



Optimal identification experiment design for LPV systems using the local approach

Debarghya Ghosh, Xavier Bombois, Julien Huillery, Gérard Scorletti,
Guillaume Mercère

► To cite this version:

Debarghya Ghosh, Xavier Bombois, Julien Huillery, Gérard Scorletti, Guillaume Mercère. Optimal identification experiment design for LPV systems using the local approach. *Automatica*, 2018, 87, pp.258 - 266. 10.1016/j.automatica.2017.10.013 . hal-01720097

HAL Id: hal-01720097

<https://hal.science/hal-01720097>

Submitted on 25 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Optimal identification experiment design for LPV systems using the local approach

D. Ghosh^aX. Bombois^aJ. Huillery^aG. Scorletti^aG. Mercère^b

^aLaboratoire Ampère UMR CNRS 5005, Ecole Centrale de Lyon, 36 avenue Guy de Collongue, Ecully, France

^bUniversité de Poitiers, LIAS, 2 rue Pierre Brousse, Poitiers, France

Abstract

In this paper, we consider the problem of optimally designing the experimental conditions for LPV system identification with the local approach. Such an LPV system identification experiment is characterized by a number of local LTI identification experiments performed at constant values of the scheduling variable. The main contribution of this paper is to determine these constant values of the scheduling variable as well as the input spectra of the corresponding local LTI identification experiments in order to obtain a user-defined model accuracy with the least input energy.

Key words: Optimal identification experiment design, LPV system identification

1 Introduction

In this paper, we consider the problem of optimally designing the identification experiment leading to a model of a Linear Parameter Varying (LPV) system. A LPV system is a system whose parameters vary as a function of an external variable, the so-called scheduling variable [21]. When the scheduling variable is kept constant, the LPV system reduces to a Linear Time-Invariant (LTI) system whose dynamics depend on the chosen value for the constant scheduling variable. A constant scheduling variable is often called an *operating point* in the literature (see e.g. [3]).

It is important to note that efficient control design methods have been developed for LPV systems [1,2,19]. These control design methods deliver controllers whose parameters are, such as the system itself, a function of the scheduling variable. Methodologies have also been developed for the identification of such LPV systems. There exist two mainstream approaches for the identification of LPV systems, namely the local [24,15,21,20] and the global approaches [4,21,3,14,9]. In this paper, we will restrict attention to the so-called local approach for LPV system identification. In the first step of this local approach, the scheduling variable is kept constant at successive operating points and the corresponding local LTI models are identified using classical LTI identification [16]. In a second step, these identified models are interpolated to deliver a model of the LPV system. This second step is generally done by determining an estimate of the time-invariant coefficients θ^0 of the (e.g. polyno-

mial) function describing the dependence of the system parameters on the scheduling variable [21,3].

In order to design such a LPV system identification procedure in an optimal way, we need to determine optimal values 1) for the number of local LTI identification experiments that will be performed, 2) for the operating points at which these local LTI identification experiments will be performed and 3) for the input spectra that will be used in each of these local LTI identification experiments. These experimental conditions will be determined in such a way that we can guarantee a user-defined model accuracy with the least input energy.

The main contribution of this paper is to propose a convex optimization problem to tackle this rather complex optimal experiment design problem. For this purpose, an intermediate contribution is to determine the minimal variance estimate for the parameter vector θ^0 describing the LPV system (see above), as well as an expression for the covariance matrix P_θ of this estimate. The obtained expression for P_θ is such that its inverse is a summation of the contributions of each of the local LTI identification experiments. This property allows one to transform classical accuracy constraints into constraints that are linear in the input spectra used in the local experiments. Consequently, similarly as for optimal experiment design in the LTI case [6,10], one can use convex optimization to determine the optimal input spectra that have to be used at each operating point when the values of these operating points have been fixed a-priori. If this convex optimization problem is performed for a

set of operating points that constitutes a fine grid of the scheduling space, the obtained solution will typically be *sparse* i.e. many of the input spectra will be identically zero. The optimal operating points among this fine grid of the scheduling space are consequently those corresponding to a nonzero input spectra and local LTI identification experiments will only be performed at those operating points.

Such an optimal LPV identification experiment design framework is novel. However, we had already tackled, in [13,23], the sub-problem of selecting optimal operating points for LPV system identification using the local approach. In [13], this problem is tackled supposing that each local LTI identification experiment yields a perfect model of the local LTI system (i.e. the variance is supposed equal to zero). With this assumption, the optimal operating points are defined as those leading to the best approximation of a complex parameter dependence on the scheduling function by a polynomial function of low order. In [23], an iterative methodology is developed to determine a set of operating points for which the corresponding local LTI systems represent a sufficient sample of the different local dynamics of the LPV system. However, unlike in the present paper, those operating points are not determined as the solution of an optimization problem and the variance of the identified local LTI models is also neglected.

Notations. The matrix

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

will be denoted $\text{diag}(X_1, \dots, X_N)$ if the elements X_i ($i = 1, \dots, N$) are scalar quantities while it will be denoted $\text{bdiag}(X_1, \dots, X_N)$ if the elements X_i ($i = 1, \dots, N$) are matrices.

2 Identification Procedure

2.1 True System

We consider the identification of a Single-Input Single-Output (SISO) LPV system i.e. a SISO system whose parameters vary with time as a function of an exogenous variable, the so-called scheduling variable $p(t)$. We assume that $p(t)$ is a scalar signal that we can manipulate for identification purpose and that can vary in a given scheduling interval $[p_{\min}, p_{\max}]$. The considered SISO LPV system is as follows:

$$\begin{aligned} y(t) &= \check{y}(t) + v(t) \\ \text{with } A(z, \zeta^0(p(t))) \check{y}(t) &= B(z, \zeta^0(p(t))) u(t) \\ \text{and } D(z, \zeta^0(p(t))) v(t) &= C(z, \zeta^0(p(t))) e(t) \end{aligned} \quad (1)$$

with $u(t)$ the input signal, $y(t)$ the measured output and $e(t)$ a white noise signal of variance σ_e^2 . In (1), $A(z, \zeta^0)$, $B(z, \zeta^0)$, $C(z, \zeta^0)$ and $D(z, \zeta^0)$ are polynomials in the delay operator z^{-1} and having arbitrary orders: $A(z, \zeta^0) =$

$1 + a_1^0 z^{-1} + \dots + a_{n_a}^0 z^{-n_a}$, $B(z, \zeta^0) = z^{-n_k} (b_0^0 + b_1^0 z^{-1} + \dots + b_{n_b}^0 z^{-n_b})$, $C(z, \zeta^0) = 1 + c_1^0 z^{-1} + \dots + c_{n_c}^0 z^{-n_c}$ and $D(z, \zeta^0) = 1 + d_1^0 z^{-1} + \dots + d_{n_d}^0 z^{-n_d}$. The vector ζ^0 in (1) is a column vector of dimension $n \triangleq n_a + n_b + 1 + n_c + n_d$ containing the coefficients of the four polynomials describing the LPV system i.e. $a_1^0, \dots, b_0^0, \dots, d_{n_d}^0$. All the entries ζ_i^0 ($i = 1, \dots, n$) of ζ^0 depend on the scheduling variable $p(t)$ as follows:

$$\zeta_i^0(p(t)) = \sum_{j=0}^{n_p} \zeta_{i,j}^0 f_j(p(t)) \quad i = 1, \dots, n \quad (2)$$

where n_p is an arbitrary order and $f_j(p(t))$ ($j = 0, \dots, n_p$) correspond to a set of unisolvent basis functions, e.g. $f_j(p(t)) = p^j(t)$ (see e.g. [8]). In (2), we suppose that n_p is the same for each entry ζ_i^0 ($i = 1, \dots, n$) in order to simplify the notations (this is nevertheless not a necessity). Let us define by $\theta^0 = (\zeta_{1,0}^0, \zeta_{1,1}^0, \dots, \zeta_{n,n_p}^0)^T$, the vector of dimension $k \triangleq (n_p + 1)n$ which contains all the time-invariant coefficients of these polynomial expansions. The relation between θ^0 and the time-varying parameter vector $\zeta^0(p(t))$ can thus be expressed as follows:

$$\zeta^0(p(t)) = T(p(t)) \theta^0 \quad (3)$$

for a given matrix $T(p(t))$ of dimension $n \times k$ which is uniquely function of $f_j(p(t))$ ($j = 0, \dots, n_p$). In the sequel, we will assume that the structure of the LPV system is known i.e. that n_a, n_b, n_k, n_c and n_d are known in (1) and that n_p is known in (2). Under this assumption, the parameter vector θ^0 entirely describes the LPV system.

