LEAST DISTURBING CLOSED-LOOP IDENTIFICATION EXPERIMENT FOR CONTROL

X. Bombois * G. Scorletti** P. Van den Hof*

* Delft Center for Systems and Control, Delft University of Technology, The Netherlands
** GREYC, Equipe AUTO, ISMRA, Caen, France

Abstract: This paper presents results on closed-loop identification experiment design for control. The objective is to design the least costly closed-loop identification experiment such that the controller designed with the identified model stabilizes and achieves a prescribed level of $H_\infty$ performance with the unknown true system.

Copyright © 2005 IFAC.

Keywords: experiment design, prediction error identification for control

1. INTRODUCTION

In industrial practice, one often faces the situation that the true system $G_0$ is controlled using a controller $C_{id}$ which is not fully satisfactory and which one wants to replace by a new controller. In such a situation, the new controller $\hat{C}$ for $G_0$ is generally designed from a model $\hat{G}$ which has been identified using data collected on the existing closed loop. When designing the identification experiment leading to the model $\hat{G}$, the control engineer has to make a trade-off between her/his desire of obtaining an accurate model and the economical constraint of keeping the experimental costs low.

The typical approach to this problem has been to maximize the accuracy of the identified model (possibly with a given, say, control-oriented objective in mind) for a given experiment time and under prespecified constraints on excitation power (see e.g. (Ljung, 1999; Lindqvist, 2001; Hildebrand and Gevers, 2003) and references therein). Here, we address this tradeoff from the dual perspective; namely, we seek the least costly identification experiment leading to a required model accuracy, with a control-oriented objective in mind.

More precisely, assuming that the experiment time is given, we design a closed-loop identification experiment for which it is guaranteed that the identified model $\hat{G}$ delivers a controller $\hat{C}(\hat{G})$ which achieves sufficient $H_\infty$ performance with $G_0$. Furthermore, among all experiments achieving this performance constraint, we seek the one whose excitation signal $r(t)$ induces the smallest disturbance on the normal operation of the control loop $[C_{id} G_0]$. Here, the disturbance induced by $r(t)$ on the normal operation is measured by $J_r = \alpha_y P_{y}\alpha_u P_{u}$, where $P_{y}$ (resp. $P_{u}$) is the total power of the part of the output signal $y(t)$ (resp. input signal $u(t)$) due to $r(t)$, and $\alpha_y$, $\alpha_u$ are self-chosen scalars.

This experiment design problem is solved in the following context. We assume that the identification experiment is performed on the existing closed loop $[C_{id} G_0]$ with the direct closed-loop Prediction Error (PE) identification method (Ljung, 1999). We also make a classical assumption in experiment design i.e. that the considered model structure is full-order. We finally assume that the new controller $\hat{C}$ will be designed from the identified model $G(z) = G(z, \theta_N)$ using a predefined control design method (e.g. a $H_\infty$ control
design method with fixed weights). In this particular context, we also make use of the fact that, along with the model $\hat{G}$, an identification experiment delivers a parametric uncertainty region $\mathcal{D}$ centered on $\hat{G}$ and containing the true system $G_0$ at a self-chosen probability level (Bombois et al., 2001). Therefore, the condition that $\hat{C}(\hat{G})$ achieves sufficient $H_\infty$ performance with $G_0$ can be replaced by the condition that $\hat{C}(\hat{G})$ achieves sufficient $H_\infty$ performance with all systems in $\mathcal{D}$. Note that the size of the identified $\mathcal{D}$ is a function of the covariance matrix $P_\theta$ of the identified parameter vector $\hat{\theta}_N$ and, consequently, a function of the chosen power spectrum $\Phi_r(\omega)$ of $r(t)$.

