

# A kernel based approach for LPV subspace identification

I.Proimadis\* H.J.Bijl\* J.W. van Wingerden\*

\* *Delft University of Technology, Delft, 2628 CD, The Netherlands*  
(e-mail: [jproimadis@gmail.com](mailto:jproimadis@gmail.com), {[h.j.bijl](mailto:h.j.bijl),  
[@tudelft.nl](mailto:J.W.vanWingerden)}).

**Abstract:** We present a Linear Parameter Varying (LPV) subspace identification method that takes advantage of the recent developments in the Machine Learning community. More specifically, a Radial Basis Function kernel is used to model the predictor's impulse response of an LPV model and the involved hyperparameters are estimated via a marginal likelihood maximization algorithm. This step is followed by the estimation of the predictor's impulse response coefficients, evaluated at the training points. Finally, these values are used to estimate the related coefficients of the LPV model. From this point, the algorithm follows the same steps as in the LPV-PBSID<sub>opt</sub> algorithm. Simulation results verify that this algorithm can improve the accuracy of the estimated model with respect to the state-of-the-art LPV subspace methods.

© 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: LPV, subspace identification, ARX models

## 1. INTRODUCTION

Modelling and identification of Linear Parameter Varying (LPV) systems has attracted the interest of the scientific community over the past few years [Lovera et al., 2013, Tóth, 2010]. Various identification methods have been developed based on prediction error or subspace identification techniques. Prediction error methods deliver an Input/Output (I/O) model, while the asymptotical properties of the estimated variables can also be derived. Nonetheless, modern control methods for LPV systems are based on a state-space description. Recently, a mathematical framework was developed to describe the equivalence relations between the I/O and the state-space descriptions of the LPV models [Tóth et al., 2012], showing that, in the general case, a dynamic dependency on the scheduling parameters has to be introduced in the state-space model in order to have a state-minimal realization, thus increasing the complexity of the identified model.

On the other hand, subspace methods directly deliver a state-space model. One of the drawbacks of these methods is the so-called curse of dimensionality [Verdult and Verhaegen, 2002]. Nonetheless, the state-of-the-art subspace identification algorithms offer an attractive framework for the estimation of the LPV models in state-space form, while methods for decreasing the complexity of the algorithms have also been investigated [Lovera and Mercere, 2007, van Wingerden and Verhaegen, 2009].

In the very recent years, the interest on the identification of LTI systems with the use of a Bayesian framework has been revived, mainly due to developments in the machine learning community (e.g. within the context of Gaussian processes [Rasmussen and Williams, 2005]) and the development of new kernel methods that incorporate simple prior information about the unknown system, such as BIBO stability or smoothness [Pillonetto and De Nico-

lao, 2010]. In this paper we will model the predictor's impulse response coefficients as Gaussian processes and we will develop an algorithm that delivers an LPV state-space model, partially facilitating the LPV-PBSID<sub>opt</sub> algorithm from van Wingerden and Verhaegen [2009]. Given an appropriate choice of the related parameters, the new algorithm leads to a more accurate estimation of the unknown system compared to the state-of-the-art algorithms, due to the reduction of the mean squared error through the intrinsic regularization of Gaussian process estimation methods.

This paper is organized as follows. In Section 2 we formulate the identification problem and we make some necessary assumptions. In Section 3 we introduce a framework to describe the predictor's impulse response coefficients with the use of Gaussian processes and we show how we can estimate their values based on a Bayesian framework. In Section 4 we explain how we can estimate the unknown coefficients that are needed in the LPV-PBSID<sub>opt</sub> algorithm and we also give a summary of the proposed algorithm. Simulation results are given in Section 5 in order to verify the validity of the proposed algorithm. Moreover, we perform a comparison of the new algorithm with already existing subspace identification methods for LPV systems. Finally, we present our conclusions in Section 6.