2.2 LPV system identification using the local approach

We will now show how we can deduce an estimate $\hat{\theta}$ of θ^0 using the local approach for LPV system identification and how we can evaluate the accuracy of this estimate $\hat{\theta}$. The first step of this local approach consists of performing a certain number (say M) of so-called local LTI identification experiments ($m = 1, \dots, M$) [21].

Let us describe one of these local LTI identification experiments. We first bring the scheduling variable to a given constant value (or operating point) \mathbf{p}_m . Then, while maintaining the scheduling variable at this operating point $p(t) = \mathbf{p}_m$, we apply an input sequence $u(t) = u_m(t)$ ($t = 1, \dots, N_p$) of spectrum Φ_{u_m} to the LPV system (1). The duration of the experiment is thus denoted by N_p . By measuring the corresponding output $y(t) = y_m(t)$, we obtain the data set $Z_m = \{u_m(t), y_m(t) \mid t = 1, \dots, N_p\}^1$. Since $p(t)$ is kept constant to \mathbf{p}_m during the local LTI experiment, the dynamics of the true system (1) can be represented by the following LTI sys-

¹ The time index is set back to one at the beginning of each local LTI identification experiment.

tem described by the *time-invariant* parameter vector $\zeta^0(\mathbf{p}_m) = T(\mathbf{p}_m)\theta^0$:

$$y(t) = G(z, \zeta^0(\mathbf{p}_m)) u(t) + H(z, \zeta^0(\mathbf{p}_m)) e(t) \quad (4)$$

with $G(z, \zeta) = B(z, \zeta)/A(z, \zeta)$ and $H(z, \zeta) = C(z, \zeta)/D(z, \zeta)$. For the sequel, it is important to observe that, due to the independent parametrization of the four polynomials, the vectors $\zeta^0(\mathbf{p}_m)$ and θ^0 can be split into a part that uniquely appears in the plant transfer function G and one that uniquely appears in the noise transfer function H . We will distinguish these two parts by the subscripts G and H , respectively. Consequently, (4) can be rewritten as $y(t) = G(z, \zeta_G^0(\mathbf{p}_m)) u(t) + H(z, \zeta_H^0(\mathbf{p}_m)) e(t)$ with:

$$\begin{pmatrix} \zeta_G^0(\mathbf{p}_m) \\ \zeta_H^0(\mathbf{p}_m) \end{pmatrix} = \begin{pmatrix} T_G(\mathbf{p}_m) & 0 \\ 0 & T_H(\mathbf{p}_m) \end{pmatrix} \begin{pmatrix} \theta_G^0 \\ \theta_H^0 \end{pmatrix} \quad (5)$$

Based on the data set Z_m and a full order model structure $\mathcal{M} = \{G(z, \zeta) \mid \zeta \in \mathbf{R}^n\}$ for the LTI representation (4), we determine an estimate $\hat{\zeta}_m$ of $\zeta^0(\mathbf{p}_m)$ using prediction error identification [16]:

$$\hat{\zeta}_m = \arg \min_{\zeta} \frac{1}{N_p} \sum_{t=1}^{N_p} \epsilon_m^2(t, \zeta) \quad (6)$$

with $\epsilon_m(t, \zeta) = H^{-1}(z, \zeta)(y_m(t) - G(z, \zeta)u_m(t))$. This estimate has the following asymptotic distribution: $\hat{\zeta}_m \sim \mathcal{N}(\zeta^0(\mathbf{p}_m), P_{\zeta, m})$ where the covariance matrix $P_{\zeta, m} > 0$ can be estimated from the data Z_m and $\hat{\zeta}_m$ [16]. Note furthermore that $P_{\zeta, m}$ has the following theoretical expression: $P_{\zeta, m} = \frac{\sigma_e^2}{N} (\bar{E}(\psi_m(t, \zeta^0(\mathbf{p}_m)) \psi_m^T(t, \zeta^0(\mathbf{p}_m))))^{-1}$ with $\psi_m(t, \zeta) = -\frac{\partial \epsilon_m(t, \zeta)}{\partial \zeta}$ and with $\bar{E}x(t) \triangleq \lim_{N_p \rightarrow \infty} \frac{1}{N_p} \sum_{t=1}^{N_p} Ex(t)$ (E is the expectation operator) [16]. Using this expression, we observe that the covariance matrix $P_{\zeta, m}$ has a block-diagonal structure: $P_{\zeta, m} = \text{bdiag}(P_{\zeta_G, m}, P_{\zeta_H, m})$ and the block-diagonal elements of $P_{\zeta, m}$ have the following expressions:

$$P_{\zeta_G, m}^{-1} = \frac{N_p}{\sigma_e^2} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} F_{u, m}(e^{j\omega}) F_{u, m}^*(e^{j\omega}) \Phi_{u_m}(\omega) d\omega \right) \quad (7)$$

$$P_{\zeta_H, m}^{-1} = N_p \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} F_{e, m}(e^{j\omega}) F_{e, m}^*(e^{j\omega}) d\omega \right) \quad (8)$$

with $F_{u, m}(z) = \frac{1}{H(z, \zeta_H^0(\mathbf{p}_m))} \frac{\partial G(z, \zeta_G)}{\partial \zeta_G} \Big|_{\zeta_G = \zeta_G^0(\mathbf{p}_m)}$ and $F_{e, m}(z) = \frac{1}{H(z, \zeta_H^0(\mathbf{p}_m))} \frac{\partial H(z, \zeta_H)}{\partial \zeta_H} \Big|_{\zeta_H = \zeta_H^0(\mathbf{p}_m)}$.

The above procedure is repeated for M different values of \mathbf{p}_m ($m = 1, \dots, M$) yielding M independent estimates $\hat{\zeta}_m$ and their respective covariance matrix $P_{\zeta, m}$.

We will assume that the durations of these M local LTI experiments will be all equal to N_p , but that the spectrum Φ_{u_m} of the input signal can be chosen differently in these M experiments. Indeed, the LTI representation (4) will be different for each \mathbf{p}_m and it makes thus sense to use a different input spectrum for these different LTI systems. In the sequel, we will use the shorthand notations $\mathbf{P}^M \triangleq \{\mathbf{p}_1, \dots, \mathbf{p}_M\}$ for the set of M operating points at which the local LTI identification experiments are performed and $\Phi(\mathbf{P}^M) \triangleq \{\Phi_{u_1}, \dots, \Phi_{u_M}\}$ for the set containing the input spectra used in these local LTI identification experiments.

The M local LTI identification experiments constitute the first step of the local approach for LPV system identification. The estimates $\hat{\zeta}_m$ of $\zeta^0(\mathbf{p}_m)$ ($m = 1, \dots, M$) obtained in this first step will then be used in a second step to determine an estimate $\hat{\theta}$ of the time-invariant parameter vector θ^0 using the mapping (3). In the literature, the estimate $\hat{\theta}$ is generally deduced using an ordinary least-squares criterion [15, 20, 24]. This simple approach is not optimal since it does not take into account the relative accuracy of the different estimates $\hat{\zeta}_m$ (i.e. the covariance matrices $P_{\zeta, m}$). Since, in this paper, the accuracy of $\hat{\theta}$ is a central aspect, we propose to instead use a generalized least-squares (or weighted least-squares) criterion (see the following theorem). Even though such an estimator is quite classical in estimation theory, it has, to the best of our knowledge, never been used before for the second step of the local LPV identification approach. **Theorem 1** Consider expression (2) with unisolvent basis functions $f_j(p(t))$ and the mapping (3). Suppose that M local LTI identification experiments have been performed at the operating points $\mathbf{P}^M \triangleq \{\mathbf{p}_1, \dots, \mathbf{p}_M\}$ yielding M independent estimates $\hat{\zeta}_m$ of $\zeta^0(\mathbf{p}_m) = T(\mathbf{p}_m)\theta^0$ ($m = 1, \dots, M$). Suppose furthermore that these M independent estimates have the following distribution: $\hat{\zeta}_m \sim \mathcal{N}(\zeta^0(\mathbf{p}_m), P_{\zeta, m})$. Then, if and only if the set \mathbf{P}^M contains at least $n_p + 1$ different operating points, we can deduce the following estimate $\hat{\theta}$ of θ^0 :

$$\hat{\theta} = \left(\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\zeta, m}^{-1} T(\mathbf{p}_m) \right)^{-1} \left(\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\zeta, m}^{-1} \hat{\zeta}_m \right) \quad (9)$$

and this estimate $\hat{\theta}$ has the following properties:

(1) $\hat{\theta} \sim \mathcal{N}(\theta^0, P_\theta)$ with a strictly positive-definite covariance matrix P_θ given by:

$$P_\theta = \left(\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\zeta, m}^{-1} T(\mathbf{p}_m) \right)^{-1}, \quad (10)$$

(2) the estimator (9) is the one leading to the smallest variance among all unbiased estimators using the given set of observations $\hat{\zeta}_m$ ($m = 1, \dots, M$).

