Design of least costly identification experiments. The main philosophy accompanied by illustrative examples*

Xavier Bombois[1] and Gérard Scorletti[2]

[1] Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628CD Delft, The Netherlands
x.j.a.bombois@tudelft.nl

[2] Laboratoire Ampère
Ecole Centrale de Lyon
36 avenue Guy de Collongue, 69134 Ecully Cedex, France
gerard.scorletti@ec-lyon.fr

Résumé. Cet article présente, d'une manière pédagogique, les idées maîtresses derrière un nouveau paradigme de synthèse optimale de l'expérience d'identification. Ce nouveau paradigme a pour objectif de synthétiser l'expérience d'identification la moins coûteuse tout en garantissant un modèle suffisamment précis pour la commande. Le but de cet article est aussi d'illustrer au moyen d'exemples bien choisis les avantages liés à la synthèse optimale du signal d'excitation pour une identification par rapport a l'application du classique bruit blanc. Comme on le verra dans ces exemples, le fait de faconner adéquatement le signal d'excitation permet de réduire significativement le coût de l'identification.

Abstract. The goal of this paper is on the one hand to give a tutorial on the main ideas of a recently introduced paradigm for optimal experiment design whose objective is to design the least costly identification experiment while guaranteeing a sufficiently accurate model for e.g. control. On the other hand, the second goal is to illustrate with well chosen examples the advantages of designing optimally the excitation signal for an identification instead of using a classical white excitation. As we will see in these examples, shaping appropriately the excitation signal allows one to reduce significantly the cost of an identification experiment.

1 Introduction

Optimal identification experiment design is the scientific exercise of designing the excitation signal for the identification of a real-life system in an

---

optimal way. The typical approach to this problem has been to design the power spectrum of the excitation signal in order 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 [21][Chapters 12 and 13] and e.g. [22, 13, 15, 11, 20, 8, 14, 17]). In [6], we introduced a new paradigm for optimal experiment design that is somehow the dual of the classical approach. In this new paradigm, the goal is to design the power spectrum of the excitation signal in such a way that the corresponding identification experiment is the *least intrusive* for the underlying system while guaranteeing that the identified model is sufficiently accurate for the intended application (e.g. control). The optimal identification experiment is then called the *least costly identification experiment* for control. The least costly paradigm has been further developed in a series of paper: [3, 4, 1, 16].

The main goal of the present paper is to give a tutorial on the main ideas of this new optimal experiment design paradigm and to illustrate with well chosen examples the advantages of designing optimally the excitation signal instead of using a classical white excitation. As we will see in these examples, shaping appropriately the excitation signal allows one to reduce significantly the cost of an identification experiment.

The new paradigm is quite appropriate in control applications where identification experiments entail an important economical cost (e.g. in the case of industrial processes). In those applications, one is not really interested in obtaining the maximal accuracy, but in reducing the cost of the identification experiment while guaranteeing that the identified model is sufficiently accurate for its intended use. The new paradigm is also quite appropriate in relation with robust control. Indeed, imposing, such as in the least costly paradigm, that the modeling error remains below a certain threshold is equivalent with imposing bounds on the size of the uncertainty region around the identified model $\hat{G}(z)$. It is clear that a large uncertainty around the model reduces the level of performance that a controller $\hat{C}(\hat{G})$ designed with this uncertain model can achieve on the unknown true system $G_0(z)$. In other words, for the controller $\hat{C}(\hat{G})$ to achieve a certain level of performance on $G_0(z)$, the modeling error must remain below a certain threshold $r_{adm}(\omega)$:

$$|\hat{G}(e^{j\omega}) - G_0(e^{j\omega})| < r_{adm}(\omega) \quad \forall \omega \tag{1}$$

The threshold $r_{adm}(\omega)$ in (1), which represents the largest admissible uncertainty for a given level of desired performance, can be computed using robust analysis techniques such as ν-analysis [9]. See also [6, 3] for specific examples in the least costly setup. Note that the bound on the uncertainty of the model can be expressed in another domain than in the frequency domain. The bound can also be expressed as a bound on the covariance matrix of the identified parameter vector (see [4, 16]). However, we here choose the frequency domain representation for its simplicity.

Besides the desired accuracy for the identified model, another crucial ingredient in the least costly paradigm is the definition of the cost of the identification. If we limit attention to open-loop identification [21] with a fixed amount of data, a possible measure for the cost of the identification is the power of the input signal: $\mathcal{P}_u = \bar{E}u^2(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) \, d\omega$. The power of the output signal could also be considered or a combination of both input and output powers. In closed-loop identification, a quite realistic definition of the cost was introduced in [4]. In that paper, the cost is a function of the power of the perturbations induced by the excitation signal on the normal operation of the closed-loop system.

In the sequel, we will consider these different definitions of the identification cost as well as both the open-loop and closed-loop configurations in the sequel. However, as a starting point, we will focus on the open-loop configuration with a fixed amount N of data and with the cost of the identification defined as $\mathcal{P}_u$. The optimal experiment design problem can thus be formulated as follows:

Open-loop experiment design problem (fixed N). *Determine the power spectrum $\Phi_u(\omega)$ of the input signal of minimal power such that the model $\hat{G}(z)$ identified with this input signal satisfies the accuracy constraint (1) for a given threshold $r_{adm}(\omega)$.*

2 Prediction Error Identification aspects (open-loop case)

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

$$
\mathcal{S}: \quad
\begin{aligned}
y(t) &= G_0(z)u(t) + \overbrace{H_0(z)e(t)}^{=v(t)} \\
&= G(z,\theta_0)u(t) + H(z,\theta_0)e(t) \quad\quad (2)
\end{aligned}
$$

for some unknown parameter vector $\theta_0 \in \mathbf{R}^k$, and with $e(t)$ a white noise of variance σ_e^2. In open-loop identification, the identification experiment consists of applying to (2) an input sequence $u(t)$ ($t = 1...N$) and of collecting the corresponding output $y(t)$: $Z^N = \{y(t) \ u(t) \mid t = 1...N\}$. When designing this experiment, if we suppose the duration N of the identification experiment fixed a-priori, we see that the user has to make the following choices:

1. the power of the signal $u(t)$ that will be applied

2. the frequency content of this signal $u(t)$

These two items are encompassed in the power spectrum $\Phi_u(\omega)$ of $u(t)$ which is the design variable of our optimal experiment design problem.

Once the data set Z^N has been collected, prediction error identification can be used to determine a consistent estimate $\hat{\theta}_N$ of the true parameter vector θ_0 using the following criterion:

$$\hat{\theta}_N \triangleq \arg \min_{\theta} \frac{1}{N} \sum_{t=1}^{N} \epsilon^2(t, \theta) \tag{3}$$

with $\epsilon(t, \theta) \triangleq H(z, \theta)^{-1} (y(t) - G(z, \theta)u(t))$.

