Robust optimal identification experiment design for multisine excitation

Xavier Bombois, Federico Morelli, H. Hjalmarsson, Laurent Bako, Kévin Colin

Robust optimal identification experiment design for multisine excitation

X. Bombois [a,b] F. Morelli [a] H. Hjalmarsson [c] L. Bako [a] K. Colin [a]

[a] *Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon, 36 avenue Guy de Collongue, Ecully, France*

[b] *Centre National de la Recherche Scientifique (CNRS), France*

[c] *Automatic control, KTH, Stockholm, Sweden*

Abstract

In least costly experiment design, the optimal spectrum of an identification experiment is determined in such a way that the cost of the experiment is minimized under some accuracy constraint on the identified parameter vector. Like all optimal experiment design problems, this optimization problem depends on the unknown true system, which is generally replaced by an initial estimate. One important consequence of this is that we can underestimate the actual cost of the experiment and that the accuracy of the identified model can be lower than desired. Here, based on an a-priori uncertainty set for the true system, we propose a convex optimization approach that allows to prevent these issues from happening. We do this when the to-be-determined spectrum is the one of a multisine signal.

Key words: Optimal experiment design, System identification

1 Introduction

We consider in this paper the problem of optimally designing the spectrum Φ_u of the excitation signal u of an open-loop identification experiment. By optimal spectrum, we here mean the spectrum yielding the smallest experiment cost while guaranteeing that the accuracy of the identified parameter vector of the plant transfer function is larger than a given threshold. We thus consider the *least costly experiment design* framework [5], but the approach can easily be adapted to other (dual) frameworks [10,17,13]. The experiment cost $\mathcal{J}$ can be defined as a linear combination of the power of the excitation signal u and of the power of the part of the output signal induced by u. The experiment cost will therefore be a function of the spectrum Φ_u, but also of the unknown true parameter vector θ_0 (we therefore denote the cost as $\mathcal{J}(\theta_0, \Phi_u)$). Likewise, the accuracy constraint will also depend on θ_0 and on Φ_u since the classical accuracy constraints are of the type $P^{-1}(\theta_0, \Phi_u) \geq R_{adm}$ where $P(\theta_0, \Phi_u)$ is the covariance matrix of the to-be-identified parameter vector (which depends on θ_0 and Φ_u) and R_{adm} a matrix reflecting the desired accuracy. The dependency of the optimal spectrum $\Phi_{u,opt}$ on the unknown true parameter vector θ_0 is the so-called *chicken-and-egg* issue encountered in optimal experiment design. This issue is generally circumvented by replacing θ_0 by an initial estimate $\hat{\theta}_{init}$ of θ_0 since, in this case, the optimal experiment design problem boils down to a convex optimization problem (see e.g. [10,5]). However, this approach has the drawback that the optimal spectrum is not guaranteed to yield the desired accuracy and that the experiment cost computed with $\hat{\theta}_{init}$ and $\Phi_{u,opt}$ can underestimate the actual experiment cost. These observations are at the root of the research area on *robust optimal experiment design* (see [19] for a good survey).

In robust experiment design, different lines of research have been considered. In [18], a spectrum that yields good accuracy for a very broad set of systems (also of different orders) is discussed. However, in the engineering literature, the most widely used approach is the one that consists in considering an uncertainty set U containing the unknown true parameter vector θ_0 (the so-called min-max design [19,16]). The optimal experiment design problem can then be formulated as determining the spectrum Φ_u minimizing the value of a scalar γ under the constraints that $\mathcal{J}(\theta, \Phi_u) \leq \gamma \ \forall \theta \in U$ and that $P^{-1}(\theta, \Phi_u) \geq R_{adm} \ \forall \theta \in U$. If we denote by γ_{opt} and $\Phi_{u,opt}$ the solution of this optimization problem, we have the guarantee that $P^{-1}(\theta_0, \Phi_{u,opt}) \geq R_{adm}$ and that γ_{opt} is an upper bound for the actual experiment cost $\mathcal{J}(\theta_0, \Phi_{u,opt})$. However, finding a tractable approach to deal with such a robustified optimal experiment design

problem is still an open research question. While this optimization problem can be exactly solved in very particular and simple situations (see e.g. [19,1]), the general approach when facing more complex systems is to replace the initial uncertainty set (containing an infinite number of elements) by a number n of grid points of this uncertainty set U (see e.g. [10,5,19]). Consequently, the cost constraint and the accuracy constraint over the set U in the robustified optimal experiment design problem are replaced each by n constraints (one for each grid point). Even though it is obviously better from a robustification point-of-view than just replacing θ_0 by one grid point i.e. $\hat{\theta}_{init}$, this relaxation of the original robustified optimal experiment design problem cannot yield the guarantees linked to the original problem and it can become computationally heavy for large values of n.

In [12,7,11], approaches are presented to uniquely tackle the robustified cost constraint $\mathcal{J}(\theta, \Phi_u) \leq \gamma$ $\forall \theta \in U$ (i.e. in the accuracy constraint, θ_0 is replaced by $\hat{\theta}_{init}$). However, these approaches all entail some approximation: a first-order approximation in [7], a second-order approximation in [12] and an approximation based on the unscented transform in [11].

Our main contribution in this paper is to present an approach in order to tackle the robust optimal experiment design problem without approximation. For this purpose, we observe that, except for its dependence on the to-be-determined spectrum, the robustified cost constraint and the robustified accuracy constraint are similar to constraints that are treated in robustness analysis. Based on this observation and on the separation of graph framework [20,9,14], we derive constraints that are linear in the decision variables of the optimal experiment design problem and that imply the original robustified cost and accuracy constraints. We do that for one of the most commonly used parametrization of the to-be-determined spectrum Φ_u i.e. the one corresponding to a multsine [10]. We however restrict attention to Box-Jenkins (BJ) model structures and to accuracy constraints on the parameters of the plant transfer functions (we can indeed not robustify the noise part of the covariance matrix using the tools presented in this paper).

Notations. The matrix

$$\begin{pmatrix} X_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & X_N \end{pmatrix}$$

will be denoted $diag(X_1, ..., X_N)$ if the elements X_i $(i = 1, ..., N)$ are scalar quantities while it will be denoted $bdiag(X_1, ..., X_N)$ if the elements X_i $(i = 1, ..., N)$ are matrices. In addition, the Fourier transform of a signal $x(t)$ is denoted $x(e^{j\omega})$, I_n represents the identity matrix of dimension n and $\otimes$, the Kronecker product. Finally, $\mathbf{0}$ represents a vector or a matrix containing only zeros and A^* is the conjugate transpose of the complex matrix A.

2 Identification

We consider a single-input single-output true system with input u and output y:

$$y(t) = G_0(z)u(t) + \underbrace{H_0(z)e(t)}_{=v(t)} \tag{1}$$

where $v(t) = H_0(z)e(t)$ is the disturbance acting on the system. In (1), $e(t)$ is a white noise with variance σ_e^2 and $G_0(z)$ and $H_0(z)$ are stable transfer functions. In addition, $H_0(z)$ is also assumed to be inversely stable and monic.