Proof. Denoting $\hat{\zeta} \triangleq (\hat{\zeta}_1^T, \dots, \hat{\zeta}_M^T)^T$ and $\mathcal{T} \triangleq (T^T(\mathbf{p}_1), \dots, T^T(\mathbf{p}_M))^T$, we can write that $\hat{\zeta} \sim \mathcal{N}(\mathcal{T}\theta^0, P_\zeta)$ with P_ζ a strictly positive definite block-diagonal matrix given by $\text{bdiag}(P_{\zeta,1}, \dots, P_{\zeta,M})$. We can also equivalently write that $\hat{\zeta} = \mathcal{T}\theta^0 + \delta$ with $\delta \sim \mathcal{N}(0, P_\zeta)$. By multiplying this equation on both sides by the square root S of the inverse of P_ζ (i.e. $P_\zeta^{-1} = S^T S$), we obtain:

$$S\hat{\zeta} = S\mathcal{T}\theta^0 + \tilde{\delta} \quad (11)$$

with $\tilde{\delta} \triangleq S\delta$ having a distribution $\tilde{\delta} \sim \mathcal{N}(0, I_{nM})$ (I_{nM} denotes the identity matrix of dimension nM). Due to the distribution of $\tilde{\delta}$, the best linear unbiased estimator (BLUE) for θ^0 is the ordinary least-squares estimator based on the observation $S\hat{\zeta}$ and the regressor $S\mathcal{T}$ [12,16]. This estimator is given by

$$\hat{\theta} = (\mathcal{T}^T S^T S \mathcal{T})^{-1} (\mathcal{T}^T S^T S \hat{\zeta}) \quad (12)$$

and we observe that this estimator is equivalent to (9). In [16, Appendix 2], it is shown that, since $\tilde{\delta}$ is normally distributed, the estimate $\hat{\theta}$ has also a normal distribution. By inserting (11) in (12), we see that the expected value of $\hat{\theta}$ is equal to θ^0 and that its covariance matrix P_θ is equal to $(\mathcal{T}^T S^T S \mathcal{T})^{-1}$ which is equivalent to (10) in the statement of the theorem. The condition that \mathbf{P}^M contains at least $n_p + 1$ different operating points \mathbf{p}_m is an important condition since it guarantees that $\hat{\theta}$ can effectively be determined. Indeed, this condition is a necessary and sufficient condition for \mathcal{T} to be full column rank when (2) is constructed with unisolvant basis functions. A full column rank \mathcal{T} in turn ensures that $\mathcal{T}^T S^T S \mathcal{T} = \mathcal{T}^T P_\zeta^{-1} \mathcal{T}$ is strictly positive definite and thus invertible in (12). This also equivalently ensures that P_θ is strictly positive definite.

We still have to prove that the best linear unbiased estimator (9) is also the best unbiased estimator i.e. the estimator leading not only to the smallest variance among all linear unbiased estimators, but, in fact, among all unbiased estimators. For this purpose, we prove that the covariance matrix (10) of (12) corresponds to the Cramer-Rao bound which is the lower bound for the covariance matrix of all unbiased estimators. For this purpose, we recall that the Cramer-Rao bound is given by \mathbf{J}^{-1} with $\mathbf{J} = E \left(-\frac{\partial^2}{\partial \theta^2} \log \bar{f}(\theta) |_{\theta=\theta^0} \right)$ with E the expectation operator and $\bar{f}(\theta)$ the probability density function of the observation vector $S\hat{\zeta}$. In our case, since $\tilde{\delta} \sim \mathcal{N}(0, I_{nM})$, $\bar{f}(\theta) = \frac{1}{(2\pi)^{\frac{nM}{2}}} e^{-0.5(\hat{\zeta} - \mathcal{T}\theta)^T S^T S (\hat{\zeta} - \mathcal{T}\theta)}$. Thus, $\mathbf{J} = E(\mathcal{T}^T P_\zeta^{-1} \mathcal{T}) = \mathcal{T}^T P_\zeta^{-1} \mathcal{T}$ since \mathcal{T} is fully deterministic. We thus effectively see that \mathbf{J}^{-1} is given by (10) and thus that the second property given in the statement of the theorem holds. ■

Remark 1. For the properties of Theorem 1 to hold, N_p must be chosen sufficiently large since these properties require $\hat{\zeta}_m$ to be normally distributed and the latter is an asymptotic property in N_p .

Remark 2. It follows from the proof of Theorem 1 that, if the basis functions in (2) are not unisolvant, the result in Theorem 1 will still hold, but then under the condition that the considered set \mathbf{P}^M yields a full column rank matrix $\mathcal{T} \triangleq (T^T(\mathbf{p}_1), \dots, T^T(\mathbf{p}_M))^T$. In this case, choosing a set \mathbf{P}^M with $n_p + 1$ different operating points will not necessarily be sufficient to guarantee this property of \mathcal{T} .

The estimate $\hat{\theta}$ in (9) and its covariance matrix P_θ have some interesting structural properties that will be used for the design of the optimal experimental conditions. The first structural property is the following one. Suppose that, based on M_1 local LTI identification experiments at operating points $\{\mathbf{p}_1, \dots, \mathbf{p}_{M_1}\}$ (containing at least $n_p + 1$ different values), we have obtained, via (9), an estimate $\hat{\theta}_1$ for θ_0 with a certain covariance matrix P_{θ_1} that can be computed with (10). Suppose also that, subsequently, we have performed additional local LTI identification experiments at $\{\mathbf{p}_{M_1+1}, \dots, \mathbf{p}_{M_1+M_2}\}$ (those operating points should not be necessarily different from the ones in $\{\mathbf{p}_1, \dots, \mathbf{p}_{M_1}\}$). Then, the expression (10) for the covariance matrix P_θ of the estimator (9) corresponding to the $M = M_1 + M_2$ operating points $\{\mathbf{p}_1, \dots, \mathbf{p}_{M_1}, \mathbf{p}_{M_1+1}, \dots, \mathbf{p}_{M_1+M_2}\}$ can be rewritten as:

$$P_\theta = \left(P_{\theta_1}^{-1} + \sum_{m=M_1+1}^{M_1+M_2} T^T(\mathbf{p}_m) P_{\zeta,m}^{-1} T(\mathbf{p}_m) \right)^{-1} \quad (13)$$

where $P_{\zeta,m}$ is the covariance matrix of the estimate $\hat{\zeta}_m$ of the m^{th} experiment (such as in Theorem 1).

The second property pertains to the structure of P_θ (see (10)). First, observe that its inverse P_θ^{-1} is the summation of a contribution of each local LTI experiment. Moreover, P_θ has, like $P_{\zeta,m}$, a block-diagonal structure. Indeed, using (5), $P_\theta = \text{bdiag}(P_{\theta_G}, P_{\theta_H})$ with:

$$\begin{aligned} P_{\theta_G}^{-1} &= \sum_{m=1}^M T_G^T(\mathbf{p}_m) P_{\zeta_{m,G}}^{-1} T_G(\mathbf{p}_m) \\ P_{\theta_H}^{-1} &= \sum_{m=1}^M T_H^T(\mathbf{p}_m) P_{\zeta_{m,H}}^{-1} T_H(\mathbf{p}_m) \end{aligned} \quad (14)$$

with $P_{\zeta_{m,G}}$ and $P_{\zeta_{m,H}}$ as defined in (7) and (8), respectively. The matrix $P_{\theta_G}^{-1}$ is, via (7), a linear function of Φ_{u_m} ($m = 1, \dots, M$). This property will be important to formulate the optimal experiment design problem in a convex way.

