Optimal experimental design for LPV identification using a local approach *

Amol A. Khalate ∗ Xavier Bombois ∗ Roland Tóth ∗ Robert Babuška ∗

∗ Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.
(email: {a.a.khalate, x.j.a.bombois, r.toth, r.babuska}@dcs.c.tudelft.nl)

Abstract: A common approach for dealing with non-linear systems is to describe the system by a model with parameters that vary as a function of the operating point. Consequently, the non-linear system is seen as a combination of local Linear Time-Invariant (LTI) systems, one for each value of the operating point. Such representations are called Linear Parameter-Varying (LPV) models. Due to the importance of this representation for the control of nonlinear systems, numerous algorithms have recently been developed to identify LPV models. However, the optimal design of such identification experiments remains completely unexplored. In this paper, we consider the so-called local approach for the LPV identification. In the local approach, the local linear models, corresponding to a series of fixed operating points, are identified by performing one identification experiment at each of these operating points. The LPV nature of the system is then retrieved by interpolating the value of the parameters at other operating points with, for example, a polynomial function which is fitted through the parameters identified at the operating points considered. We present an approach to choose optimally the value of the operating points at which the local identification experiments will be performed. By optimal, we mean that the value of the operating points are optimized in such a way that the LPV model obtained after interpolation has a maximum accuracy.

Keywords: Optimal experiment design, Linear Parameter-Varying systems

1. INTRODUCTION

A common approach for dealing with non-linear systems is to describe the system by a model with parameters that vary as a function of the operating point, or as a function of other exogenous variables (i.e. the scheduling variables). Consequently, the non-linear system is seen as a collection of local linear systems, one for each value of the operating point (or one for each value of the exogenous variables).

Such representations are called Linear Parameter-Varying (LPV) models. A practical use of LPV representations is stimulated by the fact that LPV control design is well worked out, extending the results of optimal and robust LTI control to nonlinear plant. These control design techniques deliver an “LPV controller”, i.e., a controller whose parameters are also a function of the value of the operating point, or a function of the value of the exogenous variable (see e.g. Becker and Packard [1994], Scherer [2001], Scorletti and Ghaoui [1998], Dinh et al. [2005]).

Due to the importance of the LPV representation, numerous algorithms have recently been developed to identify LPV models. However, the subject to be investigated in this paper - the optimal design of such identification experiments - remains completely unexplored. Before presenting the two mainstream approaches to identify an LPV model:

- Local approach (Steinbuch et al. [2003], Wassink et al. [2004], Tóth et. al. [2007]). In the local approach, the local linear models corresponding to a series of fixed operating points (or corresponding to a series of fixed values of the exogenous variable) are identified by performing one identification experiment at each of these operating points. The LPV nature of the system is then retrieved by interpolating the value of the parameters at other operating points with, for example, a polynomial function which is fitted through the parameters identified at the operating points considered.

- Global approach (Lee and Poolla [1999], Bamieh and L. Giarré [2002], Felici et al. [2006], Tóth et. al. [2007], Wingerden et al. [2009]). The global approach consists of exciting all the non-linearities of the system via one single experiment passing through a large number of operating points (or a large number of values of the exogenous variables) and of directly identifying the functional dependence of the parameters on the value of the operating points (or on the value of the exogenous variables) based on the collected input-output data.

Both approaches have significant advantages and disadvantages with the common need for the investigation of

* This work has been carried out as part of the Octopus project with Océ Technologies B.V. under the responsibility of the Embedded Systems Institute. This project is partially supported by the Netherlands Ministry of Economic Affairs under the Bzik program.
optimal experiment design should be investigated for both approaches. In this paper, as a first step, we consider optimal experiment design for the local approach and our objective is to determine the experimental conditions of this method to maximize the accuracy of the identified LPV model.

Optimal experiment design has been extensively investigated for LTI systems (Ljung [1999], Jansson [2004], X. Bombois et. al [2006]). For LTI systems, generally, the objective is to determine the excitation signal \( u(t) \) maximizing the accuracy of the identified model under the constraint that the power of \( u(t) \) remains below some threshold. The current paper is an extension of this framework to the case of LPV systems. In the LPV case, the design variable is not only the excitation signal \( u(t) \) used for each local LTI identification experiment, but also the value of the operating points at which the identification experiments are performed. Hence optimal identification experiment design for LPV systems is a complex problem.

In this paper, as a first step, we will focus on the optimal determination of the operating points at which the linear identifications are performed. These values also determine the interpolation points used to determine the values of the parameter vector at other operating points via interpolation. This will be done by assuming that the excitation signal \( u(t) \) is sufficiently powerful or sufficiently long to neglect any variance effect in the linear identifications and thus by assuming that the linear identification experiments allow one to get fully accurate models at the operating points where these linear identifications are performed.