## 2. PROBLEM FORMULATION AND ASSUMPTIONS

In this paper we will consider LPV systems described by an affine parameter dependence in the state equation. In this paper we will focus on LPV systems whose output equation does not depend on the scheduling parameter. The state-space equations are given by

$$x_{k+1} = \sum_{i=1}^m \mu_k^{(i)} \left( A^{(i)} x_k + B^{(i)} u_k + K^{(i)} e_k \right), \quad (1)$$

$$y_k = C x_k + D u_k + e_k, \quad (2)$$

where  $x_k \in \mathbb{R}^n$ ,  $u_k \in \mathbb{R}^{n_u}$ ,  $y_k \in \mathbb{R}^{n_y}$  and  $e_k \in \mathbb{R}^{n_y}$  denote the state, the input, the output and the zero mean white innovation process, respectively. The innovation process will be assumed to be normally distributed, with covariance described by  $\text{cov}(e_k, e_{k'}) = \sigma^2 \delta_{k-k'}$ , where the function  $\delta$  represents the Kronecker delta. Using an affine representation, we always assume that  $\mu_k^{(1)} = 1$ . Without loss of generality (as long as specific conditions concerning the delay of the system are fulfilled, see van der Veen et al. [2013], Ljung [1999]), we will assume that  $D = 0$ . Taking this into account we can rewrite (1)-(2) in the predictor form

$$x_{k+1} = \sum_{i=1}^m \mu_k^{(i)} \left( \tilde{A}^{(i)} x_k + B^{(i)} u_k + K^{(i)} y_k \right), \quad (3)$$

$$y_k = C x_k + e_k, \quad (4)$$

where  $\tilde{A}^{(i)} = A^{(i)} - K^{(i)} C$ . Now we are in the position to formulate the identification problem: given the input, output and scheduling parameter sequences  $u_k$ ,  $y_k$  and  $\mu_k$  respectively, with  $k \in \{1, \dots, N\}$ , identify, up to a global similarity transformation, the system matrices  $A^{(i)}$ ,  $B^{(i)}$ ,  $K^{(i)}$  and  $C$ , with  $i \in \{1, \dots, m\}$ .

One of the most interesting algorithms that can handle both open and closed loop systems was presented in van Wingerden and Verhaegen [2009]. In this paper the LPV-PBSID<sub>opt</sub> algorithm was developed, following a rationale similar to the PBSID<sub>opt</sub> algorithm for LTI systems [Chiuso, 2007].

As we already mentioned, in this paper we will follow a Bayesian paradigm for the identification of an LPV state-space model, following the development of similar ideas in Rasmussen and Williams [2005], Golabi et al. [2014]. More specifically, instead of postulating a parametric description for the predictor's impulse response, we employ a non-parametric Gaussian regression approach. To this end, the following definition will be useful.

*Definition 1.* Given (3)-(4), the output is described by

$$y_k = \sum_{t=1}^{\infty} h^u(\mu_{k-t}, \dots, \mu_{k-1}; t) u_{k-t} + \sum_{t=1}^{\infty} h^y(\mu_{k-t}, \dots, \mu_{k-1}; t) y_{k-t} + e_k, \quad (5)$$

where the exact relationship between the predictor's impulse response coefficients  $h^u$ ,  $h^y$  and the quantities in (3)-(4) will be established in Lemma 4. With the notation in (5) we want to make clear that the predictor's impulse response is a function of the instant  $t$ , as well as of various  $\mu$  coefficients, the number of which is depending on  $t$ .

*Remark 1.* From this point onwards to keep the notation tractable we will present the framework for a SISO system. For a MIMO system we can include the additional signals in (5) in an obvious way. Moreover, the steps described

in this paper have to be followed for each output channel separately.

The relationship between the predictor's impulse response coefficients  $h^u$  and  $h^y$  and the scheduling parameters together with the system matrices will become clear with the use of the following definitions.

*Definition 2.* By first defining the matrices

$$\begin{aligned} \mathcal{L}_1^u &= \left[ B^{(1)}, \dots, B^{(m)} \right], \\ \mathcal{L}_1^y &= \left[ K^{(1)}, \dots, K^{(m)} \right], \end{aligned} \quad (6)$$

we extend this notation, as it is performed in van Wingerden and Verhaegen [2009].

$$\begin{aligned} \mathcal{L}_i^u &= \left[ \tilde{A}^{(1)} \mathcal{L}_{i-1}^u, \dots, \tilde{A}^{(m)} \mathcal{L}_{i-1}^u \right], \\ \mathcal{L}_i^y &= \left[ \tilde{A}^{(1)} \mathcal{L}_{i-1}^y, \dots, \tilde{A}^{(m)} \mathcal{L}_{i-1}^y \right], \end{aligned} \quad (7)$$

where  $\mathcal{L}_i^y, \mathcal{L}_i^u \in \mathbb{R}^{n \times m^i}$ .