Remark 3. As already mentioned, in the above identification procedure, we have supposed that we know the structure of the LPV system. For this purpose, we not only have to know the order of $G(z, \zeta)$ and $H(z, \zeta)$ in (4), but we also have to correctly know the order n_p in (2). Like for the choice of the orders of $G(z, \zeta)$

and $H(z, \zeta)$ [16], the choice of n_p can be validated a-posteriori by testing the whiteness of the residual signals $\epsilon_m(t, T(\mathbf{p}_m)\hat{\theta})$ ($m = 1, \dots, M$) and/or the whiteness of the residual vector $S(\hat{\zeta} - T\hat{\theta})$ (see the notation of (11)).

3 Optimal experiment design for LPV system identification using the local approach

3.1 Definition of the optimal experiment design problem

The local approach for LPV system identification presented in Section 2.2 involves a number of (important) experimental choices. Firstly, we have to determine the number M and the set of operating points $\mathbf{P}^M = \{\mathbf{p}_1, \dots, \mathbf{p}_M\}$ at which local LTI identification experiments will be performed. Second, we have also to determine the set $\Phi(\mathbf{P}^M) = \{\Phi_{u_1} \dots \Phi_{u_M}\}$ of spectra that will be used in these local LTI experiments. The duration N_p of the local experiment will be considered fixed a-priori (note nevertheless that N_p will have to be chosen larger than a given threshold (see Section 3.2)).

In the sequel, we will develop a method in order to make these experimental choices in an optimal way. An objective of this optimal design will be to guarantee that the estimate $\hat{\theta}$ satisfies some accuracy constraint(s). Like in [5, 17, 10, 18], the considered accuracy constraint(s) will be of the form $P_\theta^{-1} > R_{adm}$ where $R_{adm} \geq 0$ is a given matrix of dimension k . If R_{adm} is chosen as the diagonal matrix $\text{diag}\left(\frac{1}{\sigma_1^2}, \dots, \frac{1}{\sigma_k^2}\right)$, we impose that the standard deviation of each entry θ_i of $\hat{\theta}$ is smaller than σ_i . We can also impose constraints on the accuracy of the local LTI system corresponding to an arbitrary operating point \mathbf{p} . This local LTI system is described by (4) with a parameter vector $\zeta^0(\mathbf{p}) = T(\mathbf{p})\theta^0$. The estimate of this parameter vector is $\hat{\zeta}_p = T(\mathbf{p})\hat{\theta}$ and its covariance matrix is equal to $T(\mathbf{p})P_\theta T^T(\mathbf{p})$. We can e.g. impose an accuracy constraint on the frequency response of this local LTI model by slightly modifying the result in [5], but a constraint on the accuracy of $\hat{\zeta}_p$ can also be formulated as $(T(\mathbf{p})P_\theta T^T(\mathbf{p}))^{-1} > R_{adm, \zeta}$ where $R_{adm, \zeta} > 0$ is a given matrix of dimension n . This accuracy constraint can be rewritten, using Schur complements [7], as $P_\theta^{-1} > R_{adm}$ with $R_{adm} = T^T(\mathbf{p})R_{adm, \zeta}T(\mathbf{p}) \geq 0$. Such a constraint can of course be imposed for different values of \mathbf{p} . We may thus have multiple constraints $P_\theta^{-1} > R_{adm}^j$ ($j = 1, \dots, J$). We will assume that all $R_{adm}^j \geq 0$ have a block diagonal structure like P_θ i.e. $R_{adm}^j = \text{bdiag}(R_{adm, G}^j, R_{adm, H}^j)$. Consequently, the accuracy constraints that $\hat{\theta}$ will have to satisfy are:

$$P_{\theta_G}^{-1} > R_{adm, G}^j \quad j = 1, \dots, J \quad (15)$$

$$P_{\theta_H}^{-1} > R_{adm, H}^j \quad j = 1, \dots, J \quad (16)$$

with P_{θ_G} and P_{θ_H} as defined in (14).

We are now ready to formalize our optimal experiment design problem for the local LPV identification approach.

Optimal experiment design problem: *determine the experimental conditions M , \mathbf{P}^M , and $\Phi(\mathbf{P}^M)$ in such a way that the corresponding LPV identification experiment is the experiment that uses the least input energy $\mathcal{J} = N_p \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$ while guaranteeing that the covariance matrix P_θ of the estimate $\hat{\theta}$ satisfies the accuracy constraints (15) and (16).*

Like in the LTI case, the solution of the above optimal experiment design problem requires the knowledge of the true parameter vector θ^0 and the true noise variance σ_e^2 since P_{θ_G} and P_{θ_H} in (15)-(16) depend on θ^0 and σ_e^2 (via (7)-(8)). These unknown quantities have thus to be replaced by initial guesses obtained e.g. via an initial LPV identification experiment. We will discuss this further in Section 4.

It is also to be noted that, since $R_{adm}^j = \text{bdiag}(R_{adm, G}^j, R_{adm, H}^j) \geq 0$, the constraints (15) and (16) imply that P_θ^{-1} is strictly positive-definite. The set \mathbf{P}^M solving the above optimal experiment design problem will therefore contain at least $n_p + 1$ different operating points since the latter is a necessary (and sufficient) condition for (10) to be strictly positive definite (see Theorem 1).

3.2 Determination of the optimal spectra for fixed operating points

As we will see in the sequel, it is interesting to first consider the above optimal experiment design problem for the particular case where M and \mathbf{P}^M have been determined a-priori. The optimal spectra $\Phi_{opt}(\mathbf{P}^M)$ can then be determined as the solution of the following optimization problem:

$$\min_{\Phi(\mathbf{P}^M)} \mathcal{J} \quad \text{subject to (15) and (16)} \quad (17)$$

The properties of this optimization problem will be given in Theorem 2. Note beforehand that, in order to satisfy (16), the duration N_p of each local LTI identification experiment must in any case satisfy:

$$N_p > \max_j \lambda_{max} \left(\left(\sum_{m=1}^M X^H(\mathbf{p}_m) \right)^{-1} R_{adm, H}^j \right) \quad (18)$$

with $\lambda_{max}(A)$ the largest eigenvalue of A and $X^H(\mathbf{p}_m) \triangleq T_H^T(\mathbf{p}_m) \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} F_{e, m}(e^{j\omega}) F_{e, m}^*(e^{j\omega}) d\omega \right) T_H(\mathbf{p}_m)$ ($m = 1, \dots, M$).

Theorem 2 *Suppose that both M and the set $\mathbf{P}^M = \{\mathbf{p}_1, \dots, \mathbf{p}_M\}$ are fixed. Suppose also that this set \mathbf{P}^M contains at least $n_p + 1$ different operating points (see Theorem 1). Consider the optimization problem (17) and denote by $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ the optimal cost of this optimization problem for the fixed M and \mathbf{P}^M . Then, a necessary and sufficient condition for (17) to have a solution is the condition (18). When (18) holds, the optimization problem (17) is equivalent to the following convex optimization*

tion problem:

$$\min_{\Phi(\mathbf{P}^M)} \mathcal{J} \quad \text{subject to (15)} \quad (19)$$

If (18) holds, the optimal cost $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ of (17) is thus also the optimal cost of (19) and the value of this optimal cost does not depend on the chosen value for N_p . *Proof.* The constraints (16) do not depend on the decision variables Φ_{u_m} ($m = 1, \dots, M$) and have thus only to be satisfied. A necessary and sufficient condition for (16) to be satisfied is that (18) holds [18]. Consequently, when (18) holds, the optimization problem (17) is equivalent to (19). Note that both \mathcal{J} and $P_{\theta_G}^{-1}$ are linear in the products $N_p \Phi_{u_m}$ ($m = 1, \dots, M$). Consequently, for fixed N_p , the optimization problem (19) is convex. Moreover, if we make, in (19), the changes of variables $\bar{\Phi}_{u_m} = N_p \Phi_{u_m}$ ($m = 1, \dots, M$), we obtain an optimization problem that is independent of N_p . Consequently, the value of the optimal cost $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ will not depend on the chosen value for N_p . ■

The fact that the optimal cost $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ does not depend on the choice of N_p justifies why N_p is not considered as a design variable in this paper.