Optimal experimental design always requires the knowledge of the true system or at least some a-priori knowledge of this true system and this paper is not different: the determination of the optimal location of the operating points will require the knowledge of the true LPV system. Even though this requirement seems unrealistic, the true LPV system can be replaced in practice by an initial estimate deduced from an (un-optimized) identification or from a first principle model.

The paper is organised as follows: In Section 2, the local approach for LPV identification is presented. In Section 3, we propose our method to determine the operating points optimally to maximize the accuracy of the identified LPV model. In Section 4, a numerical example is presented to demonstrate the efficacy of the proposed method. Finally, conclusions and future directions are presented in Section 5.

2. LOCAL APPROACH FOR LPV SYSTEM IDENTIFICATION

Based on the LPV modeling concept, the dynamics of the real-system vary as a function of the operating point \( p \) at which the system is operated. In this paper, we consider that the operating points are represented by a scalar variable, i.e. \( p \in \mathbb{R} \), that can vary in the interval \( [p_{\text{min}}, p_{\text{max}}] \). In particular, we consider a “true” single-input single-output LPV system that can be represented at each fixed value of the operating point \( p \in \mathbb{R} \) by the following \( p \)-dependent difference equation:

\[
S: y(t) = \phi^T(t) \theta_0(p) + v(t) \tag{1}
\]

where \( \phi(t) = [-y(t-1), -y(t-2), ..., u(t-1), u(t-2), \ldots]^T \in \mathbb{R}^k \) is a regression vector and \( v(t) \) is a stochastic noise. Both \( \phi(t) \) and \( v(t) \) are independent of the scheduling variable \( p \). The system \( S \) is only dependent on \( p \) via the value of the “true” parameter vector \( \theta_0(p) \in \mathbb{R}^k \). In other words, we assume that the local linear systems corresponding to each frozen value of the operating point \( p(t) \equiv \bar{p} \), where \( \bar{p} \in \mathbb{P} \), have the same order. In the sequel, we will suppose that this order is known and thus that, in order to identify a model of \( (1) \), it is sufficient \(^1\) to find an estimate of the \( p \)-dependent true parameter vector \( \tilde{\theta}_0(p) \). Note that we do not make any assumption on the dependence of \( \tilde{\theta}_0(p) \) on \( p \); this can be any smooth function, e.g. a nonrational.

In this paper, we will present a method in order to design optimally the identification of \( (1) \) in the case of the so-called local approach.

In the local approach, the local linear models corresponding to a series of fixed operating points are identified by performing one identification experiment at each of these operating points. The LPV nature of the system is then retrieved by interpolating the value of the parameters at other operating points with a function which is fitted through the parameters identified at the operating points considered.

In this paper, we will suppose that the number \( n \) of local linear identifications is fixed. The variables that have to be designed in the local approach are therefore:

- The value of the operating points at which we will identify the local linear models
- The experimental conditions for the linear identification at each of these operating points.

In order to be able to design optimally the identification experiment, it is very important to understand how the choice of the design variables influences the quality of the identified LPV model. For this purpose, we in the first instance suppose that we have determined the set \( \mathcal{P} = \{ \bar{p}_1, \ldots, \bar{p}_n \} \) of \( n \) operating points at which a linear identification will be performed and that we have chosen the input signal \( u_i(t) \) (\( i = 1, \ldots, n \)) as well as the length \( N_i \) (\( i = 1, \ldots, n \)) of these linear identification. Based on this choice, we will show how the LPV model is identified and then analyze the accuracy of the identified LPV model and the influence of the experimental variables on this accuracy.

The LPV model can be identified as follows. We first fix \( p \) to the operating point \( \bar{p}_1 \), i.e. \( p = \bar{p}_1 \) and we collect data by applying the input signal \( u_1(t) \) for a duration \( N_1 \) to the system \( (1) \):

\[
y(t) = \phi(t) \tilde{\theta}_0(\bar{p}_1) + v(t)
\]

Note that the data-generating system is now linear time-invariant. Based on the collected data and a full order model structure, we can identify a consistent estimate \( \hat{\theta}_0(\bar{p}_1) \) of \( \theta_0(\bar{p}_1) \) using e.g. prediction error identification or instrumental variable methods (Ljung [1999]). This procedure is then repeated for the other operating points i.e. \( \bar{p}_2, \bar{p}_3, \ldots, \bar{p}_n \). After these \( n \) linear identification experiments, we have thus estimated the function \( \theta_0(p) \) at each fixed value of the operating point \( \bar{p} \), where \( \bar{p} \in \mathbb{P} \), have the same order. In the sequel, we will suppose that this order is known and thus that, in order to identify a model of \( (1) \), it is sufficient \(^1\) to find an estimate of the \( p \)-dependent true parameter vector \( \tilde{\theta}_0(p) \). Note that we do not make any assumption on the dependence of \( \tilde{\theta}_0(p) \) on \( p \); this can be any smooth function, e.g. a nonrational.