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 [21]: $P_\theta = \frac{\sigma_e^2}{N} \left(\bar{E} \left(\psi(t, \theta_0) \psi(t, \theta_0)^T \right) \right)^{-1}$ 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 [21]:

$$\begin{aligned}
P_\theta^{-1} &= \left(\frac{N}{\sigma_e^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} F_u(e^{j\omega}, \theta_0) F_u(e^{j\omega}, \theta_0)^* \Phi_u(\omega) d\omega \right) \\
&+ \left(N \frac{1}{2\pi} \int_{-\pi}^{\pi} F_e(e^{j\omega}, \theta_0) F_e(e^{j\omega}, \theta_0)^* d\omega \right)
\end{aligned} \tag{4}$$

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 $\Lambda_H(z, \theta) = \frac{\partial H(z, \theta)}{\partial \theta}$.

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:

$$\mathcal{D}_{r_u}(\hat{\theta}_N) = \left\{ G(z) \in H_\infty \mid \left| G(e^{j\omega}) - G(e^{j\omega}, \hat{\theta}_N) \right| < r_u(\omega) \; \forall \omega \right\} \tag{5}$$

Consider the following first order approximation[1] 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 (4). Using this approximation, the size $r_u(\omega)$ of $\mathcal{D}_{r_u}(\hat{\theta}_N)$ can then be written as :

$$r_u(\omega) = \alpha \sqrt{\lambda_1 \left(T(e^{j\omega}, \theta_0) P_\theta T(e^{j\omega}, \theta_0)^T \right)} \tag{6}$$

where $T(e^{j\omega}, \theta_0) \triangleq \begin{pmatrix} Re(\Lambda_G^T(e^{j\omega}, \theta_0)) \\ Im(\Lambda_G^T(e^{j\omega}, \theta_0)) \end{pmatrix} \in \mathbf{R}^{2 \times k}$, $\lambda_1(A)$ denotes the largest eigenvalue of A and α is a real constant dependent on the chosen probability level: if we want $Pr(G_0 \in \mathcal{D}_{r_u}(\hat{\theta}_N)) = 0.95$, then α is chosen such that

---

[1]The first order approximation is not absolutely necessary to compute the function $r_u(\omega)$ (see e.g. [17, 6]). However, we have here chosen this approach for the sake of simplicity.

$Pr(\chi^2(k_G) < \alpha^2) = 0.95$ (k_G is the number of parameters in $G(z, \theta)$).

The size $r_u(\omega)$ of the uncertainty region $\mathcal{D}_{r_u}$ is thus a measure of the modeling error between the identified model $G(z, \hat{\theta}_N)$ and the unknown true system $G_0(z) = G(z, \theta_0)$. We have indeed that, modulo a certain probability level,:

$$|G(e^{j\omega}, \hat{\theta}_N) - G_0(e^{j\omega})| < r_u(\omega) \quad \forall \omega \tag{7}$$

In the sequel, we will refer to $r_u(\omega)$ as the bound on the modeling error or, for the sake of simplicity, as the modeling error. Note that, after an identification experiment, P_θ and $r_u(\omega)$ can be estimated by replacing the actual θ_0 and σ_e^2 by their estimate $\hat{\theta}_N$ and $\hat{\sigma}_e^2 = \frac{1}{N}\sum_{t=1}^{N} \epsilon^2(t, \hat{\theta}_N)$.

An important observation at this stage is also that the modeling error $r_u(\omega)$ is a function of the covariance matrix P_θ and thus, by (4), a function of the input signal $u(t)$ used during the identification experiment. Obviously, for signals $u(t)$ having the same frequency content, the more powerful $u(t)$, the smaller the modeling error. Moreover, for signals $u(t)$ having the same power, different frequency contents will lead to different modeling errors $r_u(\omega)$. This phenomenon is crucial for optimal experiment design and will be illustrated by the following example.

Example 1 *We consider the following true system*

$$\mathcal{S}: \quad y(t) = \frac{3.6z^{-1}}{1 - 0.7z^{-1}} u(t) + \left(1 - 0.9z^{-1}\right) e(t) \tag{8}$$

with $\sigma_e^2 = 1$ and we choose the following full order model structure:

$$\mathcal{M}: \quad G(z, \theta) = \frac{bz^{-1}}{1 - fz^{-1}} \qquad H(z, \theta) = 1 + cz^{-1} \qquad \theta = \begin{pmatrix} b \\ c \\ f \end{pmatrix} \tag{9}$$

In this example, we want to compare the accuracy of the model identified in this model structure with two different input signals. The first input signal $u_1(t)$ is a realization of length $N = 500$ of a white noise with variance 1.7 and the second input signal $u_2(t)$ is generated as the 500 first samples of a cosine of amplitude 0.75 at $\omega = 0.15$ rad/s i.e. $u_2(t) = 0.75cos(0.15t)$ ($t = 1...500$). The two signals are represented in Figure 1. These two input signals have clearly a different frequency content and it is also clear that the power of $u_1(t)$ ($\mathcal{P}_{u_1} = 1.7$) is approximatively six times larger than the power of u_2 ($\mathcal{P}_{u_2} \approx 0.28$). We could therefore think that the input signal u_1 would lead to a more accurate model than the input signal u_2. However, it is not the case. The model $\hat{G}_2$ identified with u_2 has indeed a slightly better accuracy than the model $\hat{G}_1$ identified with u_1 at each frequency: the modeling error $r_{u_1}(\omega)$ obtained with u_1 is, as can be seen in Figure 2, larger that the modeling error $r_{u_2}(\omega)$ obtained with u_2 at each frequency: $r_{u_1}(\omega) \geq r_{u_2}(\omega) \quad \forall \omega$.
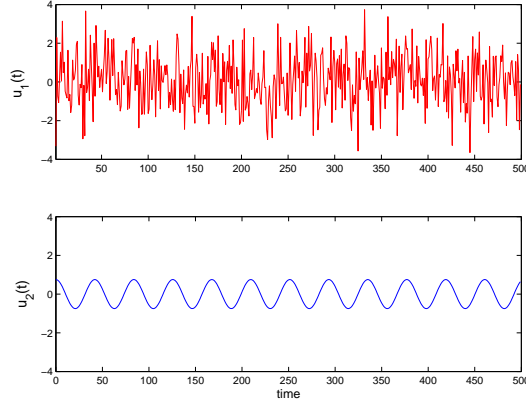
Figure 1: Two input signals: $u_1(t)$ (top) and $u_2(t)$ (bottom)

It is important to note that a cosine at another frequency than $\omega = 0.15$ rad/s will not lead to the same accuracy. In Figure 3, the modeling error obtained with another cosine i.e. $u_3(t) = 0.75cos(0.05t)$ $(t = 1...500)$ is compared with the modeling error obtained with $u_1(t)$. We see here that the modeling error with u_3 is smaller in low frequencies than the one with u_1, but much larger in high frequencies.
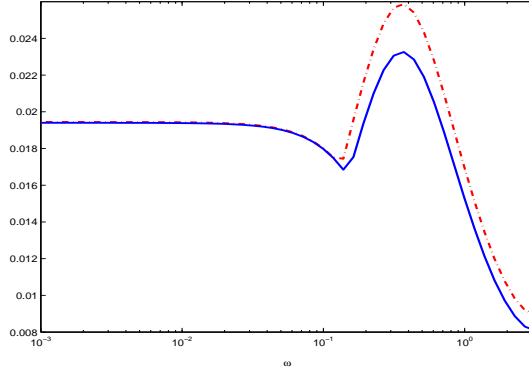