The true system (1) will be identified in a BJ model structure i.e., $\{G(z,\theta) = G(z,\rho),\ H(z,\theta) = H(z,\zeta) \mid \theta = (\rho^T, \zeta^T)^T \in \mathbf{R}^k\}$. The orders of $G(z,\theta)$ and $H(z,\theta)$ are chosen in such a way that there exists $\theta_0 = (\rho_0^T, \zeta_0^T)^T$ such that $G_0(z) = G(z,\theta_0) = G(z,\rho_0)$ and $H_0(z) = H(z,\theta_0) = H(z,\zeta_0)$. We will denote by k_G (resp. k_H) the dimension of ρ_0 (resp. ζ_0) and we have thus $k = k_G + k_H$.

If we apply a sequence $\{u(t) \mid t = 1, ..., N\}$ of spectrum Φ_u to (1) and collect the corresponding output $\{y(t) \mid t = 1, ..., N\}$, an estimate $\hat{\theta}_N$ of θ_0 can be deduced using prediction error identification [13]:

$$\begin{aligned} \hat{\theta}_N &= \arg\min_\theta \tfrac{1}{N} \sum_{t=1}^{N} \epsilon^2(t,\theta) \\ \epsilon(t,\theta) &= H^{-1}(z,\theta)\left(y(t) - G(z,\theta)u(t)\right) \end{aligned} \tag{2}$$

The estimate $\hat{\theta}_N$ is (asymptotically) normally distributed around θ_0 with a covariance matrix $P_\theta(\theta_0, \Phi_u)$ whose known expression is a function of θ_0 and of the input spectrum Φ_u used during the identification experiment [13]. It is important to note that $P_\theta^{-1}(\theta_0, \Phi_u)$ is a measure of the accuracy of $\hat{\theta}_N$ [10]. In general, we are mainly interested in the accuracy of the part $\hat{\rho}_N$ of the vector $\hat{\theta}_N = (\hat{\rho}_N^T, \hat{\zeta}_N^T)^T$ (the part that defines the model $G(z,\hat{\rho}_N)$ of $G(z,\rho_0)$). The covariance matrix P_ρ of $\hat{\rho}_N$ can be deduced from P_θ as follows:

$$P_\rho = (I_{k_G}\ \mathbf{0})\ P_\theta\ (I_{k_G}\ \mathbf{0})^T,$$

and its inverse has the following expression as a function of Φ_u and θ_0 [10,3]:

$$P_\rho^{-1}(\theta_0, \Phi_u) = \frac{N}{\sigma_e^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} F_u(e^{j\omega}, \theta_0) F_u^*(e^{j\omega}, \theta_0) \Phi_u(\omega)\, d\omega \tag{3}$$

with

$$F_u(z,\theta) = H^{-1}(z,\zeta) \frac{\partial G(z,\rho)}{\partial \rho} \tag{4}$$

In the sequel, we will suppose that an initial identification experiment has delivered an initial estimate $\hat{\theta}_{init} = (\hat{\rho}_{init}^T, \hat{\zeta}_{init}^T)^T$ with covariance matrix $P_{\theta,init}$.

Consequently, the following ellipsoid U_{init} is a η%-confidence region for the unknown true parameter vector θ_0:

$$U_{init} = \left\{ \theta \mid (\theta - \hat{\theta}_{init})^T P_{\theta,init}^{-1} (\theta - \hat{\theta}_{init}) \leq \chi \right\} \quad (5)$$

with χ such that $Pr(\chi^2(k) \leq \chi) = \eta$ (say 95 %). The significance of this set is that, if the initial experiment and the estimation of $\hat{\theta}_{init}$ is repeated, the true parameter θ_0 will belong to U_{init} in the fraction η of these essays. For this reason, U_{init} can be used as a description of the uncertainty of the initial estimate. From now on, we will assume that θ_0 indeed belongs to the ellipsoid U_{init} constructed based on the initial estimate $\hat{\theta}_{init}$ and its covariance matrix $P_{\theta,init}$.

We will suppose that the accuracy $P_{\rho,init}^{-1}$ of $\hat{\rho}_{init}$ is not sufficient for the purpose of the identified model ($P_{\rho,init} = (I_{k_G}\ \mathbf{0}) P_{\theta,init} (I_{k_G}\ \mathbf{0})^T$). In the sequel, we will deem an estimate of ρ_0 sufficiently accurate when the inverse of its covariance matrix satisfies $P_\rho^{-1} \geq R_{adm}$ for a given positive-definite matrix $R_{adm} \in \mathbf{R}^{k_G \times k_G}$.

In order to obtain a sufficiently accurate estimate of θ_0, we need to perform a second experiment that yields an estimate $\hat{\theta}_N = (\hat{\rho}_N^T, \hat{\zeta}_N^T)^T$ having the property that the covariance matrix $P_\rho(\theta_0, \Phi_u)$ of $\hat{\rho}_N^T$ is such that $P_\rho^{-1}(\theta_0, \Phi_u) \geq R_{adm}$ (Φ_u is the spectrum of the input signal used during the second experiment). Note that the initial and the second identification experiments can also be combined [13] and the inverse $P_\rho^{-1}(\theta_0, \Phi_u)$ of the covariance matrix corresponding to the second experiment must then satisfy: $P_\rho^{-1}(\theta_0, \Phi_u) + P_{\rho,init}^{-1} \geq R_{adm}$.

For the sequel, it will be important to make the following assumptions:

Assumption 1 *We restrict attention to parametrizations $G(z,\theta)$ and $H(z,\theta)$ that are rational functions of the parameter vector θ:*

$$G(z,\theta) = \frac{Z_N(z)\theta}{1 + Z_D(z)\theta} \quad (6)$$

$$H(z,\theta) = \frac{1 + Z_{N,H}(z)\theta}{1 + Z_{D,H}(z)\theta} \quad (7)$$

where $Z_N(z)$, $Z_D(z)$, $Z_{N,H}(z)$ and $Z_{D,H}(z)$ are row vectors of transfer functions. ■

Assumption 2 *The uncertainty U_{init} defined in (5) is small enough to guarantee that, like $G(z,\theta_0)$ and $H^{-1}(z,\theta_0)$, $G(z,\theta)$ and $H^{-1}(z,\theta)$ are stable transfer functions for all $\theta \in U_{init}$. Due to (4), this also implies that $F_u(z,\theta)$ is a vector of stable transfer functions for all $\theta \in U_{init}$.* ■

Note that the classical BJ parametrization used in prediction error identification satisfy Assumption 1 (see Appendix A). Moreover, we can easily verify whether a given U_{init} satisfies the property mentioned in Assumption 2 using the results in [4].

3 Optimal experiment design

As done in the *least costly experiment design* framework [5], we will design the spectrum Φ_u of the excitation signal u of the second experiment in such a way that the accuracy constraint is met with the least perturbation on the system (i.e. with the least identification cost). The perturbation on the system induced by u will be here measured by a linear combination of the power of the input signal and of the power of $\breve{y}(t) = G(z,\theta_0)u(t)$:

$$\mathcal{J}(\theta_0, \Phi_u) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 + \beta\ |G(e^{j\omega}, \theta_0)|^2\right)\ \Phi_u(\omega)\ d\omega \quad (8)$$

where β is an user-chosen constant that weighs the two terms in $\mathcal{J}$. We observe that the cost $\mathcal{J}(\theta_0, \Phi_u)$ is a function of the unknown θ_0 and of the chosen spectrum Φ_u. Based on the above expression, the optimal experiment design can be formulated as the problem of determining the power spectrum Φ_u of the second experiment which guarantees that $P_\rho^{-1}(\theta_0, \Phi_u) + P_{\rho,init}^{-1} \geq R_{adm}$ with the smallest cost $\mathcal{J}(\theta_0, \Phi_u)$. Like all other optimal experiment design problems, the above optimization problem unfortunately depends on the unknown true parameter vector. As mentioned in the introduction, we will here robustify this optimal experiment design problem using U_{init}. The robustified optimal experiment design problem is therefore:

$$\min_{\Phi_u,\ \gamma} \gamma \quad (9)$$

$$\text{such that } \mathcal{J}(\theta, \Phi_u) \leq \gamma\ \ \forall\ \theta \in U_{init} \quad (10)$$

$$\text{and } P_\rho^{-1}(\theta, \Phi_u) + P_{\rho,init}^{-1} \geq R_{adm}\ \ \forall\ \theta \in U_{init} \quad (11)$$