In this context, we solve our experiment design problem using Linear Matrix Inequality (LMI) optimization (Boyd et al., 1994). For this purpose, we first develop an LMI constraint, linear in $P_\theta^{-1}$, which ensures that the controller $\tilde{C}$ designed with the identified model $\hat{G}$ achieves sufficient $H_\infty$ performance with all systems in the identified uncertainty region $\mathcal{D}$. The $H_\infty$ performance is measured using an usual matricial expression involving the four closed-loop transfer functions. Then, among all spectra $\Phi_r(\omega)$ that correspond to an identification experiment ensuring this robust performance constraint, we determine the particular $\Phi_r(\omega)$ which minimizes the cost function $J_r$. To achieve this constrained minimization, we use the finite parametrization of $\Phi_r(\omega)$ presented in (Lindqvist, 2001) and which has the property that both $P_\theta^{-1}$ and $J_r$ are affine in the parameters of $\Phi_r(\omega)$.

The use of the parametric set $\mathcal{D}$ that directly follows from the identification, instead of the additive uncertainty region that embeds $\mathcal{D}$ is the main contribution of the present paper relative to (Bombois et al., 2004). Note that the uncertainty region $\mathcal{D}$ is also used for an experiment design problem in (Hildebrand and Gevers, 2003; Jansson and Hjalmarsson, 2004). However, in those papers, the considered constraint pertains to a measure of the distance between $\hat{G}$ and the plants in $\mathcal{D}$ and not to the $H_\infty$ performance. Another contribution of this paper is to extend the concept of cheap experiment design for control that has been developed for open-loop identification in (Bombois et al., 2004), to direct closed-loop identification. Note that indirect closed-loop identification can be treated as a special case of open-loop identification with an adapted expression for the uncertainty region $\mathcal{D}$ containing $G_0$ (see (Bombois et al., 2001)).

2. PE IDENTIFICATION ASPECTS

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

$$y(t) = G(z, \theta_0) u(t) + H(z, \theta_0) e(t)$$

for some unknown parameter vector $\theta_0 \in \mathbb{R}^k$, and with $e(t)$ a white noise of variance $\sigma^2$. This true system is operated in closed loop with an initial (but not fully satisfactory) controller $C_{id}$ i.e. $u(t) = C_{id}(z) (r(t) - y(t))$ where $r(t)$ is the reference signal which is assumed to be zero in normal operation and which can be used to excite the system for a closed-loop identification. The closed-loop system can thus be written as:

$$y_r(t) = \frac{1}{T_{id}} y_r(t) + S_{id} \delta(t)$$

$$u(t) = C_{id} S_{id} r(t) - C_{id} S_{id} v(t)$$

with $T_{id} = (C_{id} G_0)/(1 + C_{id} G_0)$ and $S_{id} = 1/(1 + C_{id} G_0)$. The controlled output and input of the true system is thus made up of a part due to the signal $r(t)$ and one due to the noise $v(t)$.

The model $\hat{G}(z) = G(z, \hat{\theta}_N), \hat{H}(z) = H(z, \hat{\theta}_N)$ of the true system is obtained from $N$ collected data $y(t)$ and $u(t)$ ($t = 1...N$) generated by a signal $r(t)$ applied to the closed loop: $\hat{\theta}_N = \arg \min_{\theta} \frac{1}{N} \sum_{t=1}^{N} e^2(t, \theta)$ with $e(t, \theta) \triangleq H(z, \theta)^{-1} (y(t) - G(z, \theta) u(t))$.

In order to keep the cost of the closed-loop identification experiment low, the power spectrum $\Phi_r(\omega)$ of $r(t)$ should be designed in such a way that the difference between $y(t)$ and $u(t)$ in normal operation (i.e. with $r(t) = 0$) and those signals during the identification is as small as possible. This objective can e.g. be obtained by minimizing the powers $P_{y_r}$ and $P_{u_r}$ of the signals $y_r(t)$ and $u_r(t)$ in (2) i.e. by minimizing the cost function

$$J_r = \alpha_y \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{y_r}(\omega) \, d\omega \right) + \alpha_u \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_r}(\omega) \, d\omega \right)$$

with $\alpha_y, \alpha_u$ self-chosen scalars and $\Phi_{y_r}(\omega)$ (resp. $\Phi_{u_r}(\omega)$) the power spectrum of $y_r(t)$ (resp. $u_r(t)$).