In order to solve the convex optimization problem (19) in an efficient manner, we need a parametrization for the spectra Φ_{u_m} ($m = 1, \dots, M$). Here, we will use the following parametrization [11] that corresponds to the spectrum of a signal that is generated by a white noise filtered by an arbitrary FIR filter of degree L (L is a user-chosen parameter):

$$\Phi_{u_m}(\omega) = c_{m,0} + 2 \sum_{i=1}^L c_{m,i} \cos(i\omega) \quad (20)$$

The positivity of (20) for all ω can be imposed by a Linear Matrix Inequality (LMI) constraint on the coefficients $c_{m,i}$ ($i = 0, \dots, L$) [11]. These coefficients will become the decision variables of the optimization problem (19). With (20), the objective function \mathcal{J} in (19) is given by $N_p \sum_{m=1}^M c_{m,0}$ [6,11] which is a linear function of the decision variables. Moreover, $P_{\zeta_{G,m}}^{-1}$ (see (7)) can also be expressed as a linear function of the decision variables $c_{m,i}$ ($i = 0, \dots, L$) [6]. This property in turn makes of $P_{\theta_G}^{-1}$ (see (14)) a linear function of the decision variables $c_{m,i}$ ($i = 0, \dots, L$) ($m = 1, \dots, M$). Consequently, with the parametrization (20), the optimization problem (19) can be solved using LMI optimization techniques [7].

3.3 Selection of the operating points

Let us now consider the original optimal experiment design problem of Section 3.1. For simplicity, we will assume that the local identification experiments can only be performed at the M_{grid} points in a fine grid $\mathbf{P}^{grid} = \{\mathbf{p}_1, \dots, \mathbf{p}_{M_{grid}}\}$ of the continuous scheduling interval $[p_{min}, p_{max}]$. In other words, we assume that the set

\mathbf{P}^M can only be chosen as a subset of \mathbf{P}^{grid} . Under the above assumption and using the notation of Theorem 2, the cost \mathcal{J}_{opt} of the optimal LPV identification experiment is the minimal value of $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ among all subsets \mathbf{P}^M of \mathbf{P}^{grid} that contains at least $n_p + 1$ operating points (see Theorem 1). We have then the following result: \mathcal{J}_{opt} can be obtained by solving (19) with $M = M_{grid}$ and with $\mathbf{P}^M = \mathbf{P}^{grid}$. In other words, $\mathcal{J}_{opt} = \mathcal{J}_{opt}(M_{grid}, \mathbf{P}^{grid})$. This result follows from the facts that the quantities which depend on the experiment in (19) are \mathcal{J} and $P_{\theta_G}^{-1}$ and that, for each LPV experiment with local LTI experiments at the points in a *strict* subset of \mathbf{P}^{grid} , we can find a (mathematically equivalent) experiment with local LTI experiments at all points in \mathbf{P}^{grid} that exactly lead to the same \mathcal{J} and $P_{\theta_G}^{-1}$. Indeed, let us consider an arbitrary experiment characterized by spectra $\Phi_{u_m}^*$ ($m = 1, \dots, M$) and a set $\mathbf{P}^{M,*}$ which is a strict subset of \mathbf{P}^{grid} (i.e. $M < M_{grid}$) and which contains at least $n_p + 1$ different operating points. The equivalent experiment with local LTI experiments at all points in \mathbf{P}^{grid} and that leads to the same \mathcal{J} and $P_{\theta_G}^{-1}$ is e.g. an experiment with input spectra equal to $\Phi_{u_m}^*$ for the \mathbf{p}_m that are also in $\mathbf{P}^{M,*}$ and equal to 0 for the other $M_{grid} - M$ points.

Since $\mathcal{J}_{opt} = \mathcal{J}_{opt}(M_{grid}, \mathbf{P}^{grid})$, a solution of the optimal experiment design problem of Section 3.1 is an experiment characterized by $M = M_{grid}$, $\mathbf{P}^M = \mathbf{P}^{grid}$ and the spectra obtained as the solution of (19) with $M = M_{grid}$ and $\mathbf{P}^M = \mathbf{P}^{grid}$. This said, it is nevertheless important to observe that, due to the particular structure of the optimization problem (19), its solution when $M = M_{grid}$ and $\mathbf{P}^M = \mathbf{P}^{grid}$ will generally be *sparse* i.e. many of the spectra Φ_{u_m} will be identically 0. Indeed, since \mathcal{J} must be minimized, power will only be injected at those operating points that allow to obtain the maximum information on θ^0 . This phenomenon can presumably also be explained by the fact that the cost function \mathcal{J} has a l_1 -norm structure and it is frequently observed that such cost functions, when minimized under convex constraints, generate a sparse solution (see e.g. [22]).

Suppose for further reference that the spectra Φ_{u_m} of the solution of (19) when $M = M_{grid}$ and $\mathbf{P}^M = \mathbf{P}^{grid}$ are nonzero for the \bar{M} operating points in the subset $\bar{\mathbf{P}}^M \triangleq \{\bar{\mathbf{p}}_1, \dots, \bar{\mathbf{p}}_{\bar{M}}\}$ of \mathbf{P}^{grid} and denote by $\bar{\Phi}(\bar{\mathbf{P}}^M)$ the set of these \bar{M} nonzero spectra ($\bar{M} \leq M_{grid}$). Observe that local LTI identification experiments for which $\Phi_{u_m} = 0$ neither contribute to the cost \mathcal{J} nor to the accuracy $P_{\theta_G}^{-1}$. Consequently, the LPV identification experiment characterized by \bar{M} , $\bar{\mathbf{P}}^M$ and $\bar{\Phi}(\bar{\mathbf{P}}^M)$ is also a solution of the optimal experiment design problem of Section 3.1 i.e. this experiment has also a cost \mathcal{J}_{opt} and satisfies the constraints (15). Since it will always be preferred to perform an LPV identification experiment with a small number of local LTI identification experiments and that performing local identifications with $\Phi_{u_m} = 0$ does not make sense, the LPV experiment character-

ized by \bar{M} , $\bar{\mathbf{P}}^{\mathbf{M}}$ and $\bar{\Phi}(\bar{\mathbf{P}}^{\mathbf{M}})$ will be the experiment that will be deemed *optimal* and that will eventually be performed.

Remark 4. To determine \bar{M} , $\bar{\mathbf{P}}^{\mathbf{M}}$ and $\bar{\Phi}(\bar{\mathbf{P}}^{\mathbf{M}})$, we thus solve the optimization problem (19) with $M = M_{grid}$ and $\mathbf{P}^{\mathbf{M}} = \mathbf{P}^{grid}$. This optimization problem has a solution for all values of N_p . However, as shown in Theorem 2, this solution is also the solution of (17) *only if* (18) holds. We can thus follow the above procedure for the selection of \bar{M} , $\bar{\mathbf{P}}^{\mathbf{M}}$ and $\bar{\Phi}(\bar{\mathbf{P}}^{\mathbf{M}})$ with an initial value of N_p . Once $\bar{\mathbf{P}}^{\mathbf{M}}$ determined, we can verify whether this initial N_p satisfies the condition (18). If that is not the case, N_p will have to be increased to a value $N_{p,2}$ satisfying (18) and the spectra in $\bar{\Phi}(\bar{\mathbf{P}}^{\mathbf{M}})$ scaled with a factor $\frac{N_p}{N_{p,2}}$.

Remark 5. When solving (19) with $M = M_{grid}$ and $\mathbf{P}^{\mathbf{M}} = \mathbf{P}^{grid}$ as proposed above, it can be of interest to add the following constraints to the optimization problem: $N_p c_{m,0} \leq \alpha_m$ ($m = 1, \dots, M_{grid}$) with α_m some user-chosen scalar constants. These constants $\alpha_m > 0$ ($m = 1, \dots, M_{grid}$) can reflect the maximal energy that is allowed to apply at each operating point \mathbf{p}_m . Moreover, choosing a relatively small α_m at operating points \mathbf{p}_m that are difficult to reach will favour the choice of other operating points. Note that, if we want to absolutely avoid some operating points, we can also remove those operating points from \mathbf{P}^{grid} . It is also to be noted that the additional constraints $N_p c_{m,0} \leq \alpha_m$ will (generally) modify the obtained solution \bar{M} , $\bar{\mathbf{P}}^{\mathbf{M}}$ and $\bar{\Phi}(\bar{\mathbf{P}}^{\mathbf{M}})$.