If we denote by $\Phi_{u,opt}^{orig}$ and γ_{opt}^{orig} the solution of this optimization problem, we have that $\gamma_{opt}^{orig} = \sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_{u,opt}^{orig})$. Moreover, the spectrum $\Phi_{u,opt}^{orig}$ is, by construction, the spectrum Φ_u leading to the smallest value of $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_u)$ while guaranteeing the robustified accuracy constraint (11). Since we assume that $\theta_0 \in U_{init}$, this robustified formulation ensures 1) that the a-priori unknown cost $\mathcal{J}(\theta_0, \Phi_{u,opt}^{orig})$ is smaller than γ_{opt}^{orig} and 2) that $P_\rho(\theta_0, \Phi_{u,opt}^{orig})$ is guaranteed to satisfy $P_\rho^{-1}(\theta_0, \Phi_{u,opt}^{orig}) + P_{\rho,init}^{-1} \geq R_{adm}$.

The above optimization problem will be a convex optimization problem if (10) and (11) can be transformed into two constraints that are linear in the decision variables Φ_u and γ. In the sequel, we will show that, as very often in robustness analysis theory, we cannot find tractable linear constraints that are equivalent to (10) and (11), but we can find one that implies (10) and another one that implies (11). This entails a certain conservatism. However, if we solve the optimization problem

with these alternative constraints and if we denote its solution by γ_{opt} and $\Phi_{u,opt}$, we still have the guarantee that

(1) $P_\rho^{-1}(\theta,\Phi_{u,opt}) + P_{\rho,init}^{-1} \geq R_{adm} \ \ \forall\ \theta \in U_{init}$ and thus $P_\rho^{-1}(\theta_0,\Phi_{u,opt}) + P_{\rho,init}^{-1} \geq R_{adm}$

(2) $\mathcal{J}(\theta_0,\Phi_{u,opt}) \leq \sup_{\theta\in U_{init}} \mathcal{J}(\theta,\Phi_{u,opt}) \leq \gamma_{opt}$.

In addition, we have also that γ_{opt} is an upper bound for γ_{opt}^{orig}, the solution of the original optimization problem (9)-(11).

We will derive the tractable alternative constraints discussed in the previous paragraph in the case of a commonly used parametrization of the to-be-determined spectrum Φ_u [10]. This spectrum parametrization corresponds to the spectrum of a multisine signal at fixed frequencies ω_m $(m = 1, ..., L)$ but with arbitrary amplitudes [1] :

$$\Phi_u(\omega) = \pi \sum_{m=1}^{L} c_m \ (\delta(\omega-\omega_m) + \delta(\omega+\omega_m)) \geq 0 \ \forall\omega \tag{12}$$

The positivity of $\Phi_u(\omega)$ for all ω can be imposed by the constraints $c_m \geq 0$ $(m = 1, \ldots, L)$. Using (12), the constraint (10) can be rewritten successively as follows:

$$\sum_{m=1}^{L} c_m \left(1 + \beta\ |G(e^{j\omega_m},\theta)|^2\right) \leq \gamma \ \ \forall\ \theta \in U_{init} \tag{13}$$

$$\left(\sum_{m=1}^{L} c_m\right) \ + \ \beta\ \mathcal{G}^*(\theta)\ \bar{C}\ \mathcal{G}(\theta) \leq \gamma \ \ \forall\ \theta \in U_{init} \tag{14}$$

with $\mathcal{G}(\theta) = (G(e^{j\omega_1},\theta), G(e^{j\omega_2},\theta), ..., G(e^{j\omega_L},\theta))^T$ and $\bar{C} = diag(c_1, c_2, ..., c_L)$. Using (3) and (12), the term $P_\rho^{-1}(\theta,\Phi_u)$ in (11) can also be rewritten as follows:

$$P_\rho^{-1}(\theta,\Phi_u) = \frac{N}{2\ \sigma_e^2} \sum_{m=1}^{L} c_m \left(F_u(e^{j\omega_m},\theta)F_u^*(e^{j\omega_m},\theta) + F_u(e^{-j\omega_m},\theta)F_u^*(e^{-j\omega_m},\theta)\right)$$

$$P_\rho^{-1}(\theta,\Phi_u) = \frac{N}{2\ \sigma_e^2} \sum_{i=1}^{k_G}\sum_{j=1}^{k_G} (e_i \otimes \mathcal{F}_i(\theta))^* \bar{C} (e_j \otimes \mathcal{F}_j(\theta)) + (e_i \otimes \mathcal{F}_j(\theta))^* \bar{C} (e_j \otimes \mathcal{F}_i(\theta)) \tag{15}$$

where e_i $(i = 1, ..., k_G)$ is a unit vector of dimension $1 \times k_G$ whose entries are all zero except the i^{th} entry which is equal to 1 and where $\mathcal{F}_i(\theta)$ $(i = 1, ..., k_G)$ is a complex vector of dimension $L \times 1$ defined as:

$$\mathcal{F}_i(\theta) = (F_{u,i}(e^{j\omega_1},\theta), F_{u,i}(e^{j\omega_2},\theta), ..., F_{u,i}(e^{j\omega_L},\theta))^T \tag{16}$$

[1] The amplitude A_m of the sinusoid at ω_m is given by $A_m = \sqrt{2c_m}$ $(m = 1, ..., L)$

with $F_{u,i}(z,\theta)$ the i^{th} entry of the vector $F_u(z,\theta)$.

Remark. In (15), the actual value of the variance σ_e^2 of the white noise e (see (1)) is generally unknown. However, we can also robustify the optimal experiment design problem against this uncertainty. For this purpose, we can replace σ_e^2 in (15) by $\sigma_{e,max}^2$ where $\sigma_{e,max}^2$ is the maximal value of the η%-confidence interval $[\sigma_{e,min}^2,\ \sigma_{e,max}^2]$ for σ_e^2 that can be constructed using the initial identification experiment (the one yielding $\hat{\theta}_{init}$) [13].

In the next two sections, we will derive, using robustness analysis tools, tractable alternatives for both (10) and (11) when the parametrization (12) is used for Φ_u. These tractable alternatives will be under the form of Linear Matrix Inequality (LMI) constraints [6] that are linear in γ and in the spectrum coefficients c_m $(m = 1, \ldots, L)$.