The identified parameter vector $\hat{\theta}_N$ is asymptotically normally distributed, $\hat{\theta}_N \sim N(\theta_0, P_\theta)$ and, given the full-order model structure assumption, the covariance matrix $P_\theta$ has the following expression (Ljung, 1999): $P_\theta = \frac{1}{\pi} \left( \frac{1}{\pi} (E \left[ \psi(t, \theta_0) \psi(t, \theta_0)^T \right])^{-1} \right)$ with $\psi(t, \theta) = \frac{\partial \hat{y}(t, \theta)}{\partial \theta}$. The dependence of $P_\theta$ on the spectrum $\Phi_r(\omega)$ used during the identification is evidenced by the following expression of $P_\theta^{-1}$.
easily deduced from (2) and the general expression of $P_0$ above:

$$P_\theta^{-1} = \left( \frac{N}{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{F_z(e^{j\omega}, \theta_0)F_z(e^{j\omega}, \theta_0)^* \Phi_\theta(\omega) d\omega}{-\pi} \right)$$

$$+ \left( \frac{N}{2\pi} \int_{-\pi}^{\pi} \frac{F_z(e^{j\omega}, \theta_0)F_z(e^{j\omega}, \theta_0)^* d\omega}{-\pi} \right)$$

Here, $F_z(z, \theta_0) = C_{id}S_{id}A_z(z, \theta_0)$, $F_z(z, \theta_0) = \frac{A_H(z, \theta_0)}{H(z, \theta_0)} - H(z, \theta_0)F_z(z, \theta_0)$, $A_G(z, \theta) = \frac{\partial G(z, \theta)}{\partial \theta}$ and $A_H(z, \theta) = \frac{\partial H(z, \theta)}{\partial \theta}$.

Using the asymptotic Gaussian distribution of the estimated parameter vector $\hat{\theta}_N$, it is possible to define an uncertainty region $\mathcal{D}(\hat{\theta}_N, P_0)$ around the identified model and containing the unknown true system $G(z, \theta_0)$ at any self-chosen probability level (Bombois et al., 2001):

$$\mathcal{D}(\hat{\theta}_N, P_0) = \left\{ G(z, \theta) = \frac{Z_N(z, \theta)}{1 + Z_D(z, \theta)} \mid \theta \in U, \right\}

U = \left\{ \theta | \theta - \hat{\theta}_N \right\} T P_0^{-1} (\theta - \hat{\theta}_N) < \chi \right\}$$

where $\chi$ is a real constant dependent on the chosen probability level and $Z_N, Z_D$ are row vectors containing delays and zeros. The size of the uncertainty region $\mathcal{D}(\hat{\theta}_N, P_0)$ is a function of the covariance matrix $P_0$ and thus a function of the spectrum $\Phi_\theta(\omega)$ used during the identification experiment.

We will restrict attention to spectra $\Phi_\theta(\omega)$ that can be written as (Lindqvist, 2001):

$$\Phi_\theta(\omega) = R_\epsilon(0) + 2 \sum_{i=1}^{m} R_r(i) \cos(i\omega) \geq 0$$

where $m$ is a self-chosen scalar. The parameters $R_r(i)$ ($i = 0...m$) can be interpreted as the autocorrelation sequence of a signal that has been generated by a white noise passing through a FIR filter of length $m + 1$. An important property of such parametrization is that $\mathcal{J}$ and $P_0^{-1}$ are both affine functions of the design variables $R_r(i)$ ($i = 0...m$) as shown in the following two propositions (Lindqvist, 2001).