Figure 2: Modeling error $r_u(\omega)$ obtained after an identification with $u_1(t)$ (red dashdot) and obtained after an identification with $u_2(t)$ (blue solid) (α is here chosen equal to one)

3 Optimal open-loop experiment design as an optimization problem

3.1 Convex formulation

The previous example shows that, by shaping $\Phi_u(\omega)$ appropriately, we can obtain a certain model accuracy with a (much) less powerful excitation signal $u(t)$. Finding this appropriate $\Phi_u(\omega)$ for a given true system is the
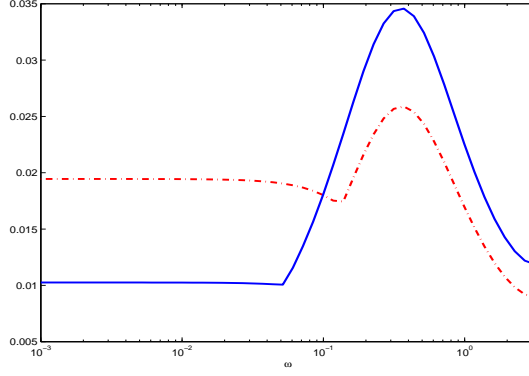
Figure 3: Modeling error $r_u(\omega)$ obtained after an identification with $u_1(t)$ (red dashdot) and obtained after an identification with $u_3(t) = 0.75cos(0.05t)$ (blue solid) (α is here chosen equal to one)

whole principle behind optimal experiment design. In this section, we will formulate the optimal experiment design problem presented at the end of Section 1 into a convex optimization problem.

A first step towards that end is to formulate the accuracy constraint (1) as a function of the decision variable $\Phi_u(\omega)$. This can be done by using the relation (7) which gives a bound $r_u(\omega)$ on the modeling error between the identified model $G(z, \hat{\theta}_N)$ and the true system $G(z, \theta_0)$:

$$\alpha \ \sqrt{\lambda_1 \left(T(e^{j\omega}, \theta_0)P_\theta T(e^{j\omega}, \theta_0)^T\right)} < r_{adm}(\omega) \quad \forall \omega \tag{10}$$

where the covariance matrix P_θ is, as shown in (4), a function of the decision variable $\Phi_u(\omega)$.

Using the new constraint (10), the optimal identification experiment defined at the end of Section 1 is the particular experiment where the power spectrum $\Phi_u(\omega)$ of $u(t)$ is the one solving:

$$\min_{\Phi_u(\omega)} \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) \ d\omega$$

$$\text{subject to} \ \ \Phi_u(\omega) \geq 0 \ \ \forall \omega \quad \text{and to} \tag{11}$$

$$\alpha \ \sqrt{\lambda_1 \left(T(e^{j\omega}, \theta_0)P_\theta T(e^{j\omega}, \theta_0)^T\right)} < r_{adm}(\omega) \quad \forall \omega$$

We can observe a couple of issues when looking at this optimization problem. The first issue is that this optimization problem has an infinite number of constraints since (10) must hold at each frequency. Even though more elaborated solutions exist (see e.g. [3]), the easiest and more efficient way to circumvent this issue is to grid the frequency range in order to obtain a finite number of constraints. A second issue is the fact that the constraint (10)

7

depends on the true parameter vector θ_0 and the true noise variance σ_e^2 (via P_θ). This is the well-known chicken-and-egg problem related to each optimal experiment design problem: the optimal input spectrum for the identification of an unknown true system depends on this true system. Note that this chicken-and-egg problem is not specific to the optimal experiment design paradigm used in this paper. It is a general problem: see e.g. [21]. To circumvent this issue, the true parameter vector θ_0 and the true noise variance σ_e^2, which are required to compute the bound $r_u(\omega)$ on the modeling error in (10), are generally replaced by initial estimates θ_{init} and $\sigma_{e,init}^2$ of those quantities. These initial estimates can be obtained by performing, prior to the design of the optimal experiment, a short identification experiment with e.g. white noise. For more elaborated techniques to circumvent the chicken-and-egg problem, we refer the reader to [4] and [12]. A third issue with the optimization problem (11) is that the constraint (10) is as such not linear in the decision variable $\Phi_u(\omega)$. However this constraint can easily be linearized using the Schur complements [7]. Indeed, as proven in the appendix, the constraint (10) can be rewritten as follows:

$$P_\theta^{-1} > R_{adm}(\omega) \quad \forall \omega \tag{12}$$

with $R_{adm}(\omega) = \frac{\alpha^2}{r_{adm}^2(\omega)} T^T(e^{j\omega}, \theta_0) T(e^{j\omega}, \theta_0)$. The constraint (12) is now affine in the decision variable $\Phi_u(\omega)$ since P_θ^{-1} has this property (see (4)).

Based on the above considerations, we can thus rewrite the optimization problem (11) as the following LMI (Linear Matrix Inequality) optimization problem:

$$\min_{\Phi_u(\omega)} \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) \, d\omega$$

$$\text{subject to} \quad \Phi_u(\omega) \geq 0 \quad \forall \omega \quad \text{and to} \tag{13}$$

$$\left(\frac{N}{\sigma_e^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} F_u(e^{j\omega}, \theta_0) F_u(e^{j\omega}, \theta_0)^* \Phi_u(\omega) d\omega \right) \ + \ ...$$

$$... \ \left(N \frac{1}{2\pi} \int_{-\pi}^{\pi} F_e(e^{j\omega}, \theta_0) F_e(e^{j\omega}, \theta_0)^* d\omega \right) > R_{adm}(\omega) \quad \forall \omega$$

where, as mentioned above, the last constraint will only be evaluated in a given frequency grid and where σ_e^2 and θ_0 are replaced by their initial estimates.

Remark 1. As mentioned in the introduction, we have chosen in this paper to express the accuracy constraint (1) for control by a bound on the modulus of the modeling error between the identified model and the true system. The accuracy constraint for control can be formulated in various other ways as shown in e.g. [4, 16]. However, it is important to note that they all eventually lead to a constraint which is very close to (12). For example, in

[16], the constraint is $P_\theta^{-1} > I_{app}$ where I_{app} is a (frequency-independent) matrix which is determined based on the desired control objectives.

Remark 2. Until now, we have considered the duration N of the identification experiment as fixed a-priori. If this is not the case, a good approach is to determine, with the optimization problem (13), the optimal spectrum $\Phi_u(\omega)$ for different values of the length N. Since, for increasing values of N, the optimal cost function $\mathcal{P}_u$ decreases, such approach allows one to find the *optimal* combination for the duration of the identification and the induced cost $\mathcal{P}_u$.