4 Tackling the robustified cost constraint using robustness analysis tools

4.1 Introduction

We will start by deriving a tractable alternative for (13)-(14). The constraints (13)-(14) are indeed equivalent to (10) when the parametrization (12) is used for Φ_u. In the paper [15], we have proposed the following tractable alternative for (13)-(14):

$$\sum_{m=0}^{L} c_m \left(1 + \beta\ \alpha_G(\omega_m)\right) \leq \gamma \tag{17}$$

with $\alpha_G(\omega_m) = \sup_{\theta\in U_{init}} |G(e^{j\omega_m},\theta)|^2$ (a computable quantity for each ω_m [4]). It is clear that (17) implies (13) and that it is linear in γ and in c_m $(m = 1, \ldots, L)$. However, the constraint (17) is rather conservative since $\alpha_G(\omega_m) = \sup_{\theta\in U_{init}} |G(e^{j\omega_m},\theta)|^2$ can be obtained for different θ at different frequencies ω_m $(m = 1, \ldots, L)$. In this paper, we propose an alternative approach which, as will be shown in the example section, will generally be less conservative since it will explicitly take into account the dependency on θ of the frequency response elements $G(e^{j\omega_m},\theta)$ in $\mathcal{G}(\theta)$ (see (14)).

4.2 Linear Fractional Transformation

An important step towards the developments of this result is to rewrite $\mathcal{G}(\theta)$ in the Linear Fractional Transformation (LFT) framework [21]. For this purpose, let us first observe that, due to Assumption 1, $\breve{y}(t) = G(z,\theta)u(t)$ can be written as the following LFT in θ involving the internal scalar signal q and the internal vector of signals p:

$$p = \theta\ q \ \ and \ \ \begin{pmatrix} q \\ \breve{y} \end{pmatrix} = \underbrace{\begin{pmatrix} -Z_D(z) & 1 \\ Z_N(z) & 0 \end{pmatrix}}_{M_G(z)} \begin{pmatrix} p \\ u \end{pmatrix} \tag{18}$$

Recall now that the Fourier transform $\breve{y}(e^{j\omega})$ of $\breve{y}(t) = G(z,\theta)u(t)$ is equal to $G(e^{j\omega},\theta)$ when $u(t)$ is equal to a pulse signal $\delta(t)$ (i.e. $u(e^{j\omega}) = 1$). Consequently, the frequency response $G(e^{j\omega},\theta)$ of $G(z,\theta)$ at one given frequency ω can also be deduced by solving for $\breve{y}(e^{j\omega})$ in the following system of equations:

$$p(e^{j\omega}) = \theta\ q(e^{j\omega})\ \ and$$
$$\begin{pmatrix} q(e^{j\omega}) \\ \breve{y}(e^{j\omega}) \end{pmatrix} = M_G(e^{j\omega}) \begin{pmatrix} p(e^{j\omega}) \\ 1 \end{pmatrix} \qquad (19)$$

Note that, in this system of equations, all Fourier transforms are well defined for all $\theta \in U_{init}$ due to Assumption 2.

Using the same reasoning, the vector $\mathcal{G}(\theta)$, containing the frequency response of $G(z,\theta)$ at the frequencies present in the spectrum (12), can be determined by solving for $\bar{y}$ in the system of equations (20) derived using (19):

$$\bar{p} = (I_L \otimes \theta)\ \bar{q}$$
$$\begin{pmatrix} \bar{q} \\ \bar{y} \end{pmatrix} = \underbrace{\begin{pmatrix} \bar{M}_{11,\mathcal{G}} & \bar{M}_{12,\mathcal{G}} \\ \bar{M}_{21,\mathcal{G}} & \bar{M}_{22,\mathcal{G}} \end{pmatrix}}_{\bar{M}_{\mathcal{G}}} \begin{pmatrix} \bar{p} \\ 1 \end{pmatrix} \qquad (20)$$

with $\bar{p} = (p^T(e^{j\omega_1}),...,p^T(e^{j\omega_L}))^T$, $\bar{q} = (q(e^{j\omega_1}),... ....,q(e^{j\omega_L}))^T$ and

$$\bar{M}_{11,\mathcal{G}} = -bdiag(Z_D(e^{j\omega_1}),...,Z_D(e^{j\omega_L}))$$
$$\bar{M}_{12,\mathcal{G}} = (1,...,1)^T$$
$$\bar{M}_{21,\mathcal{G}} = bdiag(Z_N(e^{j\omega_1}),...,Z_N(e^{j\omega_L}))$$
$$\bar{M}_{22,\mathcal{G}} = \mathbf{0}$$

Remark. It is important to note that the relation between $\bar{p}$ and $\bar{q}$ in (20) is in $I_L \otimes \theta = bdiag(\theta,...,\theta)$ i.e. a repetition of the same θ. This is due to the dependency on θ of the frequency response elements $G(e^{j\omega_m},\theta)$ in $\mathcal{G}(\theta)$. The approach in [15] (see Section 4.1) would in fact correspond to a relation between $\bar{p}$ and $\bar{q}$ of the form $diag(\theta(\omega_1),\theta(\omega_2),...,\theta(\omega_L))$ where $\theta(\omega_m) \in U_{init}$ for all ω_m $(m = 1,...,L)$, but can be different for each frequency ω_m $(m = 1,...,L)$. Indeed $\alpha_G(\omega_m)$ can be rewritten as $\sup_{\theta(\omega_m)\in U_{init}} |G(e^{j\omega_m},\theta(\omega_m))|^2$ $(m = 1,...,L)$.

4.3 Set of multipliers related to the uncertainty set U_{init}

Since we consider here (14), the parameter vector θ in the LFT for $\mathcal{G}(\theta)$ is restricted to be in the uncertainty set U_{init} (see (5)). In our approach, a necessary ingredient to find a tractable alternative for (14) is to associate, with the set U_{init}, a so-called set of multipliers. In a nutshell, the set of multipliers $\mathcal{A}_n$ that we will consider in this paper is an explicit and affine parametrization of the quadratic constraint satisfied by the graphs of the signals q_n and p_n when $p_n(t) = (I_n \otimes \theta)q_n(t)$ with $\theta \in U_{init}$ (n is an arbitrary integer such that $n \geq 1$) [20,9,14].

Definition 1 *Consider the set U_{init} defined in (5) satisfying Assumption 2 and an arbitrary integer $n \geq 1$. For each value of n, we define the set of multipliers $\mathcal{A}_n$ as a set of affinely parametrized Hermitian matrices A_n (of dimension $n(k+1)\times n(k+1)$) that all have the following property:*

$$\begin{pmatrix} I_n \\ I_n \otimes \theta \end{pmatrix}^T A_n \begin{pmatrix} I_n \\ I_n \otimes \theta \end{pmatrix} \geq 0\ \forall \theta \in U_{init} \qquad (21)$$

In other words, $A_n \in \mathcal{A}_n \Longrightarrow$ (21).

It is important to stress that the more extensive the parametrization of the set of multipliers, the smaller the conservatism discussed in Section 3 will be [20,9,14]. The set of multipliers $\mathcal{A}_n$ corresponding to U_{init} can be easily derived [2] from our previous contribution (see Proposition 2 of [2]).

In [2], the set of multipliers $\mathcal{A}_n$ is in fact developed for an uncertainty set of the form $U_{init} = \{\delta\theta \mid \delta\theta^T P^{-1}\delta\theta \leq \chi\}$. As shown in these papers, the use of this set of multipliers therefore entails the straightforward transformation of the LFT (18) (i.e. an LFT in θ) into an LFT in $\delta\theta = \theta - \hat{\theta}_{init}$. Another option is to adapt the multipliers of [2] to an uncertainty set U_{init} that is not centered at zero.

Remark. It is also to be noted that the set of multipliers in Definition 1 can also be derived for uncertainty sets U_{init} that are not ellipsoidal. In other words, the approach presented in this paper is valid for all uncertainty sets U_{init} for which the sets of multipliers $\mathcal{A}_n$ of Definition 1 can be constructed.