**Proposition 1.** Assume that $\Phi_\theta(\omega)$ is given by (4). Let $\hat{\theta}_N$ be the sequence of Markov parameters of $F_z, F_z^*$ i.e. $F_z(e^{j\omega}, \theta_0) = \sum_{k=-\infty}^{\infty} M_k(\theta_0)e^{-jkw}$ with $F_z(z, \theta_0)$ as defined in Section 2. Then, $P_0^{-1} \in \mathbb{R}^{k \times k}$ can be written as:

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

where $\hat{\theta}_N = \frac{N}{2\pi} \int_{-\pi}^{\pi} F_z(e^{j\omega}, \theta_0) F_z(e^{j\omega}, \theta_0)^* d\omega$ and $\mathcal{C}(\theta_0, \sigma_e^2) \in \mathbb{R}^{k \times k}$ are defined as $\mathcal{C}(\theta_0, \sigma_e^2) = \frac{N}{\pi} \hat{\theta}_N$ and $\mathcal{C}(\theta_0, \sigma_e^2) = \frac{N}{\pi^2} (\mathcal{C}(\theta_0) + \mathcal{C}^T(\theta_0)) (i = 1...m)$.

**Proposition 2.** Consider the cost function $\mathcal{J}_\theta$ and assume that $\Phi_\theta(\omega)$ is given by (4). Then, $\mathcal{J}_\theta$ can also be written as:

$$\mathcal{J}_\theta = \left( \sum_{i=m}^{\infty} \left( \alpha_i \epsilon_i(\theta_0) + \alpha_i d_i(\theta_0) \right) R_r(i) + ... \right)$$

where the coefficients $\epsilon_i(\theta_0)$ and $d_i(\theta_0)$ are the Markov parameters of $F_z, F_z^*$ and $C_{id}C_{id}^*S_{id}$ i.e. $T_e(e^{j\omega})T_i(e^{j\omega})^* = \sum_{k=-\infty}^{\infty} \epsilon_k(\theta_0)e^{-jk\omega}$ and $C_{id}C_{id}^*S_{id} = \sum_{k=-\infty}^{\infty} \delta_k e^{-jk\omega}$.

### 3. CONTROL DESIGN OBJECTIVES AND CONTROL DESIGN METHOD

As stated before, our aim is to replace the “unsatisfactory” controller $C_{id}$ 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 loop $[C G]$:

$$J(G, C, W_l, W_r) = \sup_{\omega} \mathcal{J}(\omega, G, C, W_l, W_r)$$

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

where $\sigma(A)$ denotes the largest singular value of $A$ and $W_l(z)$, $W_r(z)$ are given diagonal performance filters. The performance measure (5) 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 here in such a way that they reflect the performance specifications we want to achieve with the true system. Thus, a controller $\hat{C}$ will be deemed satisfactory if $[C G_0]$ is stable and if $J(G_0, \hat{C}, W_l, W_r) \leq 1$.

The new controller $\hat{C}$ that will replace $C_{id}$ will be designed using an identified model $\hat{G} = G(z, \theta_N)$ of $G_0$. In order to define the control design method leading to $\hat{C} = C(G(z, \theta_N))$, we make the following assumption (Bombois et al., 2004).

**Assumption 1.** We have pre-selected a fixed control design method which maps any reasonable model $G(z, \theta_0)$ to one controller
$C(G(z, \theta))$. The controller $C(G(z, \theta))$ stabilizes $G(z, \theta)$ and achieves with this model a performance level

$$J(G(z, \theta), C(G(z, \theta)), W_l(z), W_r(z)) \leq \gamma < 1,$$  

(6)

where $\gamma$ is a fixed scalar, strictly smaller than 1.

One design strategy that satisfies Assumption 1 is to choose $C(G(z, \theta))$ as the central controller of the four-block $H_\infty$ control design method with performance objective (6).

If Assumptions 1 holds, then the controller $\hat{C} = C(G(z, \theta_N))$ designed from an identified model $\hat{G} = G(z, \theta_N)$ will achieve $J(\hat{G}, \hat{C}, W_l, W_r) \leq \gamma < 1$. When this controller $\hat{C}$ is applied to the true system $G_0$, the achieved performance could be poorer than the designed performance. By choosing the design criterion (6) with $\gamma < 1$, we ensure, however, that there is a whole set of systems $G(z, \theta)$ around the to-be-identified $G(z, \theta_N)$, that are also stabilized by $\hat{C}$ and that achieve $J(G, \hat{C}, W_l, W_r) \leq 1$.