Remark 3. The cost of the identification experiment in (13) is measured using the power $\mathcal{P}_u$ of the input signal. This is of course not the only way to measure the cost of the identification. The power $\mathcal{P}_y$ of the output signal could also be used instead since it can also be formulated as an affine function of the decision variable $\Phi_u(\omega)$. A combination of both $\mathcal{P}_u$ and $\mathcal{P}_y$ could also be considered.

3.2 Parametrization of the power spectrum $\Phi_u(\omega)$

The optimization problem (13) is an LMI optimization problem with a decision variable $\Phi_u(\omega)$ of infinite dimension. To be able to solve (13) using e.g. LMI optimization [7], a finite linear parametrization of $\Phi_u(\omega)$ should be used. A common parametrization is the following one [20, 19]:

$$\Phi_u(\omega) = \sum_{r=-m}^{m} c_r \, e^{j\omega \, r} \geq 0 \tag{14}$$

with $c_r = c_{-r}$. If we replace $\Phi_u(\omega)$ by this parametrization in the optimization problem (13), this problem remains a LMI optimization problem[2] for which the decision variables are the parameters c_r $(r = 0...m)$ of this parametrization [20, 4].

The scalar m is an user choice. We observe that the larger m is chosen, the more flexible is the parametrization of the spectrum and that choosing $m = 0$ is equivalent to restrict attention to a flat spectrum (white noise): $\Phi_u(\omega) = c_0 \, \forall\omega$.

---

[2] The positivity of the spectrum $\Phi_u(\omega)$ parametrized by (14) can be guaranteed by the existence of a symmetric matrix Q satisfying the following LMI constraint that can be added to (13) [20, 4]):

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

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

$$A = \begin{pmatrix} 0 & 0 \\ I_{m-1} & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & ... & 0 \end{pmatrix}$$
$$C = \begin{pmatrix} c_1 & c_2 & ... & c_m \end{pmatrix} \quad D = \frac{c_0}{2}$$

It is also important to note that the set of signals $u(t)$ whose spectrum can be described as (14) is the set of signals that can be generated by a white noise of unit variance passing through an arbitrary FIR filter of length $m+1$. The parameters c_r $(r = 0...m)$ of this parametrization are indeed by construction the finite auto-correlation sequence of the signal corresponding to $\Phi_u(\omega)$. As such, it is also very easy to generate an input sequence $u(t)$ having the desired spectrum when using this parametrization: see [20, 19] for details.

Note that other parametrizations can also be used to obtain a finite dimension optimization problem e.g. $\Phi_u(\omega) = \sum_{r=1}^{m} c_r \left(\delta(\omega - \omega_r) + \delta(\omega + \omega_r) \right)$ corresponding to a multisine signal $u(t)$, or $\Phi_u(\omega) = \sum_{r=-m}^{m} c_r \mathcal{B}_r(e^{j\omega})$ where $\mathcal{B}_r(e^{j\omega})$ are preselected basis functions [17].

4 Optimal experiment design in closed loop

Besides open-loop identification, the optimal experiment design problem can also be formulated for *direct closed-loop identification* [4]. Direct closed-loop identification follows the same procedure as in Section 2 [21]. The main difference is nevertheless that the set $Z^N = \{y(t)\ u(t)|t = 1...N\}$ of input-output data are collected when the true system (2) is operated in closed loop (see Figure 4). Due to the closed-loop configuration, the system can no longer be excited via the input signal $u(t)$. The excitation signal is in this case an external signal $r(t)$ that can be applied at the reference or at the control input. In the latter case, we have thus that: $u(t) = -C_{id}(z)y(t)+r(t)$ (as shown in Figure 4). The decision variable is here consequently the power spectrum $\Phi_r(\omega)$ of this external signal $r(t)$.
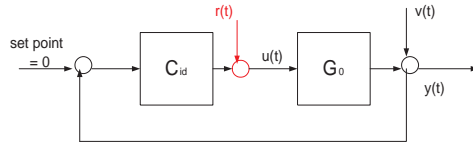

Figure 4: Direct closed-loop identification: the data set Z^N is collected by exciting the closed loop made up of the true system and a controller C_{id} using an external signal $r(t)$

Once the data set Z^N has been collected, the identified parameter vector can be identified and has the same statistical properties as in open-loop identification. The covariance matrix P_θ is also here given by: $P_\theta = \frac{\sigma_e^2}{N} \left(\bar{E} \left(\psi(t, \theta_0)\psi(t, \theta_0)^T \right) \right)^{-1}$. The inverse of this covariance matrix can also be expressed as an affine function of the design variable $\Phi_r(\omega)$ [4]:

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

$$+ \quad \left(N \frac{1}{2\pi} \int_{-\pi}^{\pi} F_v(e^{j\omega}, \theta_0) F_v(e^{j\omega}, \theta_0)^* d\omega \right) \tag{15}$$

with $F_r(z, \theta_0) = \frac{S_{id}}{H_0} \Lambda_G(z, \theta_0)$, $F_v(z, \theta_0) = \frac{\Lambda_H(z, \theta_0)}{H_0} - C_{id} S_{id} \Lambda_G(z, \theta_0)$ and $S_{id} = (1 + C_{id} G_0)^{-1}$ the sensitivity function of the closed loop $[C_{id} \ G_0]$.

The cost of the direct closed-loop identification experiment can be e.g. measured by the power $\mathcal{P}_r$ of the external signal $r(t)$. However, we will consider in the sequel an alternative definition which seems closer to the actual cost of a closed-loop experiment. Suppose that the closed-loop presented in Figure 4 represents a production unit with a product $y(t)$. In normal operation the signals $u(t)$ (control signal) and $y(t)$ (the product) are given by: $y(t) = S_{id} v(t)$ and $u(t) = -C_{id} S_{id} v(t)$ ($v(t) = H_0(z)e(t)$). 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:

$$y(t) = \overbrace{G_0 S_{id} r(t)}^{y_r(t)} + S_{id} v(t) \tag{16}$$

$$u(t) = \overbrace{S_{id} r(t)}^{u_r(t)} - C_{id} S_{id} v(t) \tag{17}$$

Those disturbances induce a loss of production quality. Consequently, we can measure the cost caused by the application of a signal with power spectrum $\Phi_r(\omega)$ using the following cost function:

$$
\begin{aligned}
\mathcal{J}_r &= \beta_y \mathcal{P}_{y_r} + \beta_u \mathcal{P}_{u_r} \\
&= \beta_y \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{y_r}(\omega) \, d\omega \right) + \beta_u \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_r}(\omega) \, d\omega \right)
\end{aligned} \tag{18}
$$

where the scalars β_u and β_y can be e.g. chosen both equal to one. As we can see, this cost function is a linear function of the decision variable $\Phi_r(\omega)$.