*Definition 3.* We define the  $\mu$  dependent vector  $P_{t|k}$  as

$$P_{t|k} = \mu_{k-1} \otimes \dots \otimes \mu_{k-t}, \quad P_{t|k} \in \mathbb{R}^{m^t}, \quad (8)$$

where  $\mu_k = \left[ 1, \mu_k^{(2)}, \dots, \mu_k^{(m)} \right]^T$  and  $\otimes$  represents the Kronecker product.

Finally, we have to make use of the transition matrix for discrete-time varying systems, defined as

$$\begin{aligned} \phi_{k,j} &= \tilde{A}_{k-1} \dots \tilde{A}_{k-j+1} \tilde{A}_{k-j} \\ \text{where } \tilde{A}_k &= \sum_{i=1}^m \mu_k^{(i)} \tilde{A}^{(i)}. \end{aligned} \quad (9)$$

Now we can show the relationship between Definitions 1, 2 and 3. Here we will present the relationship for the  $h^u$  coefficients, while a similar approach holds for  $h^y$ .

*Lemma 4.* The impulse response coefficients  $h^u(\mu_{k-t}, \dots, \mu_{k-1}; t)$  are given by

$$h^u(\mu_{k-t}, \dots, \mu_{k-1}; t) = C \mathcal{L}_t^u P_{t|k}. \quad (10)$$

**Proof.** The result in (10) can be obtained by propagating (3)-(4) to the past and taking into account that for a stable predictor matrix  $\tilde{A}_k$ ,  $\lim_{j \rightarrow \infty} \phi_{k,j} \rightarrow 0$ .

In reality, only a finite number of predictor's impulse response coefficients can be used in the algorithm, which will be called the past window ( $p$ ). We will assume, though, that the value of  $p$  is large enough to capture the dynamics of the system and that  $\phi_{k,j} \approx 0$  for  $j > p$ . The output signals that will be used in the algorithm are stacked in the matrix  $Y$ , that is to say,

$$Y = [y_{p+1}, y_{p+2}, \dots, y_N]. \quad (11)$$

Finally, we also assume that the following condition holds.

$$\text{rank}([P_{t|p+1}, P_{t|p+2}, \dots, P_{t|N}]) = m^t \text{ for each } t \in \mathbb{N}^+, \quad (12)$$

where  $p, N \in \mathbb{N}^+$ ,  $p < N$  and  $N - p \geq m^t$ . This can be perceived as a persistency of excitation condition and it is needed to uniquely determine the matrices in Definition 2.

### 3. MODELLING AND ESTIMATION OF PREDICTOR'S IMPULSE RESPONSE COEFFICIENTS

In this section we will introduce a mathematical framework based on Gaussian processes to approximate the predictor's impulse response of the LPV system. The unknown hyperparameters will be estimated by a non-convex algorithm and the values of the Gaussian processes at the data points of the training set will then be computed.

#### 3.1 Prior models for $h^u$ and $h^y$

Following a Bayesian approach, we will model the predictor's impulse response coefficients as zero-mean Gaussian processes [Rasmussen and Williams, 2005]. We will assume that coefficients that correspond to different impulse response instants  $t$  are uncorrelated and we will make the same assumption for all the predictor's impulse response coefficients that correspond to different signals. Mathematically, these conditions are expressed as

$$\begin{aligned} \mathbb{E}[h^u(\mu_{k-t}, \dots, \mu_{k-1}; t)h^u(\mu_{k'-t'}, \dots, \mu_{k'-1}; t')] &= 0, \\ &\text{for every } k, k' \in \mathbb{N}, \text{ if } t \neq t', \\ \mathbb{E}[h^u(\mu_{k-t}, \dots, \mu_{k-1}; t)h^y(\mu_{k'-t'}, \dots, \mu_{k'-1}; t')] &= 0, \\ &\text{for every } t, t', k, k' \in \mathbb{N}, \end{aligned} \quad (13)$$