Remark 6. If, for a specific situation, the spectra of the solution of (19) with $M = M_{grid}$ and $\mathbf{P}^{\mathbf{M}} = \mathbf{P}^{grid}$ are nonzero at all M_{grid} operating points (i.e. we obtain a non-sparse solution), it is important to stress that the corresponding LPV experiment with M_{grid} local LTI experiments is nevertheless the solution of the optimal experiment design problem of Section 3.1 (under the assumption that $\mathbf{P}^{\mathbf{M}}$ must be chosen as a subset of \mathbf{P}^{grid}).

4 Dependence of the solution on θ^0 and σ_e^2

As in all experiment design problem, the optimal LPV identification experiment can only be determined using the procedure in the previous section if we know the true parameter vector θ^0 and the true noise covariance σ_e^2 . Indeed, the covariance matrix P_θ depends on θ^0 and σ_e^2 (via (7)-(8)). Since θ^0 and σ_e^2 are of course unknown, they must consequently be replaced by some estimates in the expression for P_θ .

Suppose for this purpose that we have performed a first LPV identification experiment using an arbitrary set \mathbf{P}^{M_1} and arbitrary spectra $\Phi(\mathbf{P}^{M_1})$. From the M_1 estimates $\hat{\zeta}_m$ deduced from this initial LPV experiment, we can deduce an initial estimate $\hat{\sigma}_{e,1}^2$ of σ_e^2 using the estimator $\hat{\sigma}_{e,1}^2 = \frac{1}{M_1} \sum_{m=1}^{M_1} \left(\frac{1}{N_p} \sum_{t=1}^{N_p} \epsilon_m^2(t, \hat{\zeta}_m) \right)$ as well as an initial estimate $\hat{\theta}_1$ of θ^0 via the estimator (9). We can therefore also deduce an estimate of the covariance matrix P_{θ_1} of θ_1 . If $P_{\theta_1}^{-1} > R_{adm}^j \forall j$, no need to proceed with optimal experiment design. However, if it is not

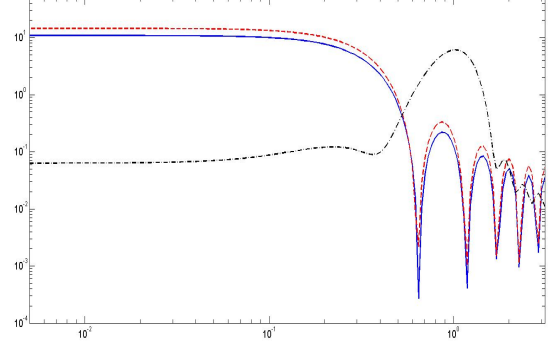


Fig. 1. Optimal power spectrum Φ_{u_m} for $\mathbf{p}_m = 0$ (blue solid), for $\mathbf{p}_m = 1.5$ (red dashed) and for $\mathbf{p}_m = 8$ (black dashdot)

the case, we will perform a second LPV experiment and we will design the local LTI identification experiments of this second LPV experiment in an optimal way. For this purpose, we will replace θ^0 by $\hat{\theta}_1$ and σ_e^2 by $\hat{\sigma}_{e,1}^2$ in the expression for P_θ and we will replace the constraints $P_\theta^{-1} > R_{adm}^j$ by $P_{\theta_1}^{-1} > R_{adm}^j - P_{\theta_1}^{-1}$ in the procedure of Section 3. Recall indeed that the covariance matrix P_θ corresponding to the combination of the initial LPV experiment (i.e. the one yielding $\hat{\theta}_1$, $\hat{\sigma}_{e,1}^2$ and P_{θ_1}) and the additional LPV experiment is given by (13).

Remark 7. The initial LPV experiment will also be generally used to determine optimal values for the orders n_a , n_b , n_c , n_d and n_p of the LPV model structure (using e.g. the procedure of Remark 3). Consequently, it is good practice to distribute the operating points in \mathbf{P}^{M_1} over the entire scheduling space and to choose M_1 not too small. Indeed, M_1 should be larger than $n_p + 1$ with the actual value of n_p in (2), value that remains uncertain when designing the initial LPV experiment.

5 Numerical example

We will now illustrate our results by considering a true LPV system having the form (1) and described by the following polynomials: $A(z, \zeta^0) = 1 + a_1^0 z^{-1}$, $B(z, \zeta^0) = b_0^0 z^{-1}$, $C(z, \zeta^0) = 1$ and $D(z, p(t)) = 1 + d_1^0 z^{-1}$. We further assume that the variance σ_e^2 of $e(t)$ in (1) is equal to 0.5 and that we have a linear dependence of the parameters on the scheduling variable $p(t)$ ($n_p = 1$ and $f_j(p(t)) = p^j(t)$): $a_1^0(p(t)) = -0.9 + 0.1 p(t)$, $b_0^0(p(t)) = 10 - 1 p(t)$ and $d_1^0(p(t)) = -0.7 + 0.08 p(t)$. We have thus $\zeta^0(p(t)) = [a_1^0(p(t)), b_0^0(p(t)), d_1^0(p(t))]^T$ and $\theta^0 = (-0.9, 0.1, 10, -1, -0.7, 0.08)^T$. The scheduling interval $[p_{min} p_{max}]$ is here given by $[0 \ 8]$.

In this example, we wish to determine an estimate $\hat{\theta}$ of θ^0 with a guaranteed standard deviation for each entry of $\hat{\theta}$. We therefore choose $R_{adm} = \text{diag} \left(\frac{1}{\sigma_1^2}, \dots, \frac{1}{\sigma_6^2} \right)$ with $\sigma_1 = 0.003$, $\sigma_2 = 3.33 \times 10^{-4}$, $\sigma_3 = 0.0333$, $\sigma_4 = 0.0033$, $\sigma_5 = 0.1167$ and $\sigma_6 = 0.0133$ (see Section 3.1).

We suppose that the local LTI identification experiments can only be performed at the $M_{grid} = 17$

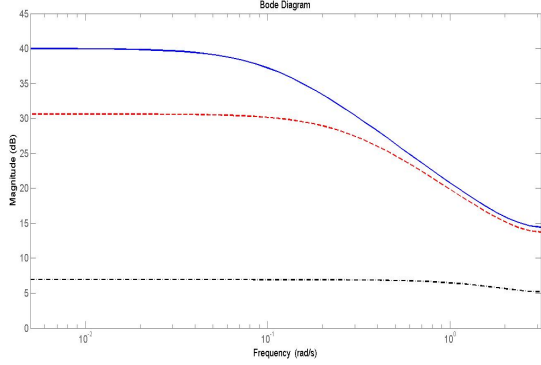


Fig. 2. Modulus of the frequency response of $G(z, \zeta_0(\mathbf{p}_m))$ for $\mathbf{p}_m = 0$ (blue solid), for $\mathbf{p}_m = 1.5$ (red dashed) and for $\mathbf{p}_m = 8$ (black dashdot)

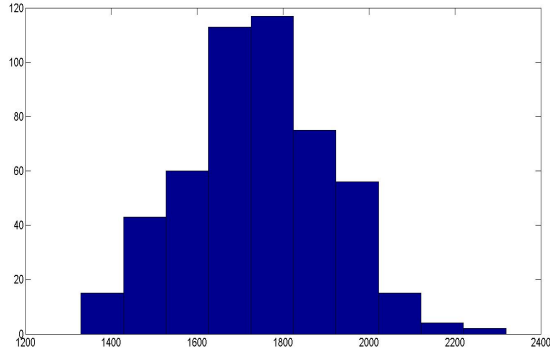


Fig. 3. Histogram of the optimal cost \mathcal{J}_{opt} of the second experiment for the 500 Monte-Carlo simulations

operating points obtained by gridding the scheduling interval $[0 \ 8]$ with a fixed step of 0.5 i.e. $\mathbf{p}_{grid} = \{0, 0.5, \dots, 7.5, 8\}$. Finally, the input power spectra Φ_{u_m} of the local LTI experiments are all parametrized as in (20) with $L = 10$ and the duration N_p of these local LTI identification experiments is initially posed to $N_p = 1000$.