4. IDENTIFICATION FOR CONTROL AT THE CHEAPEST COST

As stated in the previous sections, our objective is to design an identification experiment on the loop $[C_{id} G_0]$ 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 $G_0$ lies in the uncertainty region $D(\theta_N, P_0)$ identified along with $\hat{G}$, this performance constraint can be replaced by the constraint 1 that $\hat{C}$ achieves $J(G, \hat{C}, W_l, W_r) \leq 1$ with all $G(z) \in D(\theta_N, P_0)$. Furthermore, among all closed-loop identification experiments satisfying this performance constraint, we seek the one where $[C_{id} G_0]$ is excited with the signal $r(t)$ of spectrum $\Phi_r(e^{j\omega})$ which minimizes the cost function $J_r$. By parametrizing (4) for $\Phi_r(\omega)$, our experiment design problem can thus be formulated as:

**Cheapest experiment design problem for control:** For a direct closed-loop identification on $[C_{id} G_0]$ with $N$ sensors, determine the parameters $R_i (i = 0, \ldots, m)$ of the spectrum $\Phi_r(\omega)$ which minimize under the constraint that $\Phi_r(\omega) \geq 0 \forall \omega$ and that $J(G, \hat{C}, W_l, W_r) \leq 1$ with all $G(z) \in D(\theta_N, P_0)$. $\hat{C}$ is the controller designed with the identified model $\hat{G}$ using the control design method of Assumption 1 and $D(\theta_N, P_0)$ is the identified uncertainty region (see (3)).

We show in the sequel that this problem can be expressed as an LMI-based optimization problem (Boyd et al., 1994). For this purpose, we first express the robust performance constraint $J(\omega, \hat{G}, \hat{C}, W_l, W_r) \leq 1 \forall G \in D(\theta_N, P_0)$ at one particular frequency $\omega$ as an LMI, linear in $P_0^{-1}$. Note that $J(G, C, W_l, W_r) \leq 1 \forall G \in D(\theta_N, P_0)$ $\iff$ at each $\omega$, $J(\omega, \hat{G}, \hat{C}, W_l, W_r) \leq 1 \forall G \in D(\theta_N, P_0)$.

**Proposition 3.** Consider $\hat{C} = C(G(z, \hat{\theta}_N))$, the controller designed from the model $G(z, \hat{\theta}_N)$ using the method of Assumption 1. Consider also the set $D(\hat{\theta}_N, P_0)$ defined in (3). Then, $\hat{C}$ achieves $J(\omega, \hat{G}, \hat{C}, W_l, W_r) \leq 1$ with all $G \in D(\hat{\theta}_N, P_0)$ if and only if $\exists \tau(\omega) > 0$ and a vector $l(\omega) \in \mathbb{R}^{k+1}$ such that

$$\tau(\omega)E(\omega, \hat{\theta}_N) - P(\hat{\theta}_N) - \mathcal{L}(\omega) \leq 0$$

(7)

with $P(\hat{\theta}_N) = \left( \begin{array}{cc} I_k & \theta_N \\ -\theta_N & -\theta_N \end{array} \right) - \left( \begin{array}{cc} 0 & 0 \\ 0 & -\chi \end{array} \right) \left( \begin{array}{cc} \theta_N & 0 \\ 0 & -1 \end{array} \right) \Omega(e^{j\omega})$

and

$$\mathcal{L}(\omega) = \left( \begin{array}{cc} 0 & -l(\omega)^T \\ -l(\omega) & 0 \end{array} \right) \text{ and } j \pm \sqrt{-1} \tau(\omega) \in \mathcal{L}(\omega).$$

The symbol $\otimes$ denotes the Kronecker product.