4.4 Robustified cost contraint

Taking a set of multipliers $\mathcal{A}_n$ with n equal to the number of frequencies in the to-be-determined spectrum (i.e., $n = L$) and considering the LFT representation (20) of $\mathcal{G}(\theta)$, we have now all the ingredients to derive a tractable alternative constraint for (14).

Proposition 1 *Consider an initial uncertainty set U_{init} (see (5)) satisfying Assumption 2 and the robust cost constraint (14) obtained when the spectrum Φ_u is parametrized as in (12). Consider the LFT representation (20) for $\mathcal{G}(\theta)$ as well as the set of multipliers $\mathcal{A}_L$ associated with U_{init} (see Definition 1 with $n = L$). Then, the constraint (14) holds for a given γ if we can find a matrix $A_L \in \mathcal{A}_L$ such that*

$$\mathcal{V}^*\ A_L\ \mathcal{V} + \mathcal{L}^*\ \bar{C}\ \mathcal{L} \leq \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\gamma - \left(\sum_{m=1}^{L} c_m\right)}{\beta} \end{pmatrix} \qquad (22)$$

[2] Note that, in [2], the notations $\mathcal{A}$ (resp. $\tilde{n}$) are used instead of $\mathcal{A}_n$ (resp. n).

where $\mathcal{L} = \begin{pmatrix} \bar{M}_{21,\mathcal{G}} & \bar{M}_{22,\mathcal{G}} \end{pmatrix}$ *and*

$$\mathcal{V} = \begin{pmatrix} \bar{M}_{11,\mathcal{G}} & \bar{M}_{12,\mathcal{G}} \\ I_{kL} & \mathbf{0} \end{pmatrix}$$

We observe that the matrix inequality (22) *is linear in* γ, A_L *and in the coefficients* c_m $(m = 1, ..., L)$ *present in* $\bar{C}$.

Proof. Let us consider (20) for a given $\theta \in U_{init}$ and let us consider the corresponding signals $\bar{p}$, $\bar{q}$ and $\bar{y} = \mathcal{G}(\theta)$. Let us then pre- and post-multiply the LMI constraint (22) with $(\bar{p}^*, 1)$ and $(\bar{p}^T, 1)^T$, respectively. Using (20) and $\bar{y} = \mathcal{G}(\theta)$, this yields:

$$\begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix}^* A_L \begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix} + \mathcal{G}^*(\theta)\bar{C}\mathcal{G}(\theta) \leq \frac{\gamma - \sum_{m=1}^{L} c_m}{\beta} \quad (23)$$

Since $\bar{p} = (I_L \otimes \theta)\bar{q}$, we can rewrite (23) as follows:

$$\bar{q}^* \begin{pmatrix} I_L \\ I_L \otimes \theta \end{pmatrix}^T A_L \begin{pmatrix} I_L \\ I_L \otimes \theta \end{pmatrix} \bar{q} + \mathcal{G}^*(\theta)\bar{C}\mathcal{G}(\theta) \leq \frac{\gamma - \sum_{m=1}^{L} c_m}{\beta} \quad (24)$$

The above reasoning can be done for any value of $\theta \in U_{init}$. In other words, for the matrix $A_L \in \mathcal{A}_L$ found by the optimization problem, (24) holds true for all $\theta \in U_{init}$. Consequently, using Definition 1 with $n = L$, we have therefore also that $\mathcal{G}^*(\theta)\bar{C}\mathcal{G}(\theta) \leq \frac{\gamma - \sum_{m=1}^{L} c_m}{\beta}$ for each $\theta \in U_{init}$; which is the desired result. ■

In Proposition 1 (but also later in this paper), when we speak of finding a matrix $A_L \in \mathcal{A}_L$, we more precisely mean finding the free parameters in the affine structure of the matrix A_L.

5 Robustified accuracy constraint

Taking inspiration from what has been done for the robustified cost constraint, we will now derive a tractable alternative for the accuracy constraint (11) when the parametrization (12) is used for Φ_u. Recall the expression (15) for $P_\rho^{-1}(\theta, \Phi_u)$ and let us observe that $F_u(z, \theta)$ (see (4)) is a rational function of θ due to Assumption 1. Consequently, we can find signals p_F and q_F such that $s(t) = F_u(z, \theta)u(t)$ can be expressed as:

$$\begin{aligned} p_F &= (I_f \otimes \theta)\ q_F \\ \begin{pmatrix} q_F \\ s \end{pmatrix} &= \underbrace{\begin{pmatrix} M_{11,F} & M_{12,F} \\ M_{21,F} & M_{22,F} \end{pmatrix}}_{M_F(z)} \begin{pmatrix} p_F \\ u \end{pmatrix} \end{aligned} \quad (25)$$

where $f = 3$ as shown in Appendix A. Note that $f = 2$ in the case where $H(z, \theta) = 1$ (OE model structure).

Using a similar reasoning as in Section 4.2, we can derive from (25) an LFT expression for $\bar{s}_i = \mathcal{F}_i(\theta)$ $(i = 1, ..., k_G)$ defined in (16). If we denote by $\bar{s} = (\bar{s}_1^T, \bar{s}_2^T, ..., \bar{s}_{k_G}^T)^T$, we have indeed

$$\begin{aligned} \bar{p}_F &= (I_{fL} \otimes \theta)\ \bar{q}_F \\ \begin{pmatrix} \bar{q}_F \\ \bar{s} \end{pmatrix} &= \underbrace{\begin{pmatrix} \bar{M}_{11,F} & \bar{M}_{12,F} \\ \bar{M}_{21,F} & \bar{M}_{22,F} \end{pmatrix}}_{\bar{M}_F(z)} \begin{pmatrix} \bar{p}_F \\ 1 \end{pmatrix} \end{aligned} \quad (26)$$

$$\begin{aligned} &\bar{M}_{11,F} = bdiag(M_{11,F}(e^{j\omega_1}), ..., M_{11,F}(e^{j\omega_L})) \\ &\bar{M}_{12,F} = (M_{12,F}(e^{j\omega_1}), ..., M_{12,F}(e^{j\omega_L}))^T \\ &\bar{M}_{21,F} = (\mathcal{H}_1^T, ..., \mathcal{H}_{k_G}^T)^T \quad \bar{M}_{22,F} = (\mathcal{K}_1^T, ..., \mathcal{K}_{k_G}^T)^T \\ &\mathcal{H}_i = bdiag(M_{21,F}^{i:}(e^{j\omega_1}), ..., M_{21,F}^{i:}(e^{j\omega_L})) \quad (i = 1, ..., k_G) \\ &\mathcal{K}_i = (M_{22,F}^{i}(e^{j\omega_1}), ..., M_{22,F}^{i}(e^{j\omega_L}))^T \quad (i = 1, ..., k_G) \end{aligned}$$

where $M_{21,F}^{i:}$ (resp. $M_{22,F}^{i}$) denotes the i^{th} line of $M_{21,F}$ (resp. the i^{th} entry of $M_{22,F}$). Note that, in this system of equations, all elements are well defined for all $\theta \in U_{init}$ due to Assumption 2.

We have then the following result that gives a tractable alternative for the robustified accuracy constraint (11).