Let us now formulate the least costly identification problem for the closed-loop configuration:

Closed-loop experiment design problem (fixed N). *Determine the power spectrum $\Phi_r(\omega)$ of the excitation signal $r(t)$ corresponding to the smallest cost $\mathcal{J}_r$ while guaranteeing that the model $\hat{G}(z)$ identified with this excitation signal satisfies the accuracy constraint (1) for a given threshold $r_{adm}(\omega)$.*

The accuracy constraint (1) can here also be replaced by the constraint (10) which is in turn also equivalent with the constraint (12). Using now the fact that both P_θ^{-1} and $\mathcal{J}_r$ are affine in the decision variable $\Phi_r(\omega)$. The optimal closed-loop experiment design problem can be solved by the following LMI optimization problem.

$$\min_{\Phi_r(\omega)} \quad \mathcal{J}_r(\theta_0)$$

$$\text{subject to} \quad \Phi_r(\omega) \geq 0 \quad \forall \omega \quad \text{and to} \tag{19}$$

$$\frac{N}{\sigma_e^2} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} F_r(\theta_0) F_r^*(\theta_0) \, \Phi_r(\omega) \; + \; F_v(\theta_0) F_v^*(\theta_0) \, \sigma_e^2 \, d\omega \right) > R_{adm}(\omega) \;\; \forall \omega$$

where, similarly as for the open-loop case, the last constraint will only be evaluated in a given frequency grid and where σ_e^2 and θ_0 are replaced by their initial estimates. Since $\mathcal{J}_r$ is also function of θ_0, the initial estimate θ_{init} will also be used to evaluate the cost function.

5   Numerical illustrations

5.1   Illustration 1

In the sequel, we will present two illustrations of optimal experiment design. For the first illustration, we will consider the same true system as in the example of Section 2. This true system is given in (8) and the full order model structure for this system is given in (9).

The optimal experiment design problem we wish to solve for this true system is the following one: determine, for an open-loop experiment of duration $N = 500$, the power spectrum $\Phi_u(\omega)$ of the least powerful excitation signal $u(t)$ which leads to an identified model $G(z, \hat{\theta}_N)$ with a relative modeling error of less than 1 % at each ω i.e.,

$$|G(e^{j\omega}, \hat{\theta}_N) - G(e^{j\omega}, \theta_0)| < 0.01 \; |G(e^{j\omega}, \hat{\theta}_N)| \quad \forall \omega \tag{20}$$

Even though it is not directly related to control objectives, the above accuracy constraint has the form (1) considered in this paper with $r_{adm}(\omega) = 0.01 \; |G(e^{j\omega}, \hat{\theta}_N)|$. This problem can thus be easily transformed into the LMI optimization problem (13).

In order to be able to solve this problem, we will need an initial estimate of the true system. Here, this initial estimate has been obtained using open-loop identification with a white noise $u(t)$ of length 100 and variance 0.1. This initial estimate is required to replace θ_0 and σ_e^2 in the expression (6) of the bound $r_u(\omega)$ on the modeling error, but also to replace $\hat{\theta}_N$ in the expression of $r_{adm}(\omega)$:

$$r_{adm}(\omega) \approx 0.01 \; |G(e^{j\omega}, \theta_{init})|$$

Note that we here choose $\alpha = 2.45$ in the expression (6) of the bound $r_u(\omega)$ on the modeling error. By doing so, we ensure that this bound is valid at a probability of 95 %. The chosen parametrization for the power

spectrum $\Phi_u(\omega)$ is the parametrization (14). We have then solved the optimization problem (13) for two choices of m in this parametrization: $m = 20$ (flexible spectrum) and $m = 0$ (flat spectrum). The optimal spectra $\Phi_u(\omega)$ under these circumstances are denoted by $\Phi_{u,opt,flex}(\omega)$ when $m = 20$ and $\Phi_{u,opt,white}(\omega)$ when $m = 0$ and are depicted in Figure 5. The spectrum $\Phi_{u,opt,flex}(\omega)$ corresponds to a signal of power $\mathcal{P}_u = 0.25$ while the spectrum $\Phi_{u,opt,white}(\omega)$ corresponds to a signal of power $\mathcal{P}_u = 1.7$. This means that, if we want to identify a model of the true system satisfying (20) with a white noise input signal $u(t)$ of length $N = 500$, the power of this white noise should be of at least 1.7. On the other hand, if we shape the spectrum of the input signal appropriately such as done in $\Phi_{u,opt,flex}(\omega)$, the power of $u(t)$ required to achieve this accuracy can be reduced to 0.25 i.e. a reduction by a factor 6.
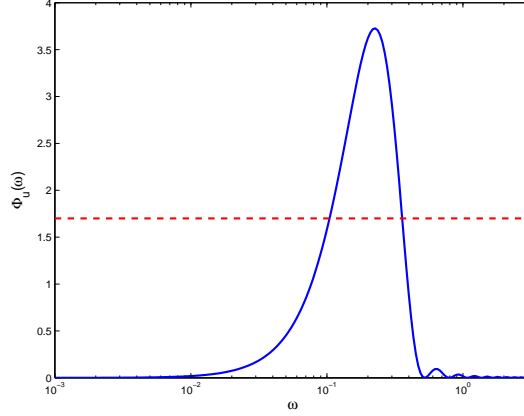
Figure 5: Optimal spectra $\Phi_{u,opt,flex}(\omega)$ i.e. when $m = 20$ (blue solid) and $\Phi_{u,opt,white}(\omega)$ i.e. when $m = 0$ (red dashed)

This result is based on the initial estimate θ_{init}. Using this initial estimate instead of the true parameter vector could lead to errors. We will verify that it is not the case here. To remain short, we do that uniquely in the case of $\Phi_{u,opt,flex}(\omega)$. We have thus generated a realization of length $N = 500$ having the optimal spectrum $\Phi_{u,opt,flex}(\omega)$. This realization is depicted in Figure 6. This realization is applied to the true system (8) and a model $G(z, \hat{\theta}_N)$ is identified. The bound $r_u(\omega)$ on the modeling error that can be computed with the covariance matrix of $\hat{\theta}_N$ must be by construction the largest possible while remaining at each ω under the threshold $0.01 \ |G(e^{j\omega}, \hat{\theta}_N)|$ (i.e. $r_{adm}(\omega)$). It is indeed the case as can be Figure 7. At a first sight, it could be surprising that $r_u(\omega) \approx r_{adm}(\omega)$ only in high frequencies and not over the whole frequency range. However, the latter is in fact not possible since $r_u(\omega)$ cannot take any possible shape due to its dependence on the parametrization of the system via P_θ.