while similar conditions hold for  $h^y$ . These conditions define a simple kernel structure, similar to the diagonal kernel for LTI systems [Chen et al., 2012]. Moreover, we will use a Radial Basis Function (RBF) to characterize the kernel function, using as inputs the scheduling parameters [Golabi et al., 2014]. Its most celebrated characteristic is the so called non-degeneracy [Rasmussen and Williams, 2005], that is to say, it can be expressed as a linear combination of possibly infinite basis functions. In practice, though, this expressibility is limited by the number of the available data points. Due to the affine property, the scheduling parameter  $\mu_k^{(1)}$  is always 1 for every  $k \in \mathbb{N}$  and for this reason it will not be used as an input in the RBF. Due to the dynamic dependency on the  $\mu$  coefficients in (10), we will define the RBF kernel as (here we show the case where  $m = 2$ )

$$\begin{aligned} [K_t^u]_{i,j} &= \mathbb{E}[h^u(\mu_{p+i-t}, \dots, \mu_{p+i-1}; t) \\ &\quad h^u(\mu_{p+j-t}, \dots, \mu_{p+j-1}; t)] \\ &= \lambda_u^2 \exp \left( - \frac{\left\| \begin{bmatrix} \mu_{p+i-t} - \mu_{p+j-t} \\ \vdots \\ \mu_{p+i-1} - \mu_{p+j-1} \end{bmatrix} \right\|_2^2}{\sigma_t^2} \right). \end{aligned} \quad (14)$$

where  $i, j = \{1, \dots, N - p\}$ . The brackets with the right bottom index refer to a specific position in a matrix. With the use of (14) we can construct the kernel matrices  $K_t^u$  and  $K_t^y$  for  $t = \{1, \dots, p\}$ , while each of these matrices will be of dimensions  $(N - p) \times (N - p)$ . All the related hyperparameters  $\eta = [\lambda_u, \lambda_y, \sigma_t]$  are unknown and have to be estimated from the available data.

At this point, we have to make the following remarks based on (14):

- We assume that  $\sigma_t$  is the same for both  $h^u$  and  $h^y$  as long as they refer to the same impulse response instant  $t$ . This is related to the fact that the characteristics of each predictor's impulse response instant are mainly attributed to the  $\tilde{A}^{(i)}$  coefficients, which enter in the same way in both equations, see (6),(7) and (10).
- The coefficients  $\lambda_u$  (and respectively for  $\lambda_y$ ) are assumed to be the same for all  $K_t^u$ , irrespective of the value of  $t$  (similarly for  $K_t^y$ ). The intuition behind this approach is that each signal is multiplied by some specific matrices, which are the same for every coefficient  $t$  ( $B^{(i)}$  for the inputs and  $K^{(i)}$  for the outputs), so we could see these matrices as scaling factors. This is exactly the role of the  $\lambda_u$  coefficients in the RBF kernel and so it is convenient to establish this correlation between them.

#### 3.2 Estimation of the hyper-parameters

In order to estimate the hyperparameters  $\eta$  we will make use of the minus log marginal likelihood and we will solve a non-convex optimization problem to minimize its value [Chen et al., 2012, Pillonetto et al., 2014].

For this purpose, we will first write down the equations for the mean and the variance of the output vector  $Y$ . In order to do so, we have to define the following matrix:

*Definition 5.* The matrix  $\mathcal{K}$  is of dimensions  $(N - p) \times (N - p)$  and its value at row  $i$  and column  $j$  is given by the expression

$$\begin{aligned} [\mathcal{K}]_{i,j} &= \sum_{t=1}^p u_{i+p-t} [K_t^u]_{i,j} u_{j+p-t} \\ &\quad + \sum_{t=1}^p y_{i+p-t} [K_t^y]_{i,j} y_{j+p-t}. \end{aligned} \quad (15)$$

Based on this definition and with the use of (14) we can construct the matrix  $\mathcal{K}$ . Subsequently, we can describe the statistical properties of the outputs:

*Lemma 6.* The mean  $m_Y$  and the covariance  $\Sigma_Y$  of the output vector  $Y$  (under the assumption that  $h^u$  and  $h^y$  are uncorrelated with the innovation sequence and that  $h^u$  and  $h^y$  are zero when  $t > p$  for a large enough  $p$ ) are expressed as

$$Y \sim \mathcal{N}(0, \mathcal{K} + \sigma^2 I_{N-p}). \quad (16)$$

**Proof.** The proof follows from straightforward calculations, by taking into account the statistical properties of the noise variance and the coefficients  $h^u$  and  $h^y$ , given

in (13), see also [Rasmussen and Williams, 2005, Papoulis and Pillai, 2002]. Similar results can also be obtained even when the condition that  $h^u$  and  $h^y$  are zero for  $t > p$  does not hold (see [Pillonetto et al., 2011] for a detailed treatment).