**Proof.** In (Bombois et al., 2001), it is shown that, for a system $G \in D(\hat{\theta}_N, P_0)$, $J(\omega, G, \hat{C}, W_l, W_r) \leq 1$ can be written as:

$$g^*(e^{j\omega}) \left( \begin{array}{cc} x^T & 0 \\ 0 & -1 \end{array} \right) g(e^{j\omega}) \leq 0$$

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

$$g(z) = \left( \begin{array}{cc} W_l & 0 \\ 0 & 1 \end{array} \right) \left( \begin{array}{cc} Z_N \theta \\ 1 + Z_D \theta \end{array} \right).$$

If we define $\bar{\theta} = \theta^T 1^T$, the expression above can be rewritten as $\bar{\theta}^T \mathcal{L}(\bar{\theta}) \mathcal{L}(\bar{\theta}) \leq 0$. On the other hand, for any $k(\omega) > 0$ and any $l(\omega)$ as defined above, $\bar{\theta}^T k(\omega) \mathcal{L}(\bar{\theta}) \leq 0$. Thus, the constraint $\bar{\theta} \in U$ (see (3)) can be rewritten as $\bar{\theta}^T (k(\omega)(P(\hat{\theta}_N) + \mathcal{L}(\omega))) \bar{\theta} \leq 0$. Based on these two quadratic constraints in $\bar{\theta}$, it is obvious that the condition stated in the proposition with $\tau(\omega) \Delta 1/k(\omega)$ implies that $J(\omega, G, \hat{C}, W_l, W_r) \leq 1 \forall G \in D(\hat{\theta}_N, P_0)$. The converse implication is also valid, but the proof is quite technical.
Based on Propositions 1, 2 and 3, the experiment design problem described above would be solvable if the parametrizations of $P_{f_0}$ and $J_f$ w.r.t. the designs variables $R_r(i)$ were not functions of $\theta_0$ and $\sigma^2$, and if condition (7) was not a function of the to-be-identified $\hat{\theta}_N$. Such difficulty is classically circumvented by using a-priori estimates for those quantities: $\theta_{o,est}$, $\sigma^2_{e,est}$ and $\hat{\theta}_{N,est}$. The problem can then be solved using the following LMI optimization problem where we use the shorthand notations: $c_i = c(\theta_{o,est})$, $d_i = d(\theta_{o,est})$, $M = M(\theta_{o,est})$, $M_i = M_i(\theta_{o,est},\sigma^2_{e,est})$ and $E(\omega) = E(\omega, \hat{\theta}_{N,est})$. Note that we often choose $\theta_{o,est} = \hat{\theta}_{N,est}$.

Theorem 1. Consider the approximations $\theta_0 \approx \theta_{o,est}$, $\hat{\theta}_N \approx \theta_{N,est}$ and $\sigma^2 \approx \sigma^2_{e,est}$ and the shorthand notations above. Then, the auto-correlation sequence $R_r(i)$ (i = 0...m) which solves the cheapest experiment design problem for control is the solution of the following LMI optimization problem:

$$\min_{R_r(i)(i=0...m)} \sum_{i=1}^{m} \left( \alpha_i c_i + \alpha_i d_i \right) R_r(i) + 2 \sum_{i=1}^{m} \left( \alpha_i c_i + \alpha_i d_i \right) R_r(i)$$

under the constraint that $\exists$ a symmetric matrix $Q$ of appropriate dimension, a frequency function $\tau(\omega)$ valued in $R$ and a frequency function $l(\omega)$ valued in $R^{2\times1}$ such that

$$\tau(\omega) E(e^{j\omega}) = - \left( \begin{array}{c} I_k \\ -\theta_{N,est} \end{array} \right) \left( \begin{array}{c} \Sigma \sum_{i=1}^{M} M_i R_r(i) \\ 0 \end{array} \right) \left( \begin{array}{c} I_k \\ -\theta_{N,est} \end{array} \right)^T + \left( \begin{array}{c} 0 \\ 0 \end{array} \right) - L(\omega) \leq 0 \quad \forall \omega$$

and that

$$\begin{pmatrix} Q - A^TQA & C^T - A^TQB \\ C - B^TQA & D + D^T - B^TQB \end{pmatrix} \geq 0$$

with $L(\omega)$ as in Proposition 3 and

$$A = \left( \begin{array}{c} 0 \\ I_{m-1} \end{array} \right), \quad B = \left( \begin{array}{c} 1 \\ 0 \end{array} \cdots \begin{array}{c} 0 \end{array} \right), \quad C = \left( R_r(1) \ R_r(2) \ldots \ R_r(m) \right), \quad D = \frac{R_r(0)}{2}$$

The optimal spectrum $\Phi_r(\omega)$ can thereafter be computed using (4).