In order to solve the optimal experiment design problem of Section 3.1, we first determine the spectra Φ_{u_m} of an LPV experiment with M_{grid} local LTI identification experiments by solving the optimization problem (19) with $M = M_{grid}$ and $\mathbf{P}^M = \mathbf{p}_{grid}$. We indeed know (see Section 3.3) that this will deliver the optimal cost \mathcal{J}_{opt} . We first perform this optimization by computing the covariance matrices (7) using the exact values of θ^0 and of σ_e^2 and we obtain $\mathcal{J}_{opt} = \mathcal{J}_{opt}(M_{grid}, \mathbf{p}_{grid}) = 3511$. Moreover, we observe that the obtained spectra Φ_{u_m} ($m = 1, \dots, 17$) are nonzero at only three values (i.e. $\mathbf{p}_m = 0, 1.5$ and 8) of the 17 operating points in \mathbf{p}_{grid} . The nonzero spectra at $\mathbf{p}_m = 0, 1.5$ and 8 are given in Figure 1. These spectra correspond to an input energy $N_p \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega = 1019$ at $\mathbf{p}_m = 0$, an input energy of 1349 at $\mathbf{p}_m = 1.5$ and an input energy of 1143 at

\mathbf{P}^M	$\mathcal{J}_{opt}(M, \mathbf{P}^M)$
$\bar{\mathbf{P}}^M = \{0, 1.5, 8\}$	3511
$\mathbf{P}^M = \{0, 5.5\}$	7904
$\mathbf{P}^M = \{0, 8\}$	18388
$\mathbf{P}^M = \{1.5, 8\}$	32764
$\mathbf{P}^M = \{0, 1.5\}$	57451

Table 1
 $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ for different choices of \mathbf{P}^M

$\mathbf{p}_m = 8$. The Bode diagrams of the corresponding LTI transfer functions $G(z, \zeta_0(\mathbf{p}_m))$ are given in Figure 2. We observe that, since the dynamics of the LTI systems are quite different for these three values of \mathbf{p}_m , the optimal spectra Φ_{u_m} are also quite different.

Let us thus consider the LPV experiment with $\bar{M} = 3$ local LTI experiments at $\bar{\mathbf{P}}^M = \{0, 1.5, 8\}$ and with the spectra $\bar{\Phi}(\bar{\mathbf{P}}^M)$ given in Figure 1. Since $N_p = 1000$ satisfies (18) for $\bar{\mathbf{P}}^M = \{0, 1.5, 8\}$, this LPV experiment is the solution of the optimal experiment design problem defined in Section 3.1. This LPV experiment is thus the experiment with the least input energy that delivers an estimate $\hat{\theta}$ whose covariance matrix P_θ satisfies the constraint $P_\theta^{-1} > R_{adm}$.

It is interesting to note that the optimal LPV experiment is characterized by three local LTI identification experiments. Consequently, in this example, the optimal LPV experiment is characterized by more local LTI identification experiments than strictly required by the condition of Theorem 1. Indeed, according to Theorem 1, $n_p + 1 = 2$ experiments at different \mathbf{p}_m are sufficient to obtain a strictly positive-definite P_θ and thus to satisfy the accuracy constraints for some N_p and Φ_{u_m} . To further illustrate this, we have computed the costs $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ that are obtained by solving (19) for all possible subsets \mathbf{P}^M of \mathbf{p}_{grid} containing two operating points (i.e. $M = 2$). The minimal value of this cost is equal to 7904 and is obtained for $\mathbf{P}^M = \{0, 5.5\}$. We observe that this cost is, as expected, (much) higher than $\mathcal{J}_{opt} = 3511$. In Table 1, we also compare the optimal cost $\mathcal{J}_{opt} = 3511$ to the costs $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ obtained for some other sets \mathbf{P}^M with $M = 2$. We observe that the costs $\mathcal{J}_{opt}(M, \mathbf{P}^M)$ can become very large in some cases.

Until now in this example, we have supposed that we know θ^0 and σ_e^2 in order to compute the expression of the covariance matrices. In practice, θ^0 and σ_e^2 are of course unknown and we have instead to follow the procedure of Section 4. This procedure starts with an arbitrary initial LPV experiment. Here, we assume that this arbitrary LPV experiment is characterized by $M_1 = 4$ local LTI experiments with $N_p = 1000$ and $\mathbf{P}^{M_1} = \{0, 3, 5.5, 8\}$ (see Remark 7). The input spectra Φ_{u_m} used at these four operating points are all equal to $\Phi_{u_m}(\omega) = 0.75 \forall \omega$. The cost \mathcal{J} of this initial LPV experiment is thus equal to 3000 and this initial LPV experiment delivers an initial estimate $\hat{\theta}_1$ of θ^0 , an initial estimate $\hat{\sigma}_{e,1}^2$ of

σ_e^2 and an initial covariance matrix P_{θ_1} that does not satisfy the desired accuracy constraint. Consequently, it makes sense to optimally design the additional LPV experiment that is necessary to obtain an estimate $\hat{\theta}$ satisfying the accuracy constraint. As proposed in Section 4, we follow almost the same procedure as in the case where θ^0 and σ_e^2 known, but we now compute the covariance matrices using the approximations $\theta^0 \approx \hat{\theta}_1$, $\hat{\sigma}_{e,1}^2 \approx \sigma_e^2$ and we replace R_{adm} by $R_{adm} - P_{\theta_1}^{-1}$. The optimal second experiment will therefore depend on the obtained initial estimates $\hat{\theta}_1$, $\hat{\sigma}_{e,1}^2$ and P_{θ_1} . In order to analyze this dependency, we have performed 500 Monte-Carlo simulations. For each Monte-Carlo simulation, we have performed the four initial local LTI experiments at $\mathbf{P}^{\mathbf{M}_1} = \{0, 3, 5.5, 8\}$ with different realizations of the white-noise input signals and different realizations of the noise signal e , we have deduced the estimates $\hat{\theta}_1$, $\hat{\sigma}_{e,1}^2$ and P_{θ_1} and performed the optimal design of the second LPV experiment as presented in Section 4. We have observed that, for these Monte-Carlo simulations, the obtained set $\bar{\mathbf{P}}^{\mathbf{M}}$ can vary, e.g. $\bar{\mathbf{P}}^{\mathbf{M}} = \{0, 1.5, 2, 8\}$ or $\bar{\mathbf{P}}^{\mathbf{M}} = \{0, 1.5, 8\}$. The optimal cost \mathcal{J}_{opt} of this second LPV experiment will also vary for different Monte-Carlo simulations as shown in the histogram presented in Figure 3. The obtained values for this cost have a mean of 1741 while the maximal value is 2319. We observe that this cost is smaller than previously. The desired accuracy R_{adm} is indeed here obtained by combining this additional LPV experiment with the initial one. In other words, this second LPV experiment is the one with the least input energy that guarantees that the estimate $\hat{\theta}$, which is determined by combining both this second LPV experiment and the initial one, has a covariance matrix satisfying $P_{\hat{\theta}}^{-1} > R_{adm}$. Let us also note that, for the worst of the 500 Monte-Carlo simulations, the total input energy used to obtain a sufficiently accurate estimate is equal to $3000 + 2319 = 5319$. This energy is of course larger than the cost $\mathcal{J}_{opt} = 3511$ that we have obtained in the ideal case where θ^0 and σ_e^2 are known. However, it is much smaller than the cost that is necessary to obtain $P_{\hat{\theta}}^{-1} > R_{adm}$ using the experimental conditions of the initial experiment. Indeed, in order to satisfy $P_{\hat{\theta}}^{-1} > R_{adm}$ using local LTI experiments at $\mathbf{P}^{\mathbf{M}_1} = \{0, 3, 5.5, 8\}$ with a white noise having the same power for all four experiments, we have to choose $\Phi_{u_m}(\omega) = 2.47 \forall \omega$ ($m = 1, \dots, 4$) when $N_p = 1000$. This corresponds to a cost $\mathcal{J} = 9880$.

