Cheapest open-loop identification for control

X. Bombois[(1)], G. Scorletti[(2)], M. Gevers[(3)], R. Hildebrand[(4)], P. Van den Hof[(1)]

[(1)] Delft Center for Systems and Control, Delft University of Technology, The Netherlands

[(2)] GREYC, Equipe AUTO, ISMRA, Caen, France

[(3)] CESAME, Université Catholique de Louvain, Belgium

[(4)] LMC, Université Joseph Fourier, Grenoble, France

***Abstract*— This paper presents a new method of identification experiment design for control. Our objective is to design the open-loop identification experiment with minimal excitation such that the controller designed with the identified model stabilizes and achieves a prescribed level of H_∞ performance with the unknown true system G_0.**

I. Introduction

A controller for a real-life system G_0 is usually designed on the basis of a model $\hat{G}$ of G_0 identified using data collected from the true system. When designing the identification experiment, the control engineer often 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. Obtaining an accurate model requires a long identification experiment and a powerful input signal, while keeping the experimental costs low corresponds to a short experiment time and the excitation of G_0 with a low power signal.

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 input power (see e.g. [10], [9], [7] and references therein). In this paper, 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, we assume that the experiment time is fixed, and we then define the least costly identification experiment for control as the experiment on G_0 whose input signal power $\mathcal{P}_u$ is minimized under the constraint that the controller $\hat{C}$ designed from the identified model $\hat{G}$ is guaranteed to stabilize and to achieve sufficient performance with the unknown true G_0. In this paper, the desired performance on G_0 is expressed by magnitude bounds on one (or several) closed-loop transfer functions of $[\hat{C}\ G_0]$ (H_∞ performance constraints).

This experiment design problem is solved in the following context. We assume that the identification experiment is performed in open loop using Prediction Error identification, and with a model structure $G(z,\theta)$ to which the true G_0 belongs [10]. This yields a model $\hat{G} = G(z,\theta_N)$ and an uncertainty region, centered on $\hat{G}$, containing the true G_0 at a user-chosen probability level. We use an additive description of this uncertainty region, whose size r_u is a function of the input signal. In order to highlight this dependence, the uncertainty set is denoted $\mathcal{D}_{r_u}(\hat{\theta}_N)$. It can be estimated from the data. Finally, the controller $\hat{C}$ to be applied to the true system is designed from $\hat{G}$ using a pre-defined H_∞ control design method with fixed weights.

We propose the following two-step methodology to solve the experiment design problem with the cheapest experimental cost. In a first step, we determine the size $r_{adm}(\omega)$ of the largest additive uncertainty region that we can a-priori tolerate around the to-be-identified model $\hat{G}$ for the controller $\hat{C} = C(\hat{G})$ to achieve the required H_∞ performance level with all systems in this uncertainty region. In a second (identification design) step, we then deduce the least powerful quasi-stationary input signal $u(t)$ such that the size $r_u(\omega)$ of the identified uncertainty region $\mathcal{D}_{r_u}(\hat{\theta}_N)$ is at each frequency smaller than the largest admissible uncertainty radius $r_{adm}(\omega)$. The solution of the second step is based on results of [12] which show that, for each quasi-stationary input signal $u(t)$, one can define a finite-sized vector x_u of moments of the input power spectrum $\Phi_u(\omega)$ weighted with a special weight depending on the true system G_0, with the property that both the inverse P_θ^{-1} of the covariance matrix of the parameter vector identified with $u(t)$ and the power $\mathcal{P}_u$ of $u(t)$ are affine functions of x_u. We show that the optimization of the power $\mathcal{P}_u$ of $u(t)$ under the constraint $r_u(\omega) \leq r_{adm}(\omega) \quad \forall\omega$ can therefore be reduced to a tractable LMI[1] optimization problem on the finite-sized vector of moments x_u. A quasi-stationary input signal $u(t)$ can easily be constructed from the optimal moment vector x_u.

II. Prediction Error Identification aspects

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

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

[1] Linear Matrix Inequality.

the true system. Thus, the true system is given by:

$$y(t) = G(z,\theta_0)u(t) + H(z,\theta_0)e(t) \quad (1)$$

for some unknown parameter vector $\theta_0 \in \mathbf{R}^k$, and with $e(t)$ a white noise of variance σ_e^2.

A model $\hat{G} = G(z,\hat{\theta}_N), \hat{H} = H(z,\hat{\theta}_N)$ of the true system is obtained from N input-output data $y(t)$ and $u(t)$ ($t = 1...N$), using a Prediction Error criterion: $\hat{\theta}_N \triangleq \arg\ \min_\theta \frac{1}{N}\sum_{t=1}^{N} \epsilon^2(t,\theta)$ with $\epsilon(t,\theta) \triangleq H(z,\theta)^{-1}\left(y(t) - G(z,\theta)u(t)\right)$.