Now that we have defined all the required quantities, we will determine the unknown hyperparameters via the minimization of minus the logarithm of the marginal likelihood function. This is a common way to set the values of the hyperparameters, due to the apprehensible framework. Moreover, due to the assumption on the normal distribution of all the involved random variables, the marginal likelihood can be computed in an analytical way. The objective function that we will minimize is given by

$$J(\eta) := -\log(p(Y|\eta)) = \frac{N-p}{2} \log(2\pi) + \frac{1}{2} \log(\det(\Sigma_Y(\eta))) + \frac{1}{2} Y \Sigma_Y^{-1}(\eta) Y^T. \quad (17)$$

*Remark 2.* The noise variance  $\sigma^2$  can also be treated as a hyperparameter and it can be estimated together with  $\eta$ . Alternatively we can pre-estimate its value in a way similar to the one described for LTI systems in Goodwin et al. [1992].

### 3.3 Estimation of $h^u$ and $h^y$

The next step is to compute the values of  $h^u(\mu_{k-t}, \dots, \mu_{k-1}; t)$  and  $h^y(\mu_{k-t}, \dots, \mu_{k-1}; t)$  for  $t = \{1, \dots, p\}$ , evaluated at the training points  $k = \{p+1, \dots, N\}$ . To do so, we are looking for the a posteriori estimates  $p(h^y|Y, \eta)$  and  $p(h^u|Y, \eta)$ . We note that the correspondence of this method with the Tikhonov regularization in a Reproducing Kernel Hilbert Space can also be proven [Rasmussen and Williams, 2005].

The a posteriori estimate can be derived in an analytical way [Chen et al., 2012]. The value of a Gaussian process in a specific training point for  $h^u$  (similarly for  $h^y$ ) is given by

$$h^u(\mu_{k-t}, \dots, \mu_{k-1}; t) = Y(\mathcal{K} + \sigma^2 I_{N-p})^{-1} \begin{bmatrix} u_{k-t} [K_t^u]_{1,k-p} \\ u_{k-t+1} [K_t^u]_{2,k-p} \\ \vdots \\ u_{k-t+N-p-1} [K_t^u]_{N-p,k-p} \end{bmatrix}. \quad (18)$$

With the use of (18) we can compute the value of  $h^u$  (similarly for  $h^y$ ) for each  $k = \{p+1, \dots, N\}$  and  $t = \{1, \dots, p\}$ .

## 4. ESTIMATION OF THE SYSTEM PARAMETERS

### 4.1 Estimation of the system parameters

Following the estimation of the hyperparameters with the use of (18), we now need to estimate the unknown coefficients which are contained in  $\mathcal{L}_1^u, \mathcal{L}_1^y, \dots, \mathcal{L}_p^u, \mathcal{L}_p^y$ . Using the result of Lemma 4 we can compute a unique

solution for each  $\mathcal{L}_t^u$  (and similarly for each  $\mathcal{L}_t^y$ ) with the use of the Least Squares (LS) method. The equality that corresponds to this problem is

$$\begin{bmatrix} h^u(\mu_{p+1-t}, \dots, \mu_p; t) \\ h^u(\mu_{p+2-t}, \dots, \mu_{p+1}; t) \\ \vdots \\ h^u(\mu_{N-t}, \dots, \mu_{N-1}; t) \end{bmatrix}^T = C \mathcal{L}_t^u \begin{bmatrix} P_{t|p+1} \\ P_{t|p+2} \\ \vdots \\ P_{t|N} \end{bmatrix}^T. \quad (19)$$

The LS problem, based on (19), has to be solved for each  $t$  (similarly for  $h^y$ ), while the condition in (12) is necessary in order to derive a unique solution. After having estimated the system parameters, the algorithm proceeds as it is explained in sections 4.2 and 4.3 of van Wingerden and Verhaegen [2009], that is to say, we estimate the state sequence and then we solve two LS problems to derive the system matrices,  $A^{(i)}, B^{(i)}, C$ ,  $i = \{1, \dots, m\}$ .