Proof. According to (Lindqvist, 2001), $\Phi_r(\omega) \geq 0 \quad \forall \omega$ is equivalent to the fact that there exists a symmetric matrix $Q$ such that the last constraint above holds. Consequently, the theorem is a direct consequence of Propositions 1, 2 and 3. □

Comment 1. Condition (8) has to be considered at every frequency. This is impossible in practice. The optimal $\Phi_r(\omega)$ can nevertheless be approximated by using a finite frequency grid. An exact, but more cumbersome solution consists of using the Kalman-Yacubovitch-Popov (KYP) lemma (Popov, 1973); see Appendix A.

Comment 2. Our experiment design problem delivers, for a given length $N$ of the identification, the spectrum $\Phi_r(\omega)$ for the identification in such a way that the part of the input-output signal due to $\Phi_r(\omega)$ has the smallest possible total power. This optimization can be achieved for different values of the length $N$ of the identification in order to choose the “optimal” combination for the length of the identification and the induced disturbance. In order to understand the trade-off that can be achieved, let us perform the following analysis. Roughly speaking, the optimal $\Phi_r(\omega)$ of Theorem 1 is the “smallest” $\Phi_r(\omega)$ which makes $P_{G}^{-1} > R_{adm}$ where $P_{G}$ is the part of $P_{0}$ corresponding to the parameters in $G(z, \theta)$ and $R_{adm}$ is a fixed matrix depending on the desired robust performance level. Consider, for simplicity, the case where $H(\theta) = 1$ (i.e. $M=OE$). Then, we obtain:

$$P_{G}^{-1} = N \left( I(F_r, \Phi_r(\omega)) + I(F_r, \sigma^2) \right)$$

with

$$I(V(z), \Phi(\omega)) = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} V(e^{j\omega})V^*(e^{j\omega})\Phi(\omega) d\omega.$$

Consequently, we can conclude that, for increasing values of $N$, the spectrum $\Phi_r(\omega)$ required to make $P_{G}^{-1} > R_{adm}$ becomes, at each $\omega$, smaller. Moreover, for $N$ larger than a given value, the spectrum $\Phi_r(\omega)$ needed to make $P_{G}^{-1} > R_{adm}$ is $= 0$. This happens when $N I(F_r, \sigma^2) > R_{adm}$. Consequently, the identification experiment for robust control can be achieved using a set of normal operation signals $^2$ $(0 < \omega < \infty)$ where $(\omega(1)) = 0$. The minimal value of $N$ for this purpose can be easily determined using a similar LMI problem as the one in Theorem 1 (Bombois et al., 2005).

Finally, let us compare the open-loop and closed-loop identification cases for $M = OE$. Using the definition of $F_r$ in Section 2 and $\Phi_u(\omega) = [C_{id}S_{id}]^{2} \Phi_r(\omega)$, the closed-loop $P_{G}^{-1}$ can be rewritten as $P_{G}^{-1} = N \left( I(A_G, \Phi_u(\omega)) + I(F_r, \sigma^2) \right)$ while, in open-loop identification, we have $P_{G}^{-1} = N \left( I(A_G, \Phi_u(\omega)) \right)$. Thus, for a fixed $N$, the spectrum $\Phi_u(\omega)$ required to make $P_{G}^{-1} > R_{adm}$ is smaller than the spectrum $\Phi_r(\omega)$ that we need to apply to $G_0$ in open loop to make $P_{G}^{-1} > R_{adm}$. This is due to the extra term $N I(F_r, \sigma^2)$ in the closed-loop identification. This result can be easily extended to BJ model structures.