The cost of the identification experiment is determined by the total power $\mathcal{P}_u$ of the chosen input signal $u(t)$:

$$\mathcal{P}_u = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_u(\omega)d\omega \quad (2)$$

where $\Phi_u(\omega)$ is the power spectrum of the input signal $u(t)$, assumed to be quasistationary. It is this power $\mathcal{P}_u$ that we shall seek to minimize.

The identified parameter vector $\hat{\theta}_N$ is asymptotically normally distributed, $\hat{\theta}_N \sim \mathcal{N}(\theta_0, P_\theta)$ and, given the full-order model structure assumption, the covariance matrix P_θ has the following expression [10]: $P_\theta = \frac{\sigma_e^2}{N}\left(\bar{E}\left(\psi(t,\theta_0)\psi(t,\theta_0)^T\right)\right)^{-1}$ with $\psi(t,\theta) = -\frac{\partial\epsilon(t,\theta)}{\partial\theta}$. The dependence of the covariance matrix P_θ on the power spectrum of the selected input signal $u(t)$ is evidenced by the following expression of the inverse of P_θ [10]:

$$\begin{aligned} P_\theta^{-1} &= \left(\frac{N}{\sigma_e^2}\frac{1}{2\pi}\int_{-\pi}^{\pi} F_u(e^{j\omega},\theta_0)F_u(e^{j\omega},\theta_0)^*\Phi_u(\omega)d\omega\right) \\ &+ \left(N\frac{1}{2\pi}\int_{-\pi}^{\pi} F_e(e^{j\omega},\theta_0)F_e(e^{j\omega},\theta_0)^* d\omega\right) \end{aligned} \quad (3)$$

Here, $F_u(z,\theta_0) = \frac{\Lambda_G(z,\theta_0)}{H(z,\theta_0)}$, $F_e(z,\theta_0) = \frac{\Lambda_H(z,\theta_0)}{H(z,\theta_0)}$, $\Lambda_G(z,\theta) = \frac{\partial G(z,\theta)}{\partial\theta}$ and $\Lambda_H(z,\theta) = \frac{\partial H(z,\theta)}{\partial\theta}$.

By factoring $F_u(z,\theta_0)$ as $\frac{N_{F_u}(z,\theta_0)}{d_{F_u}(z,\theta_0)}$ where $d_{F_u}(z,\theta_0)$ is the least common polynomial denominator, one can decompose $F_u(z,\theta_0)F_u(z,\theta_0)^*$ as follows:

$$\begin{aligned} F_u(z,\theta_0)F_u(z,\theta_0)^* &= \frac{N_{F_u}(z,\theta_0)N_{F_u}^*(z,\theta_0)}{|d_{F_u}(z,\theta_0)|^2} \\ &= \frac{1}{|d_{F_u}(z,\theta_0)|^2}\sum_{i=-n}^{n}\tilde{M}_i(\theta_0)z^i \end{aligned} \quad (4)$$

where the matrices $\tilde{M}_i(\theta_0) \in \mathbf{R}^{k\times k}$ ($i = 0...n$) satisfy $\tilde{M}_i(\theta_0) = \tilde{M}_{(-i)}(\theta_0)^T$. We now introduce the moment vector $x_u(\theta_0) \in \mathbf{R}^{n+1}$ of the input signal with respect to the true system.

Definition 2.1: Consider the true system (1), the transfer vector $F_u(z,\theta_0)$ as below (3), the degree n of the decomposition (4), and an input signal $u(t)$ with power spectrum $\Phi_u(\omega)$. Then, the moment vector $x_u(\theta_0) \triangleq \left(\ x_0(\theta_0)\ x_1(\theta_0)\ ...\ x_n(\theta_0)\ \right)^T$ of $u(t)$ is a vector in $\mathbf{R}^{n+1}$ whose elements are defined as:

$$x_i(\theta_0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\Phi_u(\omega)}{|d_{F_u}(e^{j\omega},\theta_0)|^2}\ cos(i\omega)\ d\omega \quad (i = 0...n) \quad (5)$$

One can write a compact expression of both P_θ^{-1} and $\mathcal{P}_u$ as an affine function of the elements of $x_u(\theta_0)$.

Proposition 2.1: [12] Consider an identification experiment performed on (1) using a quasi-stationary input signal $u(t)$. Then the inverse of the covariance matrix $P_\theta \in \mathbf{R}^{k\times k}$ of the estimated parameter vector can be written as:

$$P_\theta^{-1} = \bar{M}(\theta_0) + \sum_{i=0}^{n} M_i(\theta_0,\sigma_e^2)\ x_i(\theta_0) \quad (6)$$

where $\bar{M}(\theta_0) = N\frac{1}{2\pi}\int_{-\pi}^{\pi} F_e(e^{j\omega},\theta_0)F_e(e^{j\omega},\theta_0)^* d\omega$ and $M_i(\theta_0,\sigma_e^2) \in \mathbf{R}^{k\times k}$ ($i = 0...n$) are defined using (4) as $M_0(\theta_0,\sigma_e^2) = \frac{N}{\sigma_e^2}\tilde{M}_0(\theta_0)$ and $M_i(\theta_0,\sigma_e^2) = \frac{N}{\sigma_e^2}\left(\tilde{M}_i(\theta_0) + \tilde{M}_i^T(\theta_0)\right)$ ($i = 1...n$).