### 4.2 Summary of the algorithm

Here we summarize the algorithm steps, as they were presented in the previous sections. From now on we will refer to this algorithm as LPV-Kernel & PBSID<sub>opt</sub> or LPV-K&PBSID<sub>opt</sub>.

*Algorithm 1.* LPV-K&PBSID<sub>opt</sub>

- (1) Create the matrix  $Y$  and the kernel  $\mathcal{K}$  as it was described in (11) and (15).
- (2) Determine the value of the hyperparameters that minimize the value of (17).
- (3) Estimate the value of the Gaussian processes  $h^u$  and  $h^y$ , evaluated at the training points  $k = \{p+1, \dots, N\}$  based on (18).
- (4) Solve the related Least Squares problems that correspond to (19) in order to estimate the system parameters  $\mathcal{L}_t^u, \mathcal{L}_t^y$ ,  $t = \{1, \dots, p\}$ .
- (5) Proceed as it is explained in sections 4.2 and 4.3 of van Wingerden and Verhaegen [2009] to estimate the state sequence and then the unknown matrices of the model (1)-(2).

## 5. SIMULATION STUDY

In this section we will evaluate the validity of the proposed algorithm and will compare the simulation results of this algorithm with those of LPV-PBSID<sub>opt</sub> algorithm. We will also compare these algorithms with the LPV-PBSID<sub>opt</sub>, employed with Tikhonov regularization based on a Generalized Cross Validation (GCV) regularization parameter selection, as it was discussed in Verdult and Verhaegen [2005]. The system used in the first example is a second order SISO system and the one in the second example is a fourth order MIMO system, while both of them do not have any feed-through term.

### 5.1 Identification of a SISO system

As an example we consider a second order LPV model, which is a modified version of a model for the flapping dynamics of a wind turbine, also used as an example in Felici et al. [2007]. The system matrices of the data-generating LPV state-space model are given by

$$\begin{aligned} [A^{(1)}|A^{(2)}] &= \begin{bmatrix} 0 & 0.0734 & -0.0021 & 0 \\ -6.5229 & -0.4997 & -0.0138 & 0.5196 \end{bmatrix}, \\ [B^{(1)}|B^{(2)}] &= \begin{bmatrix} -0.7221 & 0 \\ -9.6277 & 0 \end{bmatrix}, \\ [C] &= [1 \ 0], \\ [K^{(1)}|K^{(2)}] &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

The LPV system is excited by a white noise input signal  $u_k$  with  $E[u_k] = 0$  and  $\text{var}(u_k) = 1$ . The scheduling parameter is given by

$$\mu_k = \left[ 1, \cos\left(\frac{2\pi}{N}k\right) + 0.2 + 0.1v_k \right]^T,$$

where  $v_k$  is a zero mean white noise sequence with auto-covariance equal to one.

We simulated the system with  $N = 1000$  data points, while we used a future window of  $f = 3$ . The Signal-to-Noise Ratio (SNR) between the noiseless system output variance and the noise variance was chosen to be 1 (0 dB) and 10 (10 dB). For each presented case study we performed 50 Monte Carlo simulations, generating a fresh input and noise signal for each one of them.

In the LPV-K&PBSID<sub>opt</sub> algorithm we pre-estimated the noise variance  $\sigma^2$  and so it was not treated as a hyperparameter in (19) (see also Remark 2).

For the evaluation of the results, we simulated the system without noise (SNR=∞) using a fresh input and noise dataset, while we used the Variance Accounted For (VAF) criterion to quantify the validation results. The equation that expresses the VAF between the measured output  $y_k$  and the estimated output obtained by simulating the identified LPV system  $\hat{y}_k$  is given by

$$VAF(y_k, \hat{y}_k) = \max \left\{ 1 - \frac{\text{var}(y_k - \hat{y}_k)}{\text{var}(y_k)}, 0 \right\} 100\%. \quad (20)$$

In Fig.1 we show the estimated eigenvalues for the models that were identified by the GCV Tikhonov based LPV-PBSID<sub>opt</sub> and the LPV-K&PBSID<sub>opt</sub> algorithms. These results correspond to the case where SNR = 10 and the past window value is  $p = 6$ . Moreover, in Table 1 we present the average VAF value of the three discussed methods for different values of the past window  $p$  (for the choice of the past and future window values see [van der Veen et al., 2013] and the references therein).