Comment 3. Theorem 1 uses a-priori estimates of $\theta_0$, $\hat{\theta}_N$ and $\sigma^2$ to determine the spectrum $\Phi_r(\omega)$

$^2$ This result also holds for other types of $M$. Note nevertheless that a direct closed-loop identification with $\Phi_r(\omega) = 0$ can then only be considered when the controller $C_{id}$ is sufficiently complex (Bombois et al., 2005).
Theorem 1 shows that the robust performance constraint $J(G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\theta_N, P_b)$ can be treated by an infinite set of LMI’s (i.e. the LMI condition (7) at each $\omega$). In this appendix, we show in Proposition 4 that we can replace this infinite set of LMI’s by a single LMI by removing the frequency dependence of condition (7) using the KYP lemma. Before applying the KYP to this condition, note that condition (7) is not the unique way to express $J(\omega, G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\theta_N, P_b)$ 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. Such an approach is correct, but can lead to a final frequency-independent LMI which has an unnecessarily large order. Consequently, we have here developed the frequency-dependent condition (7) in such a way that the corresponding frequency-independent LMI obtained via the KYP lemma has the lowest possible order. This frequency-independent LMI is given in the following proposition whose proof can be found in (Bombois et al., 2005).

Proposition 4. Consider Proposition 3 and a self-chosen scalar $b$. Define $B(z) = (1, z^{-1}, \ldots, z^{-b})^T$. Then, condition (7) holds for all $\omega$ (or equivalently $J(G, \hat{C}, W_l, W_r) \leq 1 \forall G \in \mathcal{D}(\theta_N, P_b)$) if

$$
\begin{align*}
&\exists P_f = P_f^T \text{ and } P_b = P_b^T \\
&\exists l_q \in \mathbb{R}^{(b+1)\times 1} \text{ q = 1...k} \\
&\exists \text{ scalars } \lambda_i, i = 0...b
\end{align*}
$$

such that

$$
\begin{bmatrix}
\mathcal{A}_f^T P_b A_f - P_f & \mathcal{A}_f^T P_f B_f \\
B_f^T P_f A_f & \mathcal{B}_f^T P_f B_f
\end{bmatrix}
+ \begin{bmatrix}
\mathcal{C}_f^T D_f \mathcal{D}_f
\end{bmatrix}
\leq 0
$$

$$
\begin{bmatrix}
P_a - \mathcal{A}_f^T Q_a A_f & P_a \\
C_a - B_a^T P_a A_f & D_a + P_a - B_a^T P_a B_a
\end{bmatrix}
\geq 0
$$

Here $(A_L, B_L, C_L, D_L)$ is as $(A, B, C, D)$ in Theorem 1 but with $M$ and $R_b(i)$ replaced by $b$ and $\lambda_i$, respectively. $(A_f, B_f, C_f, D_f)$ is the state-space representation of $F(z)$:

$$
F(z) = \begin{pmatrix}
I_2 \otimes B(z) \otimes (z) \\
L_{k+1} \otimes B(z)
\end{pmatrix}
$$

$$
X_f = \begin{pmatrix}
L_2 & 0 & 0 & 0 \\
0 & L_3 & 0 & 0 \\
0 & 0 & L_4 & 0 \\
0 & 0 & 0 & L_5
\end{pmatrix}
$$

$$
\Lambda = \begin{pmatrix}
\lambda_0 & \lambda_1 & \ldots & \lambda_b \\
\lambda_1 & \lambda_0 & \ldots & 0 \\
\lambda_2 & \lambda_0 & \ldots & 0 \\
\lambda_b & \lambda_0 & \ldots & 0
\end{pmatrix},
L_1 = \begin{pmatrix}
\frac{i^T}{\varphi} \\
\frac{i^T}{\psi}
\end{pmatrix},
L_2 = \begin{pmatrix}
\frac{i^T}{\varphi} \\
\frac{i^T}{\psi}
\end{pmatrix}.