\(^1\) In this paper we are not interested in the modeling of the noise disturbance \( v(t) \).
different operating points \( \bar{p}_i \); these \( n \) estimates are denoted by \( \theta(\bar{p}_i) \).

The accuracy of \( \hat{\theta}(\bar{p}_i) \) with respect to \( \theta_0(\bar{p}_i) \) is, of course, related to the duration \( N_i \) of the respective experiment and to the power of the chosen input signals \( s_i(t) \). In this paper, we assume that we can neglect the variance error induced by these \( n \) linear identification experiments and consequently:

\[
\hat{\theta}(\bar{p}_i) = \theta_0(\bar{p}_i), \quad \forall \bar{p}_i \in \mathcal{P},
\]

(2)

This is of course an important simplification that will need to be relaxed in subsequent contributions. Note that this simplification reduces the optimal identification experiment design to the optimal determination of the \( n \) operating points \( \mathcal{P} = \{\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_n\} \) at which the identification will be performed. Thus we now analyze how the choice of operating points influences the accuracy of the LPV model.

After the \( n \) local identification experiments, the next step in determining the LPV model is to interpolate the value of the parameter vector \( \theta(p) \) at other operating points \( \bar{p} \in \mathbb{P} \) than the ones in \( \mathcal{P} \). We introduce the notation \( \tilde{\theta}(\bar{p}) \) in order to distinguish the interpolated model \( \theta(p) \) from the \( n \) identified parameter vectors \( \hat{\theta}(\bar{p}_i) \) \( (i = 1, \ldots, n) \). The interpolation is generally done by fitting a function through the parameters identified at the operating points considered.

Recall that \( \theta_0(p) \) can be any smooth function. However, the choice of the interpolating function must be restricted if we want to use the model for LPV control design. Indeed, most of the methods to design an LPV controller, i.e. a controller whose parameters are also a function of the value of the set-point \( p \), requires the function \( \theta(p) \) to be a polynomial\(^2\) function of \( p \) (see e.g. Becker and Packard [1994], Scherer [2001], Scorletti and Ghaoui [1998], Dinh et al. [2005]). Consequently, we parameterize \( \theta(p) \) as follows:\(^3\):

\[
\theta(p) = \lambda_0 + \lambda_1 p + \ldots + \lambda_m p^m
\]

(3)

for a given \( m \) and for some vectors \( \lambda_j \in \mathbb{R}^k \) \( (j = 0, 1, \ldots, m) \). The vectors \( \lambda_j \) will be determined to ensure that for all \( \bar{p}_i \in \mathcal{P} \):

\[
\hat{\theta}(\bar{p}_i) = \lambda_0 + \lambda_1 \bar{p}_i + \ldots + \lambda_m \bar{p}_i^m
\]

(4)

with \( \hat{\theta}(\bar{p}_i) \) \( (i = 1, \ldots, n) \) the identified parameter vector at each of the operating points in \( \mathcal{P} \). In order to satisfy the constraint described by (4), the degree \( m \) of the polynomial function should be chosen greater or equal to \( n - 1 \). In order to avoid over parameterization and because the degree of the polynomial is directly related to the complexity of the LPV controller, we choose

\[
m \geq n - 1.
\]

(5)

\(^2\) Note that the polynomial dependence is not the only possibility: we can also consider a rational function of \( p \) or a linear combination of known basis functions in \( p \). Such parametrizations can also be used here without any difficulty.

\(^3\) The LPV control design methods generally require a model expressed in the state-space form. A state-space representation can be directly determined from an input-output model (see Tóth [2008]) with the parameter vector parameterized as in (3).

The vectors \( \lambda_j \) \( (j = 0, \ldots, n - 1) \) satisfying (4) can then be determined as the solution of a series of linear systems of \( n \) equations with \( n \) unknowns (one for each entry of the vector \( \theta \in \mathbb{R}^k \)). Let us denote by \( \bar{\lambda}_0, \ldots, \bar{\lambda}_{n-1} \) the vectors obtained in that manner.