Proposition 2 *Consider an initial uncertainty set* U_{init} *(see (5)) satisfying Assumption 2 and the robustified accuracy constraint (11) when the spectrum* Φ_u *is parametrized as in (12) and where, for this reason,* $P_\rho^{-1}(\theta, \Phi_u)$ *has the expression given in* (15)*. Consider the LFT (26) in* $I_{fL} \otimes \theta$ *which is the LFT representation for* $\bar{s} = (\mathcal{F}_1(\theta)^T, \mathcal{F}_2(\theta)^T, ..., \mathcal{F}_{k_G}(\theta)^T)^T$ *and consider the set of multipliers* $\mathcal{A}_{k_G fL}$ *associated with* U_{init} *(see Definition 1 with* $n = k_G fL$*). Then, the constraint (11) holds for a given* R_{adm} *and a given* $P_{\rho,init}$ *if we can find a matrix* $A_{k_G fL} \in \mathcal{A}_{k_G fL}$ *such that*[3]

$$\begin{aligned} &\frac{N}{2\,\sigma_e^2} \sum_{i=1}^{k_G} \sum_{j=1}^{k_G} (e_i \otimes \mathcal{X}_i)^* \bar{C} (e_j \otimes \mathcal{X}_j) + (e_i \otimes \mathcal{X}_j)^* \bar{C} (e_j \otimes \mathcal{X}_i) \\ &\quad + \left((P_{\rho,init}^{-1} - R_{adm}) \otimes \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \right) - \mathcal{M}^*\ A_{k_G fL}\ \mathcal{M} \geq 0 \end{aligned} \quad (27)$$

where $\mathcal{X}_i = \begin{pmatrix} \mathcal{H}_i & \mathcal{K}_i \end{pmatrix}$ $(i = 1, ...k_G)$ *and*

$$\mathcal{M} = \begin{pmatrix} I_{k_G} \otimes \begin{pmatrix} \bar{M}_{11,F} & \bar{M}_{12,F} \end{pmatrix} \\ I_{k_G} \otimes \begin{pmatrix} I_{fL} & \mathbf{0} \end{pmatrix} \end{pmatrix}$$

We observe that the matrix inequality (27) *is linear in* $A_{k_G fL}$ *and in the coefficients* c_m $(m = 1, ..., L)$ *present*

[3] The matrix made of zeros and a one in the second line of (27) has dimension $(fLk + 1) \times (fLk + 1)$.

in $\bar{C}$.

Proof. Let us consider (26) for a given $\theta \in U_{init}$ and let us consider the corresponding signals $\bar{p}_F$, $\bar{q}_F$ and $\bar{s} = (\mathcal{F}_1(\theta)^T, \mathcal{F}_2(\theta)^T, ..., \mathcal{F}_{k_G}(\theta)^T)^T$. Let us then pre- and post-multiply the LMI constraint (27) with $(I_{k_G} \otimes (\bar{p}_F^T, 1)^T)^*$ and $I_{k_G} \otimes (\bar{p}_F^T, 1)^T$, respectively. Using (26), (15) and the lemma in Appendix B , this yields:

$$P_\rho^{-1}(\theta, \Phi_u) + P_{\rho,init}^{-1} - R_{adm} - \ X(\theta) \geq 0 \qquad (28)$$

where $P_\rho^{-1}(\theta, \Phi_u)$ is given by (15) and $X(\theta)$ is a matrix given by

$$X(\theta) = \begin{pmatrix} I_{k_G} \otimes \bar{q}_F \\ I_{k_G} \otimes \bar{p}_F \end{pmatrix}^* A_{k_G fL} \begin{pmatrix} I_{k_G} \otimes \bar{q}_F \\ I_{k_G} \otimes \bar{p}_F \end{pmatrix}$$

Since $\bar{p}_F = (I_{fL} \otimes \theta) \ \bar{q}_F$, $X(\theta)$ can be rewritten as:

$$X(\theta) = (I_{k_G} \otimes \bar{q}_F)^* \ Y(\theta) \ (I_{k_G} \otimes \bar{q}_F)$$

$$Y(\theta) = \begin{pmatrix} I_{k_G fL} \\ I_{k_G fL} \otimes \theta \end{pmatrix}^T A_{k_G fL} \begin{pmatrix} I_{k_G fL} \\ I_{k_G fL} \otimes \theta \end{pmatrix}$$

The above reasoning can be done for any value of $\theta \in U_{init}$. In other words, for the matrix $A_{k_G fL} \in \mathcal{A}_{k_G fL}$ found by the optimization problem, (28) holds true for all $\theta \in U_{init}$. Using Definition 1 with $n = k_G fL$ and the Lemma in Appendix B, note also that:

$$X(\theta) \geq 0 \ \forall \theta \in U_{init}$$

Consequently, we have that $P_\rho^{-1}(\theta, \Phi_u) + P_{\rho,init}^{-1} \geq R_{adm}$ for each $\theta \in U_{init}$; which is the desired result. ■

6 Convex formulation of the optimal experiment design problem

Using Propositions 1 and 2, we can now straightforwardly derive a convex formulation for the robust optimal experiment design problem (9)-(11). This formulation will be under the form of a LMI optimization problem [6].

LMI formulation *Consider the parametrization (12) for the to-be-designed spectrum* Φ_u. *The LMI optimization problem has as decision variables a scalar* $\gamma > 0$, *coefficients* $c_m \geq 0$ *(*$m = 1, ..., L$*), a matrix* $A_L \in \mathcal{A}_L$ *and a matrix* $A_{k_G fL} \in \mathcal{A}_{k_G fL}$ *(see Definition 1) and consists in determining the smallest value of* γ *for which the LMI constraints (22) and (27) hold.*

As mentioned in Section 3, if we denote the solution of the above LMI optimization problem by γ_{opt} and $\Phi_{u,opt}$, γ_{opt} is an upper bound for the solution γ_{opt}^{orig} of the original robustified optimization problem (9)-(11) i.e. the minimal value of the cost that is required to guarantee the robust accuracy constraint (11). Moreover, γ_{opt} is an upper bound for $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_{u,opt})$ and, since we assume $\theta_0 \in U_{init}$, we have the guarantees 1) that $P_\rho^{-1}(\theta_0, \Phi_{u,opt}) + P_{\rho,init}^{-1} \geq R_{adm}$ and 2) that $\mathcal{J}(\theta_0, \Phi_{u,opt}) \leq \gamma_{opt}$.

Remark. In this paper, we have considered the least costly optimal experiment framework and we have supposed that the desired accuracy is represented by a given matrix R_{adm}. As mentioned in the introduction, the results of this paper can also be used for other optimal experiment design frameworks. For example, let us consider the classical E-optimality framework that, when robustified, consists in:

$$\max_{\Phi_u, \ \varepsilon} \ \varepsilon$$

$$\text{such that } \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) \ d\omega \leq \gamma$$

$$\text{and } P_\rho^{-1}(\theta, \Phi_u) + P_{\rho,init}^{-1} \geq \varepsilon I_{k_G} \ \ \forall \ \theta \in U_{init} \qquad (29)$$

where ε is a scalar decision variable and γ is a given bound on the power of the input signal. It is clear that (29) can be tackled with the tools of Section 5. Note that, here, $R_{adm} = \varepsilon I_{k_G}$ is not fixed, but a decision variable. This however does not pose a problem since the constraint (27) is also linear in R_{adm}.