Finally, we have verified that, despite the use of the initial estimate, the approach presented in Section 4 yields, in the vast majority of the Monte-Carlo simulations, an acceptable estimate $\hat{\theta}$ of θ^0 . Recall for this purpose that $P_{\hat{\theta}}^{-1} > R_{adm}$ implies that $(\hat{\theta} - \theta^0)^T R_{adm} (\hat{\theta} - \theta^0) < (\hat{\theta} - \theta^0)^T P_{\hat{\theta}}^{-1} (\hat{\theta} - \theta^0)$ and recall also that $(\hat{\theta} - \theta^0)^T P_{\hat{\theta}}^{-1} (\hat{\theta} - \theta^0) < 12.6$ with probability 95%. Consequently, we can deem that $\hat{\theta}$ is an acceptable estimate of θ^0 if $(\hat{\theta} - \theta^0)^T R_{adm} (\hat{\theta} - \theta^0) < 12.6$ and we have observed

that this is indeed the case in 499 of the 500 performed Monte Carlo simulations.

Let us now return to the case where θ^0 and σ_e^2 are known and recall that the solution is then characterized by an input energy of 1019 at $\mathbf{p}_m = 0$, an input energy of 1349 at $\mathbf{p}_m = 1.5$ and an input energy of 1143 at $\mathbf{p}_m = 8$. Let us now suppose that, for security reasons, the input energy cannot be larger than 1000 at any operating point. In this case, we have to consider the extra constraints described in Remark 6 with $\alpha_m = 1000$ ($m = 1, \dots, 17$) and to solve the optimization problem (19) with $M = M_{grid}$ and $\mathbf{P}^{\mathbf{M}} = \mathbf{P}^{grid}$ with these extra constraints. The obtained spectra Φ_{u_m} ($m = 1, \dots, 17$) are then nonzero at five values (i.e. $\mathbf{p}_m = 0, 1.5, 2, 7.5$ and 8) and the corresponding input energies at those five operating points are respectively 1000, 1000, 334, 209 and 1000. As expected, the total input energy $\mathcal{J}_{opt} = 3543$ is larger than the original one (i.e. 3511). It is also interesting to note that the maximum allowed energy is used at the operating points that had been selected in the original setting (i.e. the one without energy constraints).

Let us now also consider another situation i.e. the situation where the operating points $\mathbf{p}_m \geq 5$ are much more difficult to reach. In this case, the original solution with an experiment at $\mathbf{p}_m = 8$ can be problematic. Let us therefore solve the optimization problem by removing from \mathbf{P}^{grid} the operating points $\mathbf{p}_m \geq 5$. With this alternative setting, we obtain $\bar{\mathbf{P}}^{\mathbf{M}} = \{0, 1, 4.5\}$. We thus avoid the operating point $\mathbf{p}_m = 8$ that is more difficult to reach, but the price to pay is an important increase in the cost i.e. $\mathcal{J}_{opt} = 7357$. Consequently, the operating point $\mathbf{p}_m = 8$ seems quite important in this example to obtain the desired accuracy with low input energy.

6 Concluding remarks

This paper presents the first results of optimal experiment design for the identification of LPV systems. In this paper, we have restricted attention to the local approach for LPV system identification where a single scheduling variable is maintained constant at different operating points. Future work will consider the case of multiple scheduling variables. Indeed, if the approach of Section 3.3 based on a gridding of the scheduling space can in theory also be applied for that case, the complexity will increase and it is therefore interesting to investigate whether other approaches could be developed for the selection of the optimal operating points. We also wish to investigate optimal experiment design for the global approach for LPV system identification. Indeed, the particular choice for the variations of the scheduling variable(s) in the local approach (i.e. successive constant values) is not guaranteed to be the optimal one in order to obtain the most information on the time-invariant parameters θ^0 describing the LPV system. We will also continue our work towards accuracy constraints inspired by LPV control objectives.

References

- [1] P. Apkarian and P. Gahinet. A convex characterization of gain-scheduled H_∞ controllers. *IEEE Transactions on Automatic Control*, 40(5):853-864, 1995.
- [2] P. Apkarian, P. Gahinet, and J.M. Biannic. Gain-scheduled H_∞ control of missile via linear matrix inequalities. *Journal of Guidance, Control and Dynamics*, Vol 28(3), 1994.
- [3] A.A. Bachnas, R. Tòth, J.H.A. Ludlage, and A. Mesbah. A review on data-driven linear parameter-varying modeling approaches: A high-purity distillation column case study. *Journal of Process Control*, Vol 24(4):272-285, 2014.
- [4] B. Bamieh and L. Giarré. Identification of linear parameter varying models. *Int. Journal of Robust and Nonlinear Control*, Vol. 12, pp 841-853, 2002.
- [5] X. Bombois and G. Scorletti. Design of least costly identification experiments : The main philosophy accompanied by illustrative examples. *Journal Européen des Systèmes Automatisés*, Vol. 46(6-7), pp.587-610, 2012.
- [6] X. Bombois, G. Scorletti, M. Gevers, P. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, Vol. 42(10), pp. 1651-1662, 2006.
- [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] A. Dankers, R. Toth, P. Heuberger, X. Bombois, and P. Van den Hof. Informative data and identifiability for LPV-ARX prediction error identification. In *Proc. IEEE Conference on Decision and Control*, pages 799–804, Orlando, 2011.
- [9] M. Gilson, V. Laurain, H. Garnier, S. Payraudeau, and C. Grégoire. A new data-based modelling method for identifying parsimonious nonlinear rainfall/flow models. *Journal Européen des Systèmes Automatisés*, Vol. 6(7), pp.878-888, 2011.
- [10] H. Hjalmarsson. System identification of complex and structured systems. *European Journal of Control*, Vol. 15(3-4), pp. 275-310, 2009.
- [11] H. Jansson and H. Hjalmarsson. Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, Vol 50(10), pp. 1534-1549, 2005.
- [12] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice-Hall, Upper Saddle River, NJ, 1993.
- [13] A. Khalate, X. Bombois, R. Tòth, and R. Babuska. Optimal experimental design for LPV identification using a local approach. In *15th Symposium on System Identification*, pp. 162-167, St-Malo, 2009.
- [14] V. Laurain, M. Gilson, R. Toth, and H. Garnier. Refined instrumental variable methods for identification of LPV Box-Jenkins models. *Automatica*, Vol. 46(6), pp. 959-967, 2010.
- [15] Q. Liu, J. Gross, S. Pfeiffer, and H. Werner. A local approach for the LPV identification of an actuated beam using piezoelectric actuators and sensors. *Mechatronics*, Vol. 24, pp. 289-297, 2014.
- [16] L. Ljung. *System Identification - Theory for the User, 2nd Edition*. Prentice-Hall, Upper Saddle River, NJ, 1999.
- [17] M.G. Potters. *Experiment design for identification of structured linear systems*. PhD thesis, TU Delft, 2016.
- [18] M.G. Potters, X. Bombois, and P.M.J. Van den Hof. Experiment time minimisation under parameter accuracy constraints and time-domain signal amplitude bounds. In *15th European Control Conference*, Aalborg, 2016.
- [19] G. Scorletti and L. El Ghaoui. Improved LMI conditions for gain scheduling and related problems. *International Journal of Robust and Nonlinear Control*, Vol. 8(10):845–877, August 1998.
- [20] M. Steinbuch, M.J.G. van de Molengraft, P.S.C. Heuberger, and A.J. Van der Voort. Experimental modelling and LPV control of a motion system. In *Proceedings of the American Control Conference*, pp. 1374- 1379, Denver, 2003.
- [21] R. Tòth. *Modeling and identification of Linear Parameter Varying Systems*. Springer, Berlin, 2010.
- [22] J.A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Transactions on Information Theory*, Vol 52, pp. 1030-1051, 2006.
- [23] D. Vizer and G. Mercère. An H_∞ -norm-based approach for operating point selection and LPV model identification from local experiments. *Periodica Polytechnica*, Vol. 58(3), pp. 121-131, 2014.
- [24] M. G. Wassink, M. van de Wal, C. Scherer, and O. Bosgra. LPV control for a wafer stage: beyond the theoretical solution. *Control Engineering Practice*, Vol. 13, pp. 231-245, 2005.