The identified LPV model is then:

\[
\mathcal{M}: \quad y(t) = \phi(t)^T \tilde{\theta}(p),
\]

(6a)

\[
\tilde{\theta}(p) = \lambda_0 + \bar{\lambda}_1 p + \ldots + \bar{\lambda}_n p^n
\]

(6b)

where the subscript \( \mathcal{P} \) is introduced to stress that the vector \( \tilde{\theta}(p) \) obtained by ensuring (4) is different for different sets \( \mathcal{P} = \{\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_n\} \).

Remark. The transients generated by changing the operating point are generally important to capture the nonlinear behaviors of the system into the LPV model. However, here, we have neglected these transients effects for simplicity. Still, the considered LPV model can represent a wide range of systems. For example, in many chemical plants, the operating point does not vary as a smooth function of time. Indeed, based on the production requirements, the operating point is changed off-line. In this kind of situations, ignoring the transient effects generated by the change of operating points will hardly affect the quality of the LPV model.

3. EXPERIMENT DESIGN FOR LPV SYSTEM IDENTIFICATION

By construction, the modeling error \( \theta_0(p) - \tilde{\theta}(p) \) is equal to zero at the \( n \) operating points where the linear identification has been performed (see (2) and (4)). For other values of the operating points, the modeling error will be non-zero if \( \theta_0(p) \) is of an higher complexity than the parametrization (3).

It is important to note that the accuracy obtained at each \( p \) depends on the choice of the \( n \) operating points \( \mathcal{P} = \{\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_n\} \) at which the linear identifications are performed. For different values of the operating points \( \bar{p}_i \) \( (i = 1, \ldots, n) \), we obtain a different model (6a-b).

In the numerical example of Section 4, we show that the optimization of the operating points can lead to a significant improvement of the accuracy.

To measure the accuracy of the identified LPV model (6a-b) with respect to the true LPV system (1), we introduce the following accuracy measure:

\[
J_{\mathcal{P}} = \int_{\mathcal{P}_{\min}}^{\mathcal{P}_{\max}} \| \tilde{\theta}(p) - \theta_0(p) \|^2 dp
\]

(7)

where \( \| \cdot \| \) denotes the Euclidean norm of a vector and \( \mathcal{P} = [\mathcal{P}_{\min}, \mathcal{P}_{\max}] \) represents the range of variation of \( p \).

For this accuracy measure, the optimal location \( \mathcal{P}_{\text{opt}} \) for the operating points \( \mathcal{P} = \{\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_n\} \) is the solution of the following optimization problem:

\[
\arg \min_{\mathcal{P}} \int_{\mathcal{P}_{\min}}^{\mathcal{P}_{\max}} \| \tilde{\theta}(p) - \theta_0(p) \|^2 dp
\]

(8)

subject to the constraint

\[
\mathcal{P}_{\min} \leq \bar{p}_i \leq \mathcal{P}_{\max}, \quad \forall \bar{p}_i \in \mathcal{P}.
\]

(9)

In order to solve (8)-(9), we see that we require the true function \( \theta_0(p) \). This is an usual assumption in any optimal
experiment design problem (see e.g. Ljung [1999], Jansson [2004], X. Bombois et al [2006]). The true \( \theta_0(p) \) can always be replaced in (8)-(9) by an initial estimate of \( \theta_0(p) \) that can e.g. be deduced from a first-principle model.

The optimization problem (8)-(9) is a constrained nonlinear optimization problem that can be solved e.g. using the \texttt{Matlab}-command \texttt{fmincon} of the Optimization Toolbox. Since it is a nonlinear optimization problem, the algorithm to solve (8)-(9) will require an initial guess for \( \mathcal{P}_{\text{opt}} \) and, depending on this initial guess, may converge to a local minimum.

**Remark.** A higher accuracy can also be obtained by increasing the number \( n \) of operating points at which a linear identification is performed. Indeed, increasing the number of identified operating points allows one to increase the degree \( m \) of the interpolating polynomial function \( \theta(p) \) (see (5)). However, it is important to note that the cost of the identification is directly related to the number of local linear models being identified: higher the number of models, the longer and more intrusive the total identification experiment is. Moreover, the complexity of the LPV controller is also increases for increasing values of \( m \).