### 5.2 Identification of MIMO system

As a second example we simulated a MIMO system, the system matrices of which are given by

Table 1. Average VAF results of the three discussed LPV subspace identification algorithms for different past window and SNR ratio values.

Past Window $p = 4$			
	PBSID <sub>opt</sub>	PBSID <sub>opt</sub> +GCV Tikh.	K&PBSID <sub>opt</sub>
SNR=1	88.93 %	91.08 %	91.86 %
SNR=10	98.47 %	93.05 %	98.52 %
Past Window $p = 5$			
	PBSID <sub>opt</sub>	PBSID <sub>opt</sub> +GCV Tikh.	K&PBSID <sub>opt</sub>
SNR=1	74.22 %	91.94 %	96.40 %
SNR=10	97.98 %	93.30 %	99.24 %
Past Window $p = 6$			
	PBSID <sub>opt</sub>	PBSID <sub>opt</sub> +GCV Tikh.	K&PBSID <sub>opt</sub>
SNR=1	14.89 %	92.40 %	96.35 %
SNR=10	72.99 %	93.64 %	99.28 %

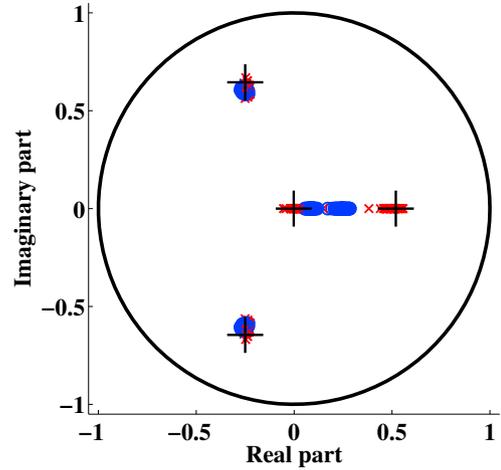


Fig. 1. Eigenvalues of the estimated  $A^{(1)}$  and  $A^{(2)}$  matrices, for 50 experiments with SNR=10 and  $p=6$ . The big crosses correspond to the real values of the matrices, while the blue circles correspond to the LPV-PBSID+GCV method and the red crosses to the LPV-K&PBSID.

$$\begin{aligned} [A^{(1)}] &= \begin{bmatrix} -1.3 & -0.6325 & -0.1115 & 0.0596 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ [A^{(2)}] &= \begin{bmatrix} -0.51 & -0.1075 & -0.007275 & -6.25 \cdot 10^{-5} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ [B^{(1)}|B^{(2)}] &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0.3 & 0.3 \end{bmatrix}, \\ [C] &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ [K] &= \begin{bmatrix} 0.16 & 0 & 0 \\ 0 & 0.16 & 0 \\ 0 & 0 & 0.16 \\ 0.16 & 0 & 0 \end{bmatrix}. \end{aligned}$$

For the identification of this system we used  $N = 1000$  data points, the past window was chosen to be  $p = 5$  and the future window  $f = 3$ . We performed 50 Monte Carlo simulations, for which the scheduling parameter is given by  $\mu_k = [1, v_k]$ , where  $v_k$  is a white noise sequence with unit variance. The SNR value is chosen to be equal to 10. The first input is a filtered version of white noise with variance equal to 1.9, while the second input is a white noise sequence with variance equal to 1. The results for the three outputs are given in Table 2.

Table 2. Average VAF results of the three discussed LPV subspace identification algorithms for each output signal of the second example.

Output	PBSID <sub>opt</sub>	PBSID <sub>opt</sub> +GCV Tikh.	K&PBSID <sub>opt</sub>
#1	55.62 %	72.75 %	88.1%
#2	54.97 %	72.59 %	88.15%
#3	55.90 %	73.06 %	88.37%

The results verify that the new algorithm can provide a more accurate model, both in terms of VAF as well as in how close the estimated poles are to the actual ones, especially when the past window is large enough to adequately model the predictor's impulse response of the LPV system.

