KYP) Lemma**

Theorems 2 and 3 show that the robust performance constraint  $J(G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\hat{\theta}_N, P_\theta)$  can be treated by an infinite set of LMIs (i.e. the LMI condition (18) at each  $\omega$ ). In this appendix, we show in Proposition 4 that we can replace this infinite set of LMIs by a single LMI by removing the frequency dependence of condition (18) using the KYP Lemma. Before applying the KYP to this condition, note that condition (18) is not the unique way to express  $\bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\hat{\theta}_N, P_\theta)$  as an LMI linear in  $P_\theta^{-1}$ : in Bombois et al. (2001), we developed another LMI condition. That LMI condition involves the real part of a frequency-dependent matrix.

Note also that, in Jansson and Hjalmarsson (2004), the authors apply the real positive (RP) Lemma (a special case of the KYP Lemma) to a LMI condition of the type in Bombois et al. (2001). For this purpose, they need to multiply the LMI by the least common denominator of its entries. While entirely correct, this approach can lead to a final frequency-independent LMI which has an unnecessarily large state-space representation. Consequently, we have here developed the frequency-dependent condition (18) in such a way that the corresponding frequency-independent LMI obtained via the KYP Lemma has the lowest possible state-space representation. This frequency-independent LMI is given in the following proposition:

**Proposition 4.** Consider Proposition 3 and an arbitrary positive integer  $b$ . Define  $\mathcal{B}(z) = (1, z^{-1}, \dots, z^{-b})^T$ . Then, condition (18) holds for all  $\omega$  (or equivalently  $J(G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\hat{\theta}_N, P_0)$ ) if there exists  $P_f = P_f^T$  and  $P_\lambda = P_\lambda^T$  of appropriate dimensions, real numbers  $\lambda_i, i = 0 \dots b$  and  $L_p \in \mathbf{R}^{(k+1) \times (b+1)(k+1)}$  of the form

$$L_p = \begin{pmatrix} 0 & l_{12}^T & l_{13}^T & \dots & l_{1(k+1)}^T \\ -l_{12}^T & 0 & l_{23}^T & \dots & l_{2(k+1)}^T \\ \dots & \dots & \dots & \dots & \dots \\ -l_{1(k+1)}^T & -l_{2(k+1)}^T & -l_{3(k+1)}^T & \dots & 0 \end{pmatrix}$$

with  $l_{ij} \in \mathbf{R}^{(b+1) \times 1}$  such that

$$\begin{pmatrix} A_f^T P_f A_f - P_f & A_f^T P_f B_f \\ B_f^T P_f A_f & B_f^T P_f B_f \end{pmatrix} + \begin{pmatrix} C_f^T \\ D_f^T \end{pmatrix} \mathcal{X}_f(C_f \ D_f) \leq 0, \tag{B.1}$$

$$\begin{pmatrix} P_\lambda - A_\lambda^T P_\lambda A_\lambda & C_\lambda^T - A_\lambda^T P_\lambda B_\lambda \\ C_\lambda - B_\lambda^T P_\lambda A_\lambda & D_\lambda + D_\lambda^T - B_\lambda^T P_\lambda B_\lambda \end{pmatrix} \geq 0.$$

Here  $(A_\lambda, B_\lambda, C_\lambda, D_\lambda)$  is as  $(A, B, C, D)$  in Theorem 2 but with  $m$  and  $R_r(i)$  replaced by  $b$  and  $\lambda_i$ , respectively.  $(A_f, B_f, C_f, D_f)$  is a state-space representation of  $\mathcal{F}(z)$ :

$$\mathcal{F}(z) = \begin{pmatrix} (I_5 \otimes \mathcal{B})\Omega(z) \\ I_{k+1} \otimes \mathcal{B} \\ I_{k+1} \end{pmatrix} \tag{B.2}$$

and

$$\mathcal{X}_f = \begin{pmatrix} \begin{pmatrix} I_4 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} \lambda_0 & \dots & \lambda_b \\ \vdots & 0 & 0 \\ \lambda_b & 0 & 0 \end{pmatrix} & 0 & 0 \\ 0 & 0 & L_p^T \\ 0 & L_p & -\mathcal{R}(\hat{\theta}_N) \end{pmatrix}. \tag{B.3}$$