Proposition 2.2: [12] Consider an input signal $u(t)$ to the true system (1) with power spectrum $\Phi_u(\omega)$. Then the total power $\mathcal{P}_u$ of $u(t)$ is a linear function of the elements $x_i(\theta_0)$ of $x_u(\theta_0)$:

$$\mathcal{P}_u = \sum_{i=0}^{n} c_i(\theta_0)x_i(\theta_0), \quad (7)$$

where the coefficients $c_i(\theta_0)$ are defined from the polynomial $d_{F_u}(e^{j\omega},\theta_0)$ as follows:

$$|d_{F_u}(e^{j\omega},\theta_0)|^2 = c_0(\theta_0) + \sum_{i=1}^{n}\left(c_i(\theta_0)cos(i\omega)\right) \quad (8)$$

Notice that the moment vector $x_u(\theta_0)$ and the parametrizations of $\mathcal{P}_u$ and P_θ with respect to this moment vector are functions of the unknown true system (via θ_0 and σ_e^2). Note also that, even though $\mathcal{P}_u$ itself is not a function of the true system, the *parametrization* of $\mathcal{P}_u$ in (7) is a function of the true system due to the dependence of $x_u(\theta_0)$ and of the coefficients $c_i(\theta_0)$ on the true system.

Using the asymptotic Gaussian distribution of the estimated parameter vector $\hat{\theta}_N$, it is possible to define an (additive) uncertainty region $\mathcal{D}_{r_u}(\hat{\theta}_N)$ around the identified model and containing the unknown true system $G_0(z)$ at any self-chosen probability level:

$$\begin{aligned} \mathcal{D}_{r_u}(\hat{\theta}_N) = \{&G(z) \in H_\infty \mid \\ &\left|G(e^{j\omega}) - G(e^{j\omega},\hat{\theta}_N)\right| < r_u(\omega)\ \ \forall\omega\Big\} \end{aligned} \quad (9)$$

Consider the following first order approximation of $G(z,\theta_0)$: $G(z,\theta_0) \approx G(z,\hat{\theta}_N) + \Lambda_G^T(z,\theta_0)(\theta_0 - \hat{\theta}_N)$ with

$\Lambda_G(z,\theta)$ as defined below (3). Using this approximation, the size $r_u(\omega)$ of $\mathcal{D}_{r_u}(\hat{\theta}_N)$ can then be written as :

$$r_u(\omega) = \sqrt{\chi \lambda_1 \left(T(e^{j\omega},\theta_0) P_\theta T(e^{j\omega},\theta_0)^T\right)} \qquad (10)$$

where χ is a real constant dependent on the chosen probability level, $T(e^{j\omega},\theta_0) \triangleq \begin{pmatrix} Re(\Lambda_G^T(e^{j\omega},\theta_0)) \\ Im(\Lambda_G^T(e^{j\omega},\theta_0)) \end{pmatrix} \in \mathbf{R}^{2\times k}$ and $\lambda_1(A)$ denotes the largest eigenvalue of A. The size $r_u(\omega)$ of the uncertainty region containing G_0 at a given probability level is a function of the covariance matrix P_θ and thus, by (3), a function of the input signal $u(t)$ used during the identification experiment. Typically, the larger the power $\mathcal{P}_u$ of $u(t)$, the smaller $r_u(\omega)$. Note also that $r_u(\omega)$ depends on θ_0.

III. Control design objectives and method

As stated before, our aim is to design a "satisfactory" controller $\hat{C}(z)$ for the unknown true system G_0 using an identified model $\hat{G} = G(z,\hat{\theta}_N)$ of G_0. A satisfactory controller must stabilize and achieve sufficient performance with G_0. In this section, we define the concept of *sufficient performance*, as well as the control design method we use to design $\hat{C}$ from the identified model. We adopt the following performance measure for a loop $[C\ G]$:

$$J(G,C,W_l,W_r) = \|W_l \overbrace{\begin{pmatrix} \frac{1}{1+GC} & \frac{G}{1+GC} \\ \frac{C}{1+GC} & \frac{GC}{1+GC} \end{pmatrix}}^{F(G,C)} W_r\|_\infty, \qquad (11)$$

where $W_l(z)$ and $W_r(z)$ are given diagonal performance filters. This performance measure 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_∞ norm smaller than one. Simpler H_∞ criteria can be chosen as special cases; e.g. $W_l(z) = diag(W(z),0)$ and $W_r = diag(1,0)$, $J(G,C,W_l,W_r) \leq 1$ corresponds to $\|W/(1+CG)\|_\infty \leq 1$.