7 Numerical illustration

In this section we present simulation results in order to show the effectiveness of our approach. We consider the following system $y(t) = G_0(z)u(t) + e(t)$ with $G_0(z) = \frac{1 \ z^{-1}}{1 - 0.7 \ z^{-1}}$ and $e(t)$ a white noise having variance $\sigma_e^2 = 1$. The true parameter vector θ_0 is thus $\theta_0 = \rho_0 = (1, -0.7)^T$. An initial estimate $\hat{\theta}_{init}$ of θ_0 and its covariance matrix $P_{\theta,init}$ have been obtained using an experiment of duration $N = 1000$ with a white noise input signal of variance 0.1: $\hat{\theta}_{init} = (0.904, -0.7161)^T$. This allows to build the initial uncertainty U_{init} (see (5)) that is a confidence region for θ_0 with probability level $\eta = 95\%$ ($\chi = 5.99$). This initial estimate does not satisfy the desired accuracy which is here that the standard deviation of the two parameters is smaller than two percents of their exact value. Based on this requirement, R_{adm} is chosen as the inverse of the following diagonal matrix $diag((0.02)^2, (0.014)^2)$ (see e.g., [8]).

We use the LMI formulation of Section 6 to determine the spectrum Φ_u of a second identification experiment of duration $N = 1000$ that will minimize a robustified version of the cost (10) under the robust accuracy constraint (11). Note that we here choose $\beta = 1$ in the expression of the cost (8). In (12), we choose $L = 9$ with ω_m ($m = 1, ..., 9$) covering the interval [0.1 3]. With these settings, the LMI optimization problem of Section 6 yields $\gamma_{opt} = 5.4948$ and an optimal multisine of spectrum $\Phi_{u,opt}$ for which all the

amplitudes are negligible except the ones at $\omega = 0.5$ and $\omega = 1$. By construction, we know that, if an excitation signal of spectrum $\Phi_{u,opt}$ is used in the second identification experiment, the obtained accuracy will be satisfactory (i.e. $P_\rho^{-1}(\theta_0, \Phi_{u,opt}) + P_{\rho,init}^{-1} \geq R_{adm}$) and the cost $\mathcal{J}(\theta_0, \Phi_{u,opt})$ of this second identification experiment will be such that $\mathcal{J}(\theta_0, \Phi_{u,opt}) \leq \gamma_{opt} = 5.4948$. The value γ_{opt} is indeed an upper bound for $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_{u,opt})$ as discussed in Section 3 and thus also an upper bound for the a-priori unknown cost $\mathcal{J}(\theta_0, \Phi_{u,opt})$ which is here equal to 4.7.

In order to check the conservatism linked to the proposed LMI formulation, we will compare the above result with the one that is obtained using the gridding approach for robust optimal experiment design. This approach considers the following optimization problem:

$$\min_{\Phi_{u,g},\, \gamma_g} \gamma_g$$

$$\text{such that } \mathcal{J}(\theta_i, \Phi_{u,g}) \leq \gamma_g \quad \forall\, \theta_i \in \Theta_n$$

$$\text{and } P_\rho^{-1}(\theta_i, \Phi_{u,g}) + P_{\rho,init}^{-1} \geq R_{adm} \quad \forall\, \theta_i \in \Theta_n$$

where Θ_n is a set containing n grid points θ_i $(i = 1, ..., n)$ such that $\theta_i \in U_{init}$. We have solved the above optimization problem for $n = 25$ using the same spectrum parametrization as above (same number L of frequencies and same frequencies ω_m) and we have obtained $\gamma_{g,opt} = 5.4866$ and a spectrum $\Phi_{u,g,opt}$ which has also contributions at two frequencies $\omega = 0.5$ and $\omega = 1$, but with (slightly) different amplitudes.

As mentioned in Sections 3 and 6, γ_{opt} (obtained with the approach proposed in this paper) is an upper bound for the solution γ_{opt}^{orig} of the original robustified optimization problem (9)-(11). It is also clear that the value $\gamma_{g,opt}$ obtained with the gridding approach is a lower bound for the same quantity. We thus observe that, in this example, the upper bound $\gamma_{opt} = 5.4948$ is almost equal to the lower bound $\gamma_{g,opt} = 5.4866$. The conservatism is thus very limited in this example. It is also important to note that, unlike the gridding approach, the approach proposed in this paper gives the guarantee that, with $\Phi_{u,opt}$, the robustified accuracy constraint (11) will be respected and also gives a guaranteed upper bound (i.e., γ_{opt}) on both $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_{u,opt})$ and the actual cost $\mathcal{J}(\theta_0, \Phi_{u,opt})$ of the second experiment.

To verify this property, we have generated 1000 grid points θ_i in U_{init} and we have computed $\mathcal{J}(\theta_i, \Phi_{u,opt})$ and $P_\rho^{-1}(\theta_i, \Phi_{u,opt})$ for these 1000 grid points. For all these grid points, we have indeed observed that $P_\rho^{-1}(\theta_i, \Phi_{u,opt}) + P_{\rho,init}^{-1} \geq R_{adm}$. We have also observed that the smallest eigenvalue of $P_\rho^{-1}(\theta_i, \Phi_{u,opt}) + P_{\rho,init}^{-1} - R_{adm}$ is, for one of these grid points θ_i, equal to 0.0004 (and thus very close to zero) and that the cost $\mathcal{J}(\theta_i, \Phi_{u,opt})$ is equal to 5.4948 for one of these grid points. This once again confirms that the conservatism of our approach is very limited in this example. Let us now compute, for these 1000 grid points, $P_\rho^{-1}(\theta_i, \Phi_{u,g,opt})$ with the spectrum $\Phi_{u,g,opt}$ obtained with the gridding approach. Here, 82 of the θ_i led to a matrix $P_\rho^{-1}(\theta_i, \Phi_{u,g,opt})$ for which $P_\rho^{-1}(\theta_i, \Phi_{u,g,opt}) + P_{\rho,init}^{-1} \geq R_{adm}$ is not satisfied. This shows the clear advantage of the approach of this paper upon the gridding approach.

Finally, let us illustrate the discussion in Section 4.1. For this purpose, we will compare the upper bound for $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_u)$ given by the left-hand side of (17) and the one corresponding to the LMI formulation proposed in this paper. We will do that for $\Phi_{u,opt}$. We know that the upper bound for $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_{u,opt})$ obtained using the tools proposed in this paper is $\gamma_{opt} = 5.4948$. Let us now compute the upper bound $\gamma_\alpha = \sum_{m=1}^{L} c_{m,opt}\,(1 + \alpha_G(\omega_m))$ for $\sup_{\theta \in U_{init}} \mathcal{J}(\theta, \Phi_{u,opt})$ (see (17)). This yields $\gamma_\alpha = 5.7408$, which shows that the approach discussed in Section 4.1 is more conservative.

8 Conclusion

In this paper, we have presented a convex relaxation that allows to robustify the least costly optimal experiment design problem using an initial uncertainty set for the unknown true parameter vector θ_0. This robustification is obtained using tools from robustness analysis. In this paper, we have restricted attention to multisine excitation signals. In the future, we will investigate whether the least costly optimal experiment design problem can also be robustified when the to-be-determined excitation spectrum is the one of a filtered white noise. This spectrum parametrization is indeed also a commonly used parametrization in optimal experiment design [10]. We will also investigate how the results presented in this paper can be extended to tackle the robustification of more complex accuracy constraints (such as the ones presented in [5]).