**Proof.** First note that, via some algebraic manipulations, constraints (18) can be rewritten as  $\mathcal{F}^*(e^{j\omega})\mathcal{X}_f\mathcal{F}(e^{j\omega}) \leq 0$  if one defines, in (18),  $\tau(\omega) = \lambda_0 + \sum_{i=1}^b \lambda_i (e^{j\omega} + e^{-j\omega})$  and  $j\mathcal{L}(\omega) = L(e^{j\omega}) + L^*(e^{j\omega})$  where  $L(z) = L_p(I_{k+1} \otimes \mathcal{B}(z))$  is a matrix of FIR transfer functions such that  $L(z) = -L(z)^T$ . Second, note that this expression of  $j\mathcal{L}(\omega)$  has the structure of Proposition 3 and that  $\tau(\omega)$  is ensured to be a positive scalar at each  $\omega$  via (B.1). Indeed,  $\tau(\omega) = \lambda_0 + \sum_{i=1}^b \lambda_i (e^{j\omega} + e^{-j\omega}) > 0 \forall \omega$

is ensured by the existence of a symmetric matrix  $P_\lambda$  such that (B.1) holds (see Lindqvist, 2001). In this context, the existence of  $\lambda_i (i = 0 \dots b)$  and of a matrix  $L_p$  such that  $\mathcal{F}^*(e^{j\omega})\mathcal{X}_f\mathcal{F}(e^{j\omega}) \leq 0$  implies that  $\bar{J}(\omega, G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\hat{\theta}_N, P_0)$ . The result of the proposition follows then from a simple application of the KYP Lemma (Popov, 1973).  $\square$

### References

Bombois, X., Gevers, M., Scorletti, G., & Anderson, B. D. O. (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. *Automatica*, 37(10), 1629–1636.

Bombois, X., & Gilson, M. (2006). Cheapest identification experiment with guaranteed accuracy in the presence of undermodeling. In *14th IFAC symposium on system identification* (pp. 505–510). Newcastle.

Bombois, X., Scorletti, G., Gevers, M., Hildebrand, R., & Van den Hof, P. M. J. (2004). Cheapest open-loop identification for control. *CD-ROM Proceedings of the 33rd IEEE conference on decision and control* (pp. 382–387). The Bahamas.

Bombois, X., Scorletti, G., Van den Hof, P. M. J., & Gevers, G. (2004). Least costly identification experiment for control: A solution based on a high-order model approximation. In *CD-ROM Proceedings of the American control conference*. Boston, MA, USA.

Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*. SIAM Studies in Applied Mathematics, Philadelphia, 1994.

Cooley, B., & Lee, J. (2001). Control-relevant experiment design for multivariable systems described by expansions in orthonormal bases. *Automatica*, 37(2), 273–281.

Ferreres, G., & Fromion, V. (1997).  $H_\infty$  control for a flexible transmission system. *CD-ROM Proceedings of the European control conference* Brussels, Belgium.

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

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

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

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

Jansson, H. (2004). *Experiment design with applications in identification for control*. Ph.D. thesis, Royal Institute of Technology, Stockholm, Sweden.

Jansson, H., & Hjalmarsson, H. (2004). Mixed  $H_\infty$  and  $H_2$  input design for identification. *CD-ROM Proceedings of the 43rd IEEE conference on decision and control* (pp. 388–393). The Bahamas.

Landau, I. D., Rey, D., Karimi, A., Voda, A., & Franco, A. (1995). A flexible transmission system as a benchmark for robust digital control. *European Journal of Control*, 1(2), 77–96.

Lindqvist, K. (2001). *On experiment design in identification of smooth linear systems*. Licentiate thesis, Royal Institute of Technology, Stockholm, Sweden.

Ljung, L. (1999). *System identification: Theory for the user*. (2nd ed.), Englewood Cliffs, NJ: Prentice-Hall.

Meinsma, G., Shrivastava, Y., & Fu, M. (1997). A dual formulation of mixed  $\mu$  and on the losslessness of  $(D, G)$  scaling. *IEEE Transactions on Automatic Control*, 42(7), 1032–1036.

Popov, V. M. (1973). *Hyperstability of control systems*. New York: Springer.

Rivera, D. E., Lee, H., Braun, M. W., & Mittelmann, H. D. (2003). Plant-friendly system identification: A challenge for the process industries. *Proceedings of the 13th IFAC symposium on system identification* (pp. 917–922). Rotterdam.

Söderström, T., & Stoica, P. (1989). *System identification*. Hemel Hempstead, Hertfordshire: Prentice-Hall International.

Tempo, R., Bai, E. W., & Dabbene, F. (1997). Probabilistic robustness analysis: Explicit bound for the minimum number of samples. *Systems and Control Letters*, 30, 237–242.

Tempo, R., Calafiore, G., & Dabbene, F. (2004). *Randomized algorithms for analysis and control of uncertain systems*. New York: Springer.

Van den Eijnde, E., & Schoukens, J. (1991). On the design of optimal excitation signals. *Proceedings of the ninth IFAC/FORS symposium on identification and system parameter estimation* (pp. 827–832). Budapest, Hungary.