We build performance filters $W_l(z)$ and $W_r(z)$ that reflect the performance specifications we want to achieve with the true system. Thus, the controller $\hat{C}$ will be deemed *satisfactory* if $J(G_0,\hat{C},W_l,W_r) \leq 1$. Here, as already mentioned, the controller $\hat{C}$ will be designed from the identified model $\hat{G} = G(z,\hat{\theta}_N)$. In order to define the control design method leading to $\hat{C} = C(G(z,\hat{\theta}_N))$, we make the following assumption.

Assumption 3.1: We have a-priori defined a set Θ of parameter vectors which we assume to contain any parameter vector $\hat{\theta}_N$ that would result from an identification under reasonable experimental conditions. We assume also that we have pre-selected a fixed control design method which maps any model $G(z,\theta)$ for $\theta \in \Theta$ to one controller $C(G(z,\theta))$. For each $\theta \in \Theta$, $C(G(z,\theta))$ stabilizes $G(z,\theta)$ and achieves with this model a performance level

$$J(G(z,\theta), C(G(z,\theta)), W_l(z), W_r(z)) \leq \gamma < 1, \qquad (12)$$

where γ is a fixed scalar, strictly smaller than 1.

One design strategy that satisfies Assumption 3.1 is to choose $C(G(z,\theta))$ as the central controller of a four-block H_∞ control design method with performance objective (12) (see, in this aspect, the recent results in [4]). If Assumption 3.1 holds, then the controller $\hat{C} = C(G(z,\hat{\theta}_N))$ designed from an identified model $\hat{G} = G(z,\hat{\theta}_N)$ 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 will in most cases be poorer than the designed performance. By choosing the design criterion (12) with $\gamma < 1$, we ensure, however, that there is a whole (additive) set of systems $G(z,\theta)$ around the to-be-identified $G(z,\hat{\theta}_N)$, characterized by a size $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$ (see next subsection), that are also stabilized by $\hat{C}$ and that achieve $J(G(z,\theta),\hat{C}(z),W_l(z),W_r(z)) \leq 1$. Before we address this input design problem, we need to properly define the largest set of systems that are stabilized and achieve the required performance with a model-based controller. We do this in the next subsection.

IV. The largest uncertainty radius $r_{adm}(\omega)$

Let us consider one model $G(z,\theta)$ for some $\theta \in \Theta$ and the controller $C(G(z,\theta))$ designed with $G(z,\theta)$ using the design rule mentioned in Assumption 3.1. For any positive function $r(\omega)$ we can define an additive uncertainty set around this model $G(z,\theta)$ (cfr. (9)):

$$\begin{aligned} \mathcal{D}_r(\theta) \quad = \quad & \{G(z) \in H_\infty \mid \\ & \left|G(e^{j\omega}) - G(e^{j\omega},\theta)\right| < r(\omega) \ \ \forall\omega\} \end{aligned} \qquad (13)$$

Consider now the set $\mathcal{R}$ of frequency functions r such that

$$i)\ [\ C(G(z,\theta)) \quad G(z)\ ] \ \text{ stable and}$$

$$ii) J(G(z), C(G(z,\theta)), W_l(z), W_r(z)) \leq 1 \ \forall\, G(z) \in \mathcal{D}_r(\theta).$$

We then define $\bar{r}_{adm}(\omega, G(z,\theta))$ at each ω as

$$\bar{r}_{adm}(\omega, G(z,\theta)) = \sup_{r\in\mathcal{R}} r(\omega). \qquad (14)$$

Given the model $G(z,\theta)$ and the controller $C(G(z,\theta))$, this largest additive uncertainty radius $\bar{r}_{adm}(\omega, G(z,\theta))$ can be computed by a classical ν-analysis problem [5]. The quantity $\bar{r}_{adm}(\omega, G(z,\theta))$ obviously depends on the value of $G(z,\theta)$, the center of the uncertainty set. In order to design an identification experiment, we would need to know the largest admissible uncertainty radius around the to-be-identified model $G(z,\hat{\theta}_N)$ i.e. $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$. Since our aim, eventually, is to do an *a priori* design of an identification input signal $u(t)$ such that the size $r_u(\omega)$ of the estimated uncertainty set (9) is smaller than $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$ for all ω, we cannot let the size of the admissible uncertainty for control, $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$,

depend upon a model that has not yet been estimated. One possibility to tackle this difficulty is to approximate the unknown $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$ by the largest admissible uncertainty radius $\bar{r}_{adm}(\omega, G(z,\theta_{init}))$ around $G(z,\theta_{init})$ which is an available model of the considered system. However, such an approach could lead to poor results if there is a lot of discrepancy between $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$ and $\bar{r}_{adm}(\omega, G(z,\theta_{init}))$. In this paper, we shall instead use the set Θ in Assumption 3.1 to determine a lower bound $r_{adm}(\omega)$ for $\bar{r}_{adm}(\omega, G(z,\hat{\theta}_N))$. This is summarized in the following result.