References

[1] S. Abrashov, R. Malti, M. Moze, X. Moreau, F. Aioun, and F. Guillemard. Simple and Robust Experiment Design for System Identification Using Fractional Models. *IEEE Transactions on Automatic Control*, 62(6):2648 – 2658, June 2017.

[2] 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.

[3] X. Bombois, B.D.O. Anderson, and G. Scorletti. Open-loop vs. closed-loop identification of box-jenkins systems in a least costly context. In *Proc. European Control Conference*, Kos, 2007.

[4] 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.

[5] 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.

[6] 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.

[7] M. Forgione, X. Bombois, and P.M.J. Van den Hof. Data-driven model improvement for model-based control. *Automatica*, 52:118–124, 2015.

[8] D. Ghosh, X. Bombois, J. Huillery, G. Scorletti, and G. Mercere. Optimal identification experiment design for LPV systems using the local approach. *Automatica*, 87:258–266, 2018.

[9] K.C. Goh and M.G. Safonov. Robust analysis, sectors and quadratic functionals. In IEEE, editor, *Proc. IEEE Conf. on Decision and Control*, New Orleans, Louisiana, 1995.

[10] 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.

[11] A. Kumar. *Optimal input signal design for plant friendly identification of process systems*. PhD thesis, Department of Chemical Engineering, Indian Institute of Technology, Madras, 2016.

[12] A. Kumar and S. Narasimhan. Robust plant friendly optimal input design. *IFAC Proceedings Volumes*, 46(32):553–558, 2013.

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

[14] A. Megretski and A. Rantzer. System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, 42:819–830, 1997.

[15] F. Morelli, X. Bombois, H. Hjalmarsson, L. Bako, and K. Colin. Optimal Experiment Design for the Identification of One Module in the Interconnection of Locally Controlled Systems. In *Proc. European Control Conference*, Naples, 2019, Draft available on HAL: hal-02063406.

[16] L. Pronzato and E. Walter. Robust experiment design via maximin optimization. *Mathematical Biosciences*, 89(2):161–176, 1988.

[17] C.R. Rojas, J.C. Agüero, J.S. Welsh, and G.C. Goodwin. On the equivalence of least costly and traditional experiment design for control. *Automatica*, 44(11):2706–2715, 2008.

[18] C.R. Rojas, J.C. Agüero, J.S. Welsh, G.C. Goodwin, and A. Feuer. Robustness in experiment design. *IEEE Transactions on Automatic Control*, 57(4):860–874, 2012.

[19] C.R. Rojas, J.S. Welsh, G.C. Goodwin, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 43(6):993–1008, 2007.

[20] M. G. Safonov. *Stability and Robustness of Multivariable Feedback Systems*. MIT Press, Cambridge, 1980.

[21] K. Zhou and J. Doyle. *Essentials of Robust Control*. Prentice Hall, Upper Saddle River, New Jersey, 1998.

A LFT representation of $F_u(z,\theta)$

The classical BJ paramatetrization is as follows:

$$G(z,\theta) = \frac{z^{-n_k}\left(b_0 + b_1 z^{-1} + ... + b_{n_b} z^{-n_b}\right)}{1 + f_1 z^{-1} + ... + f_{n_f} z^{-n_f}}$$

$$H(z,\theta) = \frac{1 + c_1 z^{-1} + ... + c_{n_c} z^{-n_c}}{1 + d_1 z^{-1} + ... + d_{n_d} z^{-n_d}}$$

with $\theta = (\rho^T, \zeta^T)^T \in \mathbf{R}^k$ with $\rho = (b_0, ..., b_{n_b}, f_1, ..., f_{n_f})^T$ and $\zeta = (c_1, ..., c_{n_c}, d_1, ..., d_{n_d})^T$ ($k_G = n_b + n_f + 1$, $k_H = n_c + n_d$, $k = k_G + k_H$). This parametrization satisfies Assumption 1. As an example, if $n_k = n_f = n_c = n_d = 1$ and $n_b = 0$ (i.e., $\theta = (b_0, f_1, c_1, d_1)^T$), we have $Z_N(z) = (z^{-1}\ 0\ 0\ 0)$, $Z_D(z) = (0\ z^{-1}\ 0\ 0)$, $Z_{N,H}(z) = (0\ 0\ z^{-1}\ 0)$ and $Z_{D,H}(z) = (0\ 0\ 0\ z^{-1})$. Moreover, we have also that $s(t) = F_u(z,\theta)u(t)$ is given by:

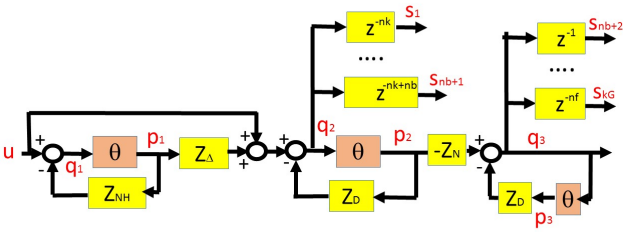

Fig. A.1. Representation of the vector $F_u(z,\theta)$ given in (A.1)

$$s(t) = \begin{pmatrix} s_1(t) \\ ... \\ s_{n_b+1} \\ s_{n_b+2} \\ ... \\ s_{k_G}(t) \end{pmatrix} = \begin{pmatrix} \frac{z^{-n_k}}{1+Z_D(z)\theta} \\ ... \\ \frac{z^{-(n_k+n_b)}}{1+Z_D(z)\theta} \\ \frac{-z^{-1}Z_N(z)\theta}{(1+Z_D(z)\theta)^2} \\ ... \\ \frac{-z^{-n_f}Z_N(z)\theta}{(1+Z_D(z)\theta)^2} \end{pmatrix} \frac{1 + Z_{D,H}(z)\theta}{1 + Z_{N,H}(z)\theta}\, u(t) \tag{A.1}$$

which is a rational function in θ. This rational function can be expressed as in (25) with $f = 3$ (see Figure A.1):

$$\underbrace{\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}}_{=p_F} = (I_3 \otimes \theta) \underbrace{\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}}_{=q_F}$$

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ s_1 \\ ... \\ s_{k_G} \end{pmatrix} = \underbrace{\begin{pmatrix} -Z_{N,H} & 0 & 0 & 1 \\ Z_\Delta & -Z_D & 0 & 1 \\ 0 & -Z_N & -Z_D & 0 \\ z^{-n_k}Z_\Delta & -z^{-n_k}Z_D & 0 & z^{-n_k} \\ ... & ... & .... & .... \\ 0 & -z^{-n_f}Z_N & -z^{-n_f}Z_D & 0 \end{pmatrix}}_{M_F(z)} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ u \end{pmatrix}$$

with $Z_\Delta(z) = Z_{D,H}(z) - Z_{N,H}(z)$. When $H(z,\theta) = 1$, the above LFT can be simplified and $f = 2$.

B Useful lemma for the proof of Proposition 2

Lemma 1 *Consider an Hermitian matrix $A = A^*$ of dimension $n \times n$ and a (complex) matrix B of dimension $n \times \tilde{n}$. Then, we have that*

$$A \geq 0 \Longrightarrow B^* A B \geq 0$$

Proof. $B^* A B \geq 0$ is equivalent to the fact that, for all complex vector x of dimension $\tilde{n}$,

$$x^* B^* A B x \geq 0$$

Denoting y the complex vector Bx of dimension n, the latter matrix inequality is equivalent to:

$$y^* A y \geq 0$$

which always holds since $A \geq 0$. ■