## 6. CONCLUSION

In this paper we proposed and developed a framework for the subspace identification of LPV models in affine state-space form, which uses the LPV-PBSID<sub>opt</sub> algorithm as a basis. Results have shown that the new algorithm outperforms the LPV-PBSID<sub>opt</sub> method, with the improvements being larger for larger past window values.

Future work includes the investigation of the effect of the non-convex optimization parameters (kernel hyperparameters) in the accuracy of the derived state-space model, as well as verification of the effectiveness of LPV-K&PBSID<sub>opt</sub> method in more complex systems. More specifically, the kernel structure is a crucial aspect of this algorithm. Therefore, we intend to establish more accurate links between the underlying predictor's impulse response coefficients and the kernel structure, as well as to enable a richer kernel structure than the one characterized by the conditions in (13). The latter requires the introduction of a correlation between the coefficients that correspond to different instants  $t = \{1, \dots, p\}$ . Finally, we are planning to provide further comparisons between the proposed algorithm and other algorithms for the identification of LPV systems.

## REFERENCES

- T. Chen, H. Ohlsson, and L. Ljung. On the estimation of transfer functions, regularizations and Gaussian processes revisited. *Automatica*, 48(8):1525–1535, 2012.
- A. Chiuso. The role of vector autoregressive modeling in predictor-based subspace identification. *Automatica*, 43(6):1034–1048, 2007.
- F. Felici, J.W. Van Wingerden, and M. Verhaegen. Subspace identification of MIMO LPV systems using a periodic scheduling sequence. *Automatica*, 43(10):1684–1697, 2007.
- A. Golabi, N. Meskin, R. Tóth, and J. Mohammadpour. A bayesian approach for estimation of linear-regression LPV models. In *Decision and Control (CDC), 2014 53th IEEE Conference on*, pages 2555–2560. IEEE, 2014.
- G. C. Goodwin, M. Gevers, and B. Ninness. Quantifying the error in estimated transfer functions with application to model order selection. *IEEE Transactions on Automatic Control*, 37(7):913–928, Jul 1992.
- L. Ljung. *System identification - Theory for the User*. Prentice-Hall, 1999.
- M. Lovera and G. Mercere. Identification for gain-scheduling: a balanced subspace approach. In *American Control Conference, 2007. ACC '07*, pages 858–863, July 2007.
- M. Lovera, M. Bergamasco, and F. Casella. LPV modelling and identification: An overview. In *Robust Control and Linear Parameter Varying Approaches*, pages 3–24. Springer, 2013.
- A. Papoulis and S.U. Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- G. Pillonetto and G. De Nicolao. A new kernel-based approach for linear system identification. *Automatica*, 46(1):81–93, 2010.
- G. Pillonetto, A. Chiuso, and G. De Nicolao. Prediction error identification of linear systems: a nonparametric Gaussian regression approach. *Automatica*, 47(2):291–305, 2011.
- G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3):657 – 682, 2014.
- C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005. ISBN 026218253X.
- R. Tóth. *Modeling and identification of linear parameter-varying systems*, volume 403. Springer Science & Business Media, 2010.
- R. Tóth, H.S. Abbas, and H. Werner. On the state-space realization of LPV input-output models: Practical approaches. *IEEE Transactions on Control Systems Technology*, 20(1):139–153, 2012.
- G. van der Veen, J.W. van Wingerden, M. Bergamasco, M. Lovera, and M. Verhaegen. Closed-loop subspace identification methods: an overview. *IET Control Theory & Applications*, 7(10):1339–1358, 2013.
- J.W. van Wingerden and M. Verhaegen. Subspace identification of bilinear and LPV systems for open-and closed-loop data. *Automatica*, 45(2):372–381, 2009.
- V. Verdult and M. Verhaegen. Subspace identification of multivariable linear parameter-varying systems. *Automatica*, 38(5):805–814, 2002.
- V. Verdult and M. Verhaegen. Kernel methods for subspace identification of multivariable LPV and bilinear systems. *Automatica*, 41(9):1557 – 1565, 2005.