Proposition 4.1: Consider the control design method of Assumption 3.1, and the performance measure $J(G, C, W_l, W_r)$ of (11) which, for each $\theta \in \Theta$, satisfies (12). Then, for any $\bar{\theta} \in \Theta$, the controller $C(G(z,\bar{\theta}))$ designed from $G(z,\bar{\theta})$ stabilizes and achieves $J(G(z), C(G(z,\bar{\theta})), W_l, W_r) \leq 1$ with all systems $G(z)$ in the additive uncertainty region $\mathcal{D}_{r_{adm}}(\bar{\theta})$ (see (13)) centered at $G(z,\bar{\theta})$ and of size $r_{adm}(\omega)$ defined as:

$$r_{adm}(\omega) \stackrel{\Delta}{=} \min_{\theta\in\Theta} \bar{r}_{adm}(\omega, G(z,\theta)). \tag{15}$$

with $\bar{r}_{adm}(\omega, G(z,\theta))$ as in (14).

Computation of $r_{adm}(\omega)$. One method for the computation of $r_{adm}(\omega)$ is to use a gridding technique: for each ω $r_{adm}(\omega)$ is computed as the smallest value of $\bar{r}_{adm}(\omega, G(z,\theta))$ over randomly selected values of θ. An alternative and more accurate method can be used if the set Θ is an ellipsoid. In such case, $r_{adm}(\omega)$ can be computed as the solution of a ν-analysis problem. Indeed this quantity can be defined equivalently as $r_{adm}(\omega) = \sup_{r\in\mathcal{R}_2} r(\omega)$ where the set $\mathcal{R}_2$ is the set of frequency functions r such that

$$\begin{aligned} &\forall\theta \in \Theta \ \text{ and } \ \forall\Delta(z) \in \{\Delta(z) \in H_\infty \mid |\Delta(e^{j\omega})| < r(\omega)\} : \\ &\quad [\ C(G(z,\theta)) \ \ ; \ \ G(z,\theta) + \Delta(z)\ ] \ \text{ stable and} \\ &\quad J(G(z,\theta) + \Delta(z), C(G(z,\theta)), W_l, W_r) \leq 1. \end{aligned} \tag{16}$$

This is a ν-analysis problem since both $C(G(z,\theta))$ and $G(z,\theta)$ can be expressed as Linear Fractional Transformations (LFT) of the variable $\theta \in \Theta$, with Θ an ellipsoid. This LFT description is indeed possible for a model $G(z,\theta)$ identified using PE identification with a resulting controller $C(G(z,\theta))$ obeying Assumption 3.1: see [1], [4].

V. Identification for control at the cheapest cost

We know from the analysis of Section II that the true G_0 lies with a probability level that we can select in the set $\mathcal{D}_{r_u}(\hat{\theta}_N)$ defined by (9), where $r_u(\omega)$ depends on the input signal spectrum and is given by (10). On the other hand, we know by Proposition 4.1 that the controller $C(G(z,\hat{\theta}_N))$ computed from an identified model $G(z,\hat{\theta}_N)$ is *satisfactory* for all models in the set $\mathcal{D}_{r_{adm}}(\hat{\theta}_N)$. By satisfactory is meant that $\hat{C} = C(G(z,\hat{\theta}_N))$ stabilizes and achieves $J(G(z), \hat{C}(z), W_l(z), W_r(z)) \leq 1$ $\forall G \in \mathcal{D}_{r_{adm}}(\hat{\theta}_N)$. Putting these two results together we conclude that a controller $C(G(z,\hat{\theta}_N))$ satisfies the stability and performance requirements with G_0 (at the desired probability level) if $\mathcal{D}_{r_u}(\hat{\theta}_N) \subseteq \mathcal{D}_{r_{adm}}(\hat{\theta}_N)$, or equivalently if the size $r_u(\omega)$ of $\mathcal{D}_{r_u}(\hat{\theta}_N)$ is at each frequency smaller than the size $r_{adm}(\omega)$ of $\mathcal{D}_{r_{adm}}(\hat{\theta}_N)$. We now seek the input signal with the cheapest cost that achieves this objective. Our cheapest experiment design problem for control can thus be re-formulated as follows:

Cheapest experiment design problem for control: *Determine the stationary input signal $u(t)$ for an identification experiment performed on G_0 with N data in such a way that the total power $\mathcal{P}_u$ of u(t) is minimized, under the constraint that the size $r_u(\omega)$ of the identified uncertainty region $\mathcal{D}_{r_u}(\hat{\theta}_N)$ is at each frequency smaller than $r_{adm}(\omega)$.*

Roughly speaking, the size $r_u(\omega)$ of the identified uncertainty region $\mathcal{D}_{r_u}(\hat{\theta}_N)$ increases when the input signal power $\mathcal{P}_u$ decreases. The cheapest identification experiment for control is thus the one for which the size $r_u(\omega)$ of the identified uncertainty region $\mathcal{D}_{r_u}(\hat{\theta}_N)$ is as large as possible under the constraint that $r_u(\omega) \leq r_{adm}(\omega) \quad \forall\omega$.