4. NUMERICAL EXAMPLE

We now demonstrate the proposed methodology on an example. Consider the following true LPV system with two parameters:

\[
y(t) = [\phi^T(t) \ u(t-1)] \begin{bmatrix} a_0(p) \\ b_0(p) \end{bmatrix},
\]

where the dependence on \( p \) of the two parameters is nonlinear:

\[
a_0(p) = -1 + \left( 0.6 e^{-\sin(0.06 p)} \right) \\
b_0(p) = 1 + \frac{8000 \cos(0.06 p)}{p^2}
\]

These two functions of \( p \) are represented by blue solid lines in Figures 1 and 2, respectively. In this example, we suppose that \( p_{\text{min}} = 60 \) and that \( p_{\text{max}} = 150 \). Note that, since we take assumption (2), we can omit the noise contribution in (10)

The goal is to identify a model of (10) using the local approach with \( n = 4 \) local linear identifications. According to (5), we therefore model the dependence on \( p \) of the parameter vector using a polynomial function of degree \( m = 3 \). Suppose that we have initially chosen the four operating points as follows:

\[
\mathcal{P} = \{ 60, 100, 120, 140 \}
\]

Using these operating points for the local approach, the identified model (6a-b) is parameterized by a \( p \)-dependent parameter vector \( \hat{\theta}_\mathcal{P}(p) = [\hat{a}_\mathcal{P}(p) \ \hat{b}_\mathcal{P}(p)]^T \) whose two entries are represented by red dashed lines in Figures 1 and 2, respectively. We observe that, as expected by our assumption (2), \( \hat{\theta}_\mathcal{P}(p_i) = \theta_0(p_i) \) for all four operating points in \( \mathcal{P} \) (see the red circles in Figures 1 and 2).

Fig. 1. Top: \( a_0(p) \) (blue solid), \( \hat{a}_\mathcal{P}(p) \) using the initial operating points (11) (red dashed), \( \tilde{a}_\mathcal{P}_{\text{opt}}(p) \) using the optimal operating points (12) (black dotted). Bottom: \( (a_0(p) - \hat{a}_\mathcal{P}(p))^2 \) (red dashed), \( (a_0(p) - \tilde{a}_\mathcal{P}_{\text{opt}}(p))^2 \) (black dotted)

![Fig. 1](image1.png)

Fig. 2. Top: \( b_0(p) \) (blue solid), \( \hat{b}_\mathcal{P}(p) \) using the initial operating points (11) (red dashed), \( \tilde{b}_\mathcal{P}_{\text{opt}}(p) \) using the optimal operating points (12) (black dotted). Bottom: \( (b_0(p) - \hat{b}_\mathcal{P}(p))^2 \) (red dashed), \( (b_0(p) - \tilde{b}_\mathcal{P}_{\text{opt}}(p))^2 \) (black dotted)

![Fig. 2](image2.png)

However, the accuracy of \( \hat{\theta}_\mathcal{P}(p) \) with respect to \( \theta_0(p) \) is relatively low especially for the interval \([60, 100]\).

In order to maximize the accuracy of the identified LPV model, we will optimize the operating points \( \mathcal{P} \) at which
we, we have improved the accuracy measure \( J_\mathcal{P} \) defined in (7) from 0.99385 to 0.45642.

The improvement of the accuracy of the identified LPV model can also be evidenced by comparing the time response of the true LPV system with the time response of the identified LPV models. Here, we consider the time response for the operating point \( \bar{p} = 74 \) when the input signal is \( u(t) = \sin(t) + \sin(4t) + \sin(8t) \). The time responses are represented in Figure 3 where we observe once again that the model identified with \( P_{n=5}^{\text{opt}} \) is the most accurate followed by the model identified with the four operating points in \( P_{\text{opt}} \), see (12).

5. CONCLUSION

We have presented a method in order to optimally design the experimental conditions for LPV identification when the local approach is used. In the local approach, the local linear models corresponding to a series of fixed operating points are identified by performing one linear identification experiment at each of these operating points. The LPV nature of the system is then retrieved by interpolating the value of the parameters with respect to other operating points using a polynomial function which is fitted through the parameters identified at the operating points considered.

The proposed methodology optimizes the location of the operating points at which the linear identifications are performed in order to maximize the accuracy of the identified LPV model. To achieve this, we have used a number of simplifications. As an example, we use here the assumption that each local linear identification experiment delivers a fully accurate model and thus that the accuracy (or the lack of accuracy) is entirely determined by the interpolation. We have also assumed that \( \theta_0(p) \) is known. Even though this requirement seems unrealistic, the true LPV system can be replaced in practice by an initial estimate deduced from an (un-optimized) identification or from a first principle model. Most of the time, aim of obtaining the LPV model is to use this model in LPV control synthesis. However, to simplify the problem formulation of the optimal experiment design for LPV identification, we have not considered the control objectives in the identification step. Another simplification is that we did not consider the control objectives in the optimal experiment design problem (while the aim of obtaining a LPV model is generally to use it for control design). All these issues will be analyzed in subsequent contributions.

REFERENCES