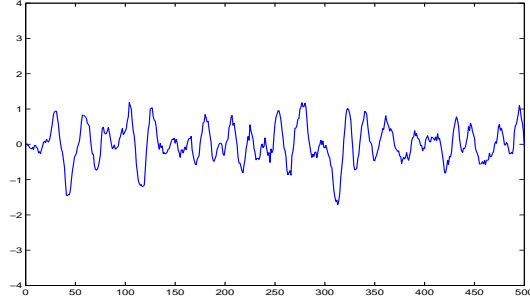Let us now finally spend some time analyzing why the optimal spectrum

13

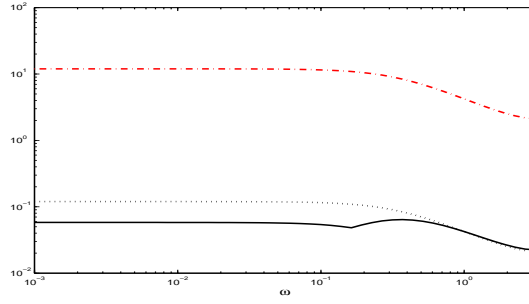Figure 6: Realization of length $N = 500$ of the optimal spectrum $\Phi_{u,opt,flex}(\omega)$

Figure 7: Modulus of $G(e^{j\omega}, \hat{\theta}_N)$ (red dashdot) with its modeling error $r_u(\omega)$ (black solid) as well as $r_{adm}(\omega) = 0.01 \ |G(e^{j\omega}, \hat{\theta}_N)|$ (back dotted)
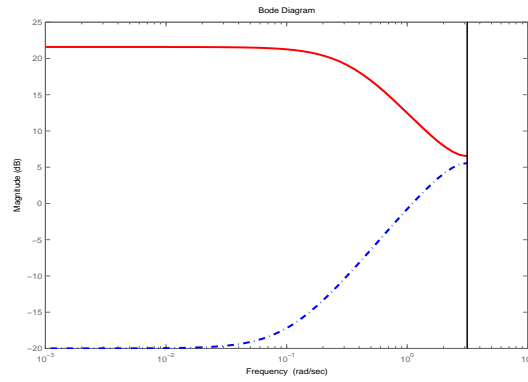

Figure 8: True system of Illustration 1: $|G_0(e^{j\omega})|$ (red solid) and of $|H_0(e^{j\omega})|$ (blue dashdot)

$\Phi_{u,opt,flex}(\omega)$ concentrates its power around $\omega = 0.15$ for this true system (see Figure 5). In fact, this can be explained by looking at the frequency response of G_0 and H_0 (see Figure 8). We see there that the optimal spectrum concentrates its power at a frequency in between the pole of G_0 (located at $\omega = 0.35$) and of the zero of H_0 (located at $\omega = 0.1$). As shown in the example of Section 2, a cosine at $\omega = 0.15$ rad/s is also a good choice. The optimal spectrum is thus dependent on the true system: this is the chicken-and-egg problem. Note however that we have determined this optimal spectrum without using the knowledge of the true system, but instead by using an initial estimate obtained after a short[3] identification experiment ($N = 100$) with relatively low input signal power ($= 0.1$).
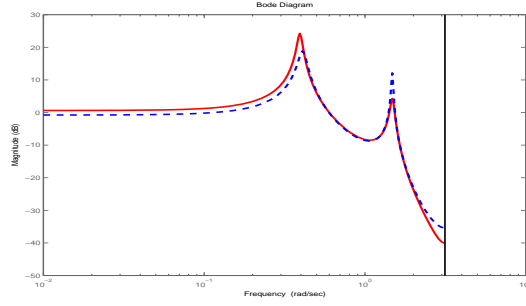
5.2 Ilustration 2

Figure 9: $|G_0(e^{j\omega})|$ (red solid) and $|G_{init}(e^{j\omega})|$ (blue dashed)

In the second illustration, we will consider the closed-loop optimal experiment design problem of Section 4. The considered true system is the following ARX system [18]:

$$y(t) = G_0(z)u(t) + H_0(z)e(t) = \frac{z^{-3}B_0(z)}{A_0(z)}u(t) + \frac{1}{A_0(z)}e(t) \qquad (21)$$

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 true system operates in closed loop with a controller C_{id} which has been designed using the 4-block H_∞ control design method of [10] and an initial estimate of the true system. This initial estimate of the true system originates from an open-loop experiment ($N = 100$, $\mathcal{P}_u = 10$) and is equal to:

$$y(t) = G_{init}(z)u(t) + H_{init}(z)e(t) = \frac{z^{-3}B_{init}(z)}{A_{init}(z)}u(t) + \frac{1}{A_{init}(z)}e(t) \qquad (22)$$

---

[3]This short initial identification experiment is not sufficient to obtain the accuracy $r_{adm}(\omega)$ considered in this example. Indeed, when $N = 500$, the minimal white noise power for this purpose is already equal to 1.7. Consequently, $\mathcal{P}_u = 0.1$ is certainly not sufficient when $N = 100$.

with $B_{init}(z) = 0.06413 + 0.2006z^{-1}$ and $A_{init}(z) = 1 - 1.975z^{-1} + 2.249z^{-2} - 1.907z^{-3} + 0.9223z^{-4}$. The estimate $\sigma_{e,init}^2$ of σ_e^2 is here 0.4893.

From a look at the parameters of G_{init} and G_0, we can see that G_{init} is not a particularly good model of G_0. This is also confirmed by looking at their Bode diagram in Figure 9. The bound $r_u(\omega)$ on the modeling error (see (6)) that can be computed from the covariance matrix corresponding to the initial estimate of the true system is represented in blue dashed in Figure 10. Note that we here choose $\alpha = 3.55$ to compute (6). By doing so, we ensure that this bound is valid at a probability of 95 %.

We want to improve the performance of the controller by re-identifying a model using a direct *closed-loop* experiment of duration $N = 500$. The desired accuracy $r_{adm}(\omega)$ for the to-be-identified model is based on robust control specifications on the sensitivity function and is represented in red solid in Figure 10. We observe that the initial estimate (22) is not accurate enough to respect this accuracy constraint.
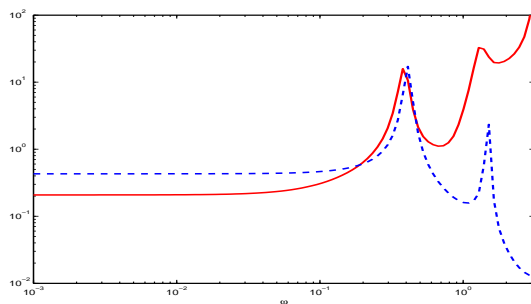

Figure 10: Modeling error $r_u(\omega)$ corresponding to the initial estimate (blue dashed) and accuracy threshold $r_{adm}(\omega)$ (red)

In order to shape optimally the spectrum $\Phi_r(\omega)$ of the excitation signal and obtain the spectrum inducing the lowest cost $\mathcal{J}_r = \mathcal{P}_{u_r} + \mathcal{P}_{y_r}$, we consider the LMI optimization problem (19). The initial estimate of the true system required to solve this LMI optimization problem is here the initial estimate given in (22). The parametrization for the power spectrum $\Phi_u(\omega)$ is here chosen similarly as in (14). We have solved the optimization problem (13) for two choices of m in this parametrization: $m = 10$ (flexible spectrum) and $m = 0$ (flat spectrum). The optimal spectra $\Phi_r(\omega)$ under these circumstances are denoted by $\Phi_{r,opt,flex}(\omega)$ when $m = 10$ and $\Phi_{r,opt,white}(\omega)$ when $m = 0$ and are depicted in Figure 11.