This experiment design problem does not have a unique solution. Indeed, if $u_0(t)$ is one solution, then all signals having the same moment vector $x_u(\theta_0)$ as $u_0(t)$ also solve the optimization problem. Our approach will be to solve for the optimal moment vector $x_u(\theta_0)$, and then to realize an input that has such moment vector. The optimization problem on the moment vector would be exactly solvable if the expressions of $r_u(\omega)$ and $\mathcal{P}_u$ as a function of the moment vector were not dependent on the unknown true system via θ_0 and σ_e^2. Such dependence on the unknown true system is inherent to all identification experiment design problems. The classical approach which we adopt here, is to replace the unknown θ_0 and σ_e^2 in the optimization problem leading to $x_u(\theta_0)$ by available estimates θ_{init} and $\sigma^2_{e,init}$ of those quantities. Consequently, the moment vector $x_u(\theta_0)$ of the input signal(s) that solve the experiment design problem described above in this section can be approximated by the optimal vector $x_{u,opt}$ of the following optimization problem on the variable $x_u = \begin{pmatrix} x_0 & x_1 & \dots & x_n \end{pmatrix}^T \in \mathbf{R}^{(n+1)}$:

$$\min_{x_u} \ \sum_{i=0}^{n} c_i x_i \tag{17}$$

subject to

$$\chi\lambda_1\left(T(e^{j\omega})\left(\bar{M} + \sum_{i=0}^{n} M_i x_i\right)^{-1} T(e^{j\omega})^T\right) \leq r^2_{adm}(\omega) \quad \forall\omega \tag{18}$$

and

$$R(x_u) \triangleq \begin{pmatrix} x_0 & x_1 & ... & x_n \\ x_1 & x_0 & ... & x_{n-1} \\ ... & ... & ... & ... \\ x_n & x_{n-1} & ... & x_0 \end{pmatrix} \geq 0 \qquad (19)$$

where we have used the shorthand notations: $c_i = c(\theta_{init})$, $T(e^{j\omega}) = T(e^{j\omega}, \theta_{init})$, $\bar{M} = \bar{M}(\theta_{init})$ and $M_i = M_i(\theta_{init}, \sigma^2_{e,init})$ and the relations for P_θ, $r_u(\omega)$ and $\mathcal{P}_u$ as a function of the moment vector (see (6), (10) and (7)). The constraint (19) is there because a real vector $x_u = (x_0 \ x_1 \ ... \ x_n)^T$ of dimension $n+1$ is the moment vector of a quasi-stationary signal $u(t)$ if and only if the Toeplitz matrix $R(x_u)$ is positive semi-definite (see e.g. [7]). The optimization problem (17)-(19) can be solved exactly provided the frequency function $r_{adm}(\omega)$ can be written as (or approximated by) a rational transfer function. Indeed, by the Kalman-Yakubovich-Popov Lemma [11], the conditions (18) defined at each frequency can be transformed into one single LMI making the optimization problem tractable. This is summarized in the following theorem.

Theorem 5.1: Assume that there is a rational transfer function $r_{adm}(z)$ such that $|r_{adm}(e^{j\omega})| = r_{adm}(\omega)$ for the frequency function $r_{adm}(\omega)$ defined in (15). Then, the solution of the optimization problem (17)-(19) is the optimal vector $x_{u,opt}$ of the following tractable LMI optimization problem:

$$\text{minimize} \sum_{i=0}^{n} c_i x_i \qquad (20)$$

$$\text{over } x_u = (x_0 ... x_n)^T \in \mathbf{R}^{(n+1)} \text{ and a matrix } P = P^*$$

$$\text{subject to } R(x_u) \geq 0 \text{ and}$$

$$\begin{pmatrix} A^*PA - P & A^*PB \\ B^*PA & B^*PB \end{pmatrix} + \begin{pmatrix} C^* \\ D^* \end{pmatrix} \mathcal{X} \begin{pmatrix} C & D \end{pmatrix} \geq 0 \qquad (21)$$

where $R(x_u)$ is defined in (19), $\mathcal{X} \in \mathbf{R}^{3(k+2)\times 3(k+2)}$ is a matrix, independent of the frequency but linearly dependent on x_u, defined as