**Xavier Bombois** was born in Brussels in 1974. He received the Electrical Engineering and Ph.D. degrees from the Université Catholique de Louvain (Belgium) in 1997 and 2000, respectively. Currently, he is an assistant professor at the Delft Center for Systems and Control (The Netherlands). His main interests are identification for control, prediction error identification.



**Gerard Scorletti** received the engineering degree from the Ecole Nationale Supérieure de Techniques Avancées (ENSTA), Paris, and obtained his Ph.D. in Automatic Control from Orsay University, France, in June 1997. His advisors were Professor Laurent El Ghaoui, Berkeley University and Professor Gilles Duc, Ecole Supérieure d'Electricité, Paris. He was a research associate at Centre for Systems Engineering and Applied Mechanics, Belgium (1997–1998) where he collaborates with Professor Michel Gevers and Professor Vincent Wertz.

Currently, he is a “maitre de conférence” (assistant professor) of Electrical Engineering at Caen University (France) and he holds a permanent joint appointment with the GREYC Equipe Auto. His research interests include robust (non-linear) control theory, gain scheduling control, identification for control, time-delay system control, structured controller design and linear matrix inequality applications with a special emphasis on CAD. He is Associate Editor to the IEEE Conference Editorial Board.



**Michel Gevers** was born in Antwerp, Belgium, in 1945. He obtained an Electrical Engineering degree from the Université Catholique de Louvain, Belgium, in 1968, and a Ph.D. degree from Stanford University, California, in 1972. He is a Fellow of the IEEE and a Distinguished Member of the IEEE Control Systems Society, and he holds a Honorary Degree (Doctor Honoris Causa) from the University of Brussels. He has been President of the European Union Control Association (EUCA) from 1997 to 1999, and Vice President of the IEEE Control Systems Society in 2000 and 2001.

Michel Gevers is Professor at the Department of Mathematical Engineering at the Université Catholique de Louvain in Louvain la Neuve, Belgium. He is the coordinator of the Belgian Interuniversity Network on Dynamical Systems and Control funded by the Federal Ministry of Science. This is a network of excellence in systems and control, which counts about 35 academics and 150 Ph.D. students and post-docs. He has spent long-term visits at the University of Newcastle, Australia, the Technical University of Vienna, and the Australian National University. His present research interests are in system identification and its interconnection with robust control design, including optimal experiment design for control.

Michel Gevers has been Associate Editor of *Automatica* and of the *IEEE Transactions on Automatic Control*. He is presently Associate Editor at Large of the *European Journal of Control* and Associate Editor of *Mathematics of Control, Signals, and Systems*. He has published about 230 papers and conference papers, and two books: “Adaptive Optimal Control—The Thinking Man’s GPC”, by R.R. Bitmead, M. Gevers and V. Wertz (Prentice-Hall, 1990), and “Parametrizations in Control, Estimation and Filtering Problems: Accuracy Aspects”, by M. Gevers and G. Li (Springer, 1993).



**Paul Van den Hof** was born in Maas-tricht, The Netherlands. He received the M.Sc. and Ph.D. degrees both from the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands in 1982 and 1989, respectively. From 1986 to 1999 he was an assistant and associate professor in the Mechanical Engineering Systems and Control Group of Delft University of Technology, The Netherlands. Since 1999 he has been a full professor in the Signals, Systems and Control Group of the Department of Applied Physics

at Delft University of Technology, and since 2003 he is Co-Director of the Delft Center for Systems and Control, with appointments in the Faculty of Mechanical Engineering and the Faculty of Applied Sciences. He has been the General Chair for the 13th IFAC Symposium on System Identification, that was held in Rotterdam, The Netherlands, in 2003. Since 2005 he is scientific director of the Dutch Institute of Systems and Control (DISC).

He has been a member of the IFAC Council (1999–2005), member of the Board of Governors of IEEE’s Control System Society (2003–2005), member of the administrative council of the European Union Control Association (EUCA), and *Automatica* Editor for Rapid Publications (1999–2005).

Paul Van den Hof’s research interests are in issues of system identification, parametrization, signal processing and (robust) control design, with applications in mechanical servo systems and industrial process control systems.



**Roland Hildebrand** was born in Leipzig, GDR, in 1975. He received the M.Sc. degree from the Free University of Berlin, Germany, in 1997 and the Ph.D. degree in mathematics from Moscow State University, Russian Federation, in 2001. From 2001 until 2003 he was postdoc at the Catholic University of Louvain, Louvain-la-Neuve, Belgium. Since 2003 he is with the Laboratory of Modeling and Calculus, Grenoble, France, where he is currently junior researcher.