Realizations $r(t)$ of length $N = 500$ of these two spectra are depicted in Figure 12. These signals $r(t)$ have been applied to the closed-loop $[C_{id}\ G_0]$ and the induced disturbances $y_r(t)$ corresponding to these two signals $r(t)$ are depicted in Figure 13. We observe that the induced perturbation $y_r(t)$ is much smaller when $\Phi_{r,opt,flex}(\omega)$ is used i.e. when the spectrum is shaped

appropriately. These observations in the time-domain are of course also confirmed when looking at the cost function $\mathcal{J}_r = \mathcal{P}_{u_r} + \mathcal{P}_{y_r}$ of the optimization problem (19). When $m = 10$, the cost $\mathcal{J}_r$ induced by the application of an external signal having the spectrum $\Phi_{r,opt,flex}(\omega)$ is equal to 9.27 while, with $\Phi_{r,opt,white}(\omega)$, $\mathcal{J}_r = 22$. Consequently, if we want to identify a model of the true system having the desired accuracy with a white noise excitation signal $r(t)$ of length $N = 500$, the induced performance degradation $\mathcal{J}_r$ is more than two times larger than when we shape the spectrum $\Phi_r(\omega)$ appropriately.

We have also used the two realizations of Figure 12 to identify a model using direct closed-loop identification and we observe that, as expected, the bound $r_u(\omega)$ on the modeling error of these two models satisfy the accuracy threshold $r_{adm}(\omega)$ (see Figure 14). However, recall that the signal corresponding to $\Phi_{r,opt,flex}(\omega)$ induced a performance degradation $\mathcal{J}_r$ which is two times smaller than the performance degradation induced by one corresponding to $\Phi_{r,opt,white}(\omega)$.

Note that the reduction of the induced cost could be even larger in practice. Indeed, the white noise $r(t)$ in Figure 12 is a white noise of power $\mathcal{P}_r = 7.6$ and this power is the smallest power for a white noise that leads to a model satisfying the accuracy constraint. In practice, if we do not perform the optimal experiment design and perform an identification with an arbitrary white noise $r(t)$, two situations can occur: either we choose a white noise with a power larger than 7.6; which means that the induced cost $\mathcal{J}_r$ will even be larger than $\mathcal{J}_r = 22$ or we choose a white noise with a power smaller than 7.6; which means that the identified model will be too inaccurate for the design of an enhanced controller and that a new experiment will have to be performed.
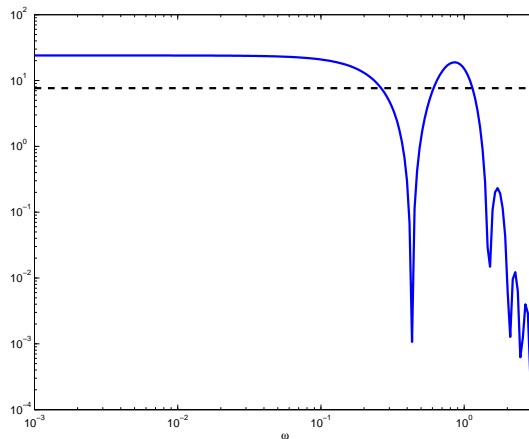

Figure 11: Optimal spectra $\Phi_{r,opt,flex}(\omega)$ i.e. when $m = 10$ (blue solid) and $\Phi_{r,opt,white}(\omega)$ i.e. when $m = 0$ (black dashed)
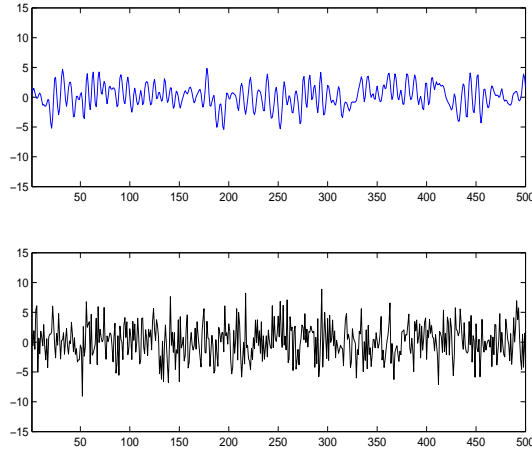
Figure 12: Realizations $r(t)$ corresponding to $\Phi_{r,opt,flex}(\omega)$ (top) and to $\Phi_{r,opt,white}(\omega)$ (bottom)
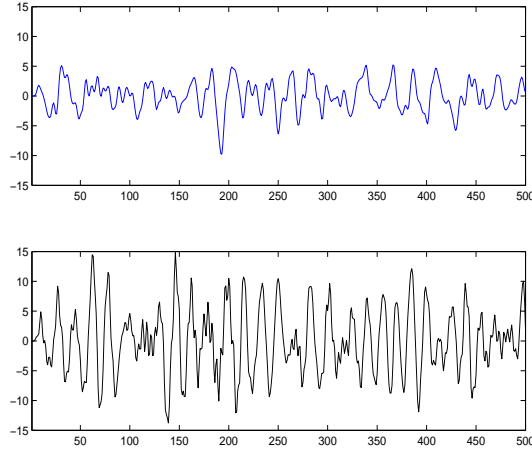

Figure 13: Induced output perturbation $y_r(t)$ corresponding to $\Phi_{r,opt,flex}(\omega)$ (top) and to $\Phi_{r,opt,white}(\omega)$ (bottom) for $N = 500$
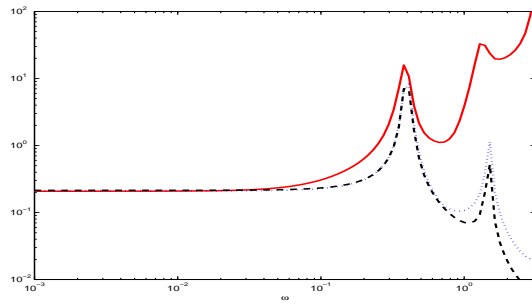

Figure 14: Comparison of $r_{adm}(\omega)$ (red solid) with the modeling error $r_u(\omega)$ obtained with $\Phi_{r,opt,flex}(\omega)$ (blue dotted) and with $\Phi_{r,opt,white}(\omega)$ (black dashed)

In order to illustrate in another way the advantage of shaping the power spectrum $\Phi_r(\omega)$ for the identification, we once again solve the optimization problem (19) with $m = 0$, but now for $N = 1000$ i.e. an identification two times longer than the previous one. The obtained optimal spectrum in this case is $\Phi_r(\omega) = 3.21 \; \forall \omega$ and corresponds thus to a white noise of power $\mathcal{P}_r = 3.21$. We have generated a realization of length $N = 1000$ of a such a white noise and compares it in Figure 15 with the realization of length $N = 500$ of the optimal spectrum $\Phi_{r,opt,flex}(\omega)$ (i.e. the realization in the top of Figure 12). The identification with the white noise lasts two times longer than the identification experiment with $\Phi_{r,opt,flex}(\omega)$, but both experiments will lead by construction to a model satisfying the accuracy constraint. Now, let us compare, in Figure 16, the perturbations y_r induced by these two experiments. In both experiments, the perturbations have a similar amplitude, but these perturbations disappears after 600 samples when $\Phi_{r,opt,flex}(\omega)$ is applied while they only disappears after 1100 samples when the white noise of power 3.21 is used. Shaping the power spectrum $\Phi_r(\omega)$ allows thus to obtain here the desired accuracy in a shorter time than when using just white noise. Note that the fact that the perturbations y_r do not stop directly after the end of the excitations $r(t)$ is due to the dynamics of the (closed-loop) system.
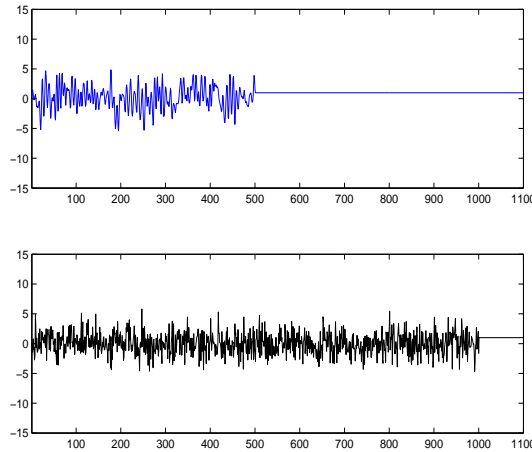