$$\mathcal{X} = \left(\begin{array}{cc|c|c} 2I_2 & 0 & 0 & 0 \\ 0 & 2\left(\bar{M} + \sum_{i=0}^{n} M_i x_i\right) & & \\ \hline 0 & & 0 & I_{k+2} \\ \hline 0 & & I_{k+2} & 0 \end{array}\right) \qquad (22)$$

and (A, B, C, D) is a state-space representation of the system $\mathcal{F}(e^{j\omega}) \in \mathcal{C}^{3(k+2)\times(k+2)}$ defined as follows:

$$\left(\begin{array}{cc} \frac{r_{adm}(e^{j\omega})}{\sqrt{\chi}} I_2 & 0 \\ 0 & I_k \\ \hline 0 & \begin{pmatrix} 1 \\ -j \end{pmatrix} \Lambda_G^T(e^{j\omega}, \theta_{init}) \\ \Lambda_G(e^{j\omega}, \theta_{init}) \begin{pmatrix} 1 & -j \end{pmatrix} & 0 \\ \hline I_2 & 0 \\ 0 & I_k \end{array}\right) \qquad (23)$$

with χ and $\Lambda_G(z, \theta)$ as defined below (10) and (3), respectively.

Proof. To prove the theorem, we prove that the infinite set of constraints (18) is equivalent to the fact that $\exists P = P^*$ such that (21) is satisfied. For this purpose, we first rewrite the set of constaints (18) using the Schur complements [3] (as is proposed in [9] for a slightly different constraint):

$$\begin{pmatrix} \frac{r^2_{adm}(e^{j\omega})}{\chi} I_2 & T(e^{j\omega}) \\ T(e^{j\omega})^T & \bar{M} + \sum_{i=0}^{n} M_i x_i \end{pmatrix} \geq 0 \quad \forall \omega \qquad (24)$$

Via some simple algebraic manipulations, (24) can then be replaced by the following equivalent set of constraints: $\frac{1}{2} \mathcal{F}(e^{j\omega})^* \mathcal{X} \mathcal{F}(e^{j\omega}) \geq 0 \ \forall \omega$ where $\mathcal{X}$ and $\mathcal{F}(e^{j\omega})$ are defined in (22) and (23), respectively. It is then a consequence of the Kalman-Yakubovich-Popov Lemma [11] that this set of constraints defined at each frequency is fulfilled if and only if $\exists P = P^*$ such that (21) is fulfilled. ■

Determination of an input signal from the optimal moment vector. The optimal vector $x_{u,opt}$ is computed using Theorem 5.1. This vector $x_{u,opt} = x_{u,opt}(\theta_{init})$ is an approximation of the moment vector of the input signals that solve the cheapest experiment design problem for control. Consequently, provided that this approximation is reliable, we can determine one of the quasi-stationary input signals that solve this experiment design problem, by determining one representation $u(t)$ of $x_{u,opt}(\theta_{init})$ i.e. a signal $u(t)$ whose moment vector is equal to $x_{u,opt}(\theta_{init})$. In [7], a method is presented that determines such an input signal as a sum of sinusoids.

VI. Numerical illustration

In order to illustrate our results, we consider as *true system* the ARX system with $B_0(z) = 0.10276 + 0.18123z^{-1}$, $A_0(z) = 1 - 1.99185z^{-1} + 2.20265z^{-2} - 1.84083z^{-3} + 0.89413z^{-4}$, corrupted by a realization of a white noise signal of variance $\sigma_e^2 = 0.5$ [8]. We consider identification experiments on this true system with $N = 500$ data points and a full-order model structure. In this example, we restrict attention to the properties of the sensitivity function and control performance criterion $J(G, C, W_l, W_r)$ is therefore defined as in (11) with the filters: $W_l(z) = diag\left(\frac{0.5165 - 0.4632z^{-1}}{1 - 0.999455z^{-1}}, 0\right)$ and $W_r(z) = diag\,(1, 0)$. The chosen control design method is

the 4-block H_∞ control design method of [6] which has the characteristics described in Section III.

Design of the cheapest input signal $u(t)$ for control. We first estimate $r_{adm}(\omega)$, then compute the optimal moment vector $x_{u,opt}$, and finally design a particular input signal $u(t)$ corresponding to $x_{u,opt}$.

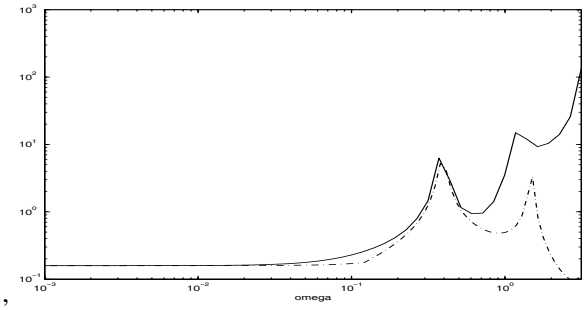

Fig. 1. $r_{adm}(\omega)$ (solid) and $r_u(\omega)$ computed with θ_{init} and the optimal $x_{u,opt}$ (dashdot)

In order to estimate $r_{adm}(\omega)$ (see (15)), we have applied the gridding technique described at the end of Section IV. The estimated frequency function $r_{adm}(\omega)$ is represented in Figure 1. The optimal moment vector $x_{u,opt}$ is determined by solving the LMI optimization problem of Theorem 5.1 for which we have used $\sigma^2_{e,init} = 0.5265$ and an initial estimate $\theta_{init} = (-1.9755, 2.1965, -1.8495, 0.8881, 0.0817, 0.172)^T$. Figure 1 shows the frequency function $r_u(\omega)$ that is computed using the left-hand side of the constraint (18) and the optimal moment vector $x_{u,opt}$. By construction, the moment vector $x_{u,opt}$ is optimized in such a way that this frequency function $r_u(\omega)$ is the largest possible under the constraint $r_u(\omega) \leq r_{adm}(\omega)\ \forall\omega$. It could therefore appear surprising that the second peak of $r_u(\omega)$ is so low and that $r_u(\omega)$ decreases after this peak. The reason is that $r_u(\omega)$ depends on the covariance matrix P_θ, which is parametrized only by $n+1$ components. Therefore the behaviour of $r_u(\omega)$ at one frequency is dependent on its behaviour at other frequencies and the second peak could not be increased without making the first peak larger than $r_{adm}(\omega)$ and/or $r_u(\omega)$ larger than $r_{adm}(\omega)$ in low frequencies.

We have designed an input signal $u(t)$ corresponding to $x_{u,opt}$ under the form of a periodic signal:

$$\begin{aligned} u(t) &= 2.07 + 2.63\ cos(0.17t) + 0.69\ cos(0.4t) \\ &\quad +0.09\ cos(1.44t) + 0.17\ cos(\pi t) \end{aligned} \tag{25}$$

It is interesting to notice that this periodic input signal $u(t)$ has components in the low frequencies (where $r_{adm}(\omega)$ is minimal) and at $\omega = 0.4$ and $\omega = 1.44$ where the two resonance peaks of $G(z, \theta_{init})$ are located.

Verification of the procedure. We have applied the input signal $u(t)$ defined in (25) to the true system. From $N = 500$ recorded input-output data, we have identified a model $\hat{G}$ along with the additive uncertainty region $\mathcal{D}_{r_u}(\hat{\theta}_N)$ around $\hat{G}$. The size $r_u(\omega)$ of this uncertainty region has not been estimated using the approximation (10), but computed exactly using an LMI optimization problem which can be found in [2]. From $\hat{G}$, we have then designed a controller $\hat{C}$ using the H_∞ control design method of [6], and we have verified whether $\hat{C}$ stabilizes and achieves the desired performance level with all systems in the identified $\mathcal{D}_{r_u}(\hat{\theta}_N)$. This was indeed the case.

VII. Conclusions

We have presented a new approach to the interplay between identification design and robust control based on an identified model and its corresponding uncertainty set. Rather than seeking an experiment design that minimizes some control-oriented quality measure of the estimated uncertainty set for a predefined constraint on allowable input power, we have in this contribution sought the experiment design with the cheapest cost (as measured by total input signal power) that delivers just enough precision on the estimated uncertainty set for the robust control specifications to be satisfied.

References

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

[2] X. Bombois, G. Scorletti, P. Van den Hof, and M. Gevers. Least costly identification experiment for control. A solution based on a high-order model approximation. In *CD-ROM Proc. American Control Conference*, Boston, June 2004.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*, volume 15 of *Studies in Appl. Math.* SIAM, Philadelphia, June 1994.

[4] M. Dinh, G. Scorletti, V. Fromion, and E. Magarotto. Parametrized H_∞ controller design for adaptive trade-off by finite dimensional LMI optimization. In *CD-ROM Proc. European Control Conference*, Cambridge, United Kingdom, 2003.

[5] G. Ferreres and V. Fromion. Computation of the robustness margin with the skewed μ-tool. *Syst. Control Letters*, 32:193–202, 1997.

[6] G. Ferreres and V. Fromion. H_∞ control for a flexible transmission system. In *CD-ROM Proc. European Control Conference*, Brussels, Belgium, 1997.

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

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

[9] K. Lindqvist and H. Hjalmarsson. Optimal input design using Linear Matrix Inequalities. In *CD-ROM Proc. IFAC Symposium on System Identification*, Santa Barbara, California, 2000.

[10] L. Ljung. *System Identification: Theory for the User, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1999.

[11] V. M. Popov. *Hyperstability of Control Systems*. Springer-Verlag, New York, 1973.

[12] M. Zarrop. *Design for Dynamic System Identification*. Lecture Notes in Control and Inform. Sci. 21, Springer Verlag, Berlin, New-York, 1979.