Figure 15: Excitation signal $r(t)$ of length $N = 500$ corresponding to to $\Phi_{r,opt,flex}(\omega)$ (top) vs. white noise $r(t)$ of power 3.21 and of length $N = 1000$ (bottom)

Let us now finally spend some time analyzing why the flexible spectrum having the shape of Figure 11 induces a much smaller cost than a white noise signal. The reason for that has to be found in the resonating nature of $G_0(z)$. The transfer function $G_0(z)$ has one important resonance peak at 0.4 rad/s (and a smaller one at 1.5 rad/s): see Figure 9. These resonances are also present in the transfer function $G_0 S_{id}$ generating the perturbation $y_r(t)$ from the external signal $r(t)$ (see (16)). A white noise excitation $r(t)$ excites
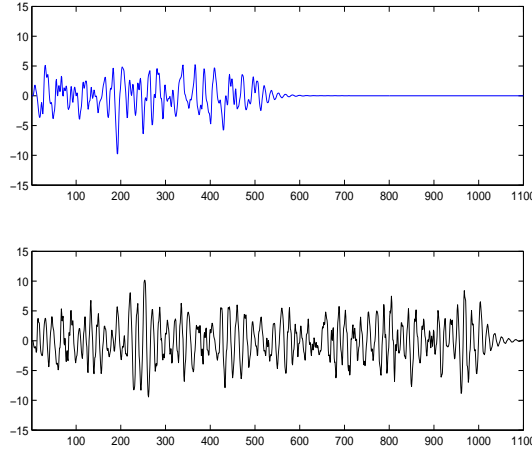
Figure 16: Induced output perturbation $y_r(t)$ corresponding to the signals $r(t)$ in Figure 15

these two resonant modes while the optimal flexible spectrum $\Phi_{r,opt,flex}(\omega)$ does not excite these modes since, as we can see in Figure 11, $\Phi_{r,opt,flex}(\omega)$ is quite small at the resonance frequencies. The optimal spectrum is thus once again dependent on the true system (chicken-and-egg problem). Note however that we have determined this optimal spectrum without using the knowledge of the true system, but instead by using a relatively poor initial estimate presenting the resonating behaviour of the system.

6   Concluding remarks

In this paper, we have presented the main ideas behind the least costly paradigm as well as illustrative examples. In order to conclude, we would like to give an overview of the different extensions. In the paper [4], the least costly paradigm is entirely developed in the parametric domain to avoid any conservatism and approximations. In this paper, the admissible uncertainty (represented by the threshold $r_{adm}(\omega)$ in this paper) is expressed based on (robust) control objectives in the H_∞ framework. These H_∞ results are extended to the multivariable case in [1]. Using the results in [5], the framework can be extended to H_2 formulations. The results in [16] allows to determine the admissible uncertainty for a large variety of other control objectives.

The assumption that the identification is performed in a full order model structure is relaxed in [2] where the bias contribution is also taken into account in the computation of $r_u(\omega)$.

References

[1] M. Barenthin, X. Bombois, H. Hjalmarsson, and G. Scorletti. Identification for control of multivariable systems: controller validation and experiment design via LMIs. *Automatica*, 44(12):3070–3078, 2008.

[2] X. Bombois and M. Gilson. Cheapest identification experiment with guaranteed accuracy in the presence of undermodeling. In *14th IFAC Symposium on System Identification, Newcastle*, 2006.

[3] X. Bombois, G. Scorletti, M. Gevers, R. Hildebrand, and P.M.J. Van den Hof. Cheapest open-loop identification for control. In *CD-ROM Proc. 43th Conference on Decision and Control, Bahamas*, 2004.

[4] X. Bombois, G. Scorletti, M. Gevers, P.M.J. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, 2006.

[5] X. Bombois, G. Scorletti, and H. Hjalmarsson. Identification for robust h2 deconvolution filtering. *Automatica*, 46(3):577–584, 2010.

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

[7] 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.

[8] B. Cooley and J. Lee. Control-relevant experiment design for multivariable systems by expansions in orthonormal bases. *Automatica*, 37:273–281, 2001.

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

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

[11] U. Forsell and L. Ljung. Some results on optimal experimental design. *Automatica*, 36(5):749–756, 2000.

[12] Lzló Gerencsér, Håkan Hjalmarsson, and Jonas Mårtensson. Identification of ARX systems with non-stationary inputs - asymptotic analysis with application to adaptive input design. *Automatica*, 45(3):623–633, March 2009.

[13] M. Gevers and L. Ljung. Optimal experiment design with respect to the intended model applications. *Automatica*, 22:543–554, 1986.

[14] 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.

[15] H. Hjalmarsson, M. Gevers, and F. De Bruyne. For model-based control design, closed-loop identification gives better performance. *Automatica*, 32(12):1659–1673, 1996.

[16] Håkan Hjalmarsson. System identification of complex and structured systems. *European Journal of Control*, 15(4):275–310, April 2009.

[17] H. Jansson and H. Hjalmarsson. Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, 50(10):1534–1549, October 2005.

[18] 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.

[19] K. Lindqvist. *On experiment design in identification of smooth linear systems.* PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2001.

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

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

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

A Linearisation of the constraint (10)

The constraint (10) is equivalent to: $T(e^{j\omega})P_\theta T(e^{j\omega})^T < \frac{r^2_{adm}(e^{j\omega})}{\alpha^2}I_2$. This new constraint can be rewritten as follows using the Schur complements [7]:

$$\begin{pmatrix} \frac{r^2_{adm}(e^{j\omega})}{\alpha^2}I_2 & T(e^{j\omega}) \\ T(e^{j\omega})^T & P_\theta^{-1} \end{pmatrix} > 0 \;\; \forall\omega$$

The latter is also equivalent with:

$$\begin{pmatrix} P_\theta^{-1} & T(e^{j\omega})^T \\ T(e^{j\omega}) & \frac{r^2_{adm}(e^{j\omega})}{\alpha^2}I_2 \end{pmatrix} > 0 \;\; \forall\omega$$

Another application of the Schur complement leads now to the result (12).