2014 European Control Conference (ECC)
June 24-27, 2014. Strasbourg, France

A Support Vector Machine-based Method for LPV-ARX Identification with Noisy Scheduling Parameters*

Farshid Abbasi[1], Javad Mohammadpour[1], Roland Tóth[2], Nader Meskin[3]

Abstract— **In this paper, we present a method that utilizes support vector machines (SVM) to identify linear parameter-varying (LPV) auto-regressive exogenous input (ARX) models corrupted by not only noise, but also uncertainties in the LPV scheduling variables. The proposed method employs SVM and takes advantage of the so-called "kernel trick" to allow for the identification of the LPV-ARX model structure solely based on the input-output data. The objective function, as defined in this paper, allows to consider uncertainties related to the LPV scheduling parameters, and hence results in a new formulation that provides a more accurate estimation of the LPV model in the presence of scheduling uncertainties. We further demonstrate the viability of the proposed LPV identification method through numerical examples, where we show that higher best fit rate (BFR) can be achieved under realistic noise conditions using the proposed method compared to the method initially proposed in [6].**

I. Introduction

Identification of linear parameter-varying (LPV) systems has attracted the attention of many researchers within the control systems community (see [1] and many references therein). The basic idea in identifying an LPV model is to introduce a parametrization of the underlying dependency of the model on the scheduling variables in terms of *a priori* chosen set of basis functions. The very first works on LPV system identification assumed a prior knowledge of the basis functions and focused on the identification of the unknown parameters [2], [3]. In those early works, the problem of finding unknown parameters was simply formulated as a least-squares (LS) problem. Making the assumption that the model structure is known is sometimes valid since the LPV model can be derived directly from the nonlinear system equations; however, this is not always the case and hence additional efforts must be devoted to identify the basis functions. Also, inaccurate selection of the basis functions leads to structural bias while over-parametrization results in a variance increase of the estimates. To resolve the high computational load and bias-variance trade-off arising from over-parameterizations based techniques for least-squares based model estimation, a semi-parametric identification approach based on least-squares support vector machines (LS-SVM) was introduced for a class of nonlinear regression models [4], [5]. Some recent works have been done to address a similar problem for LPV model identification using LS-SVM [6], [7].

Support vector machines are supervised learning tools originated in modern statistical learning theory that can effectively provide a non-parametric estimation of the dependency structure for linear regression based LPV models [16], [17]. The supervised learning method was originally proposed by [15], [16] to rebuild the inherent functional relationships and structures in the data [18]. This non-parametric functional dependence estimation is more successful in coping with the bias-variance trade-off than semi-parametric approaches like dispersion functions methods [6]. Also, considering l_2 loss functions in the LS-SVM approach gives a variation of the original SVM method that presents an effective model structure learning in the LPV setting. Finding computationally efficient and unique solution of the linear problem are the advantages of these slightly different approaches like LS-SVM over original SVM method. Hence our aim in this work is to employ an effective variation of the LS-SVM method combined with a cost function that focuses not only on prediction error, but also weighs possible uncertainties in the system variables.

Accurate knowledge of scheduling signals is a critical assumption in both LPV system identification and LPV control design. The previous works [6], [7] that use the kernel-based SVM for "model learning" assumes the perfect knowledge of the scheduling signal during the system identification process. The questions that we address in this work are: (i) how is the performance of the LPV system identification procedure proposed in [6] affected in the presence of such uncertainties? and (ii) how can we improve the LPV system identification when such uncertainties exist? We will examine the first question through simulation studies. Also, to address the latter question, we model such uncertainties in LPV parameters (that we refer to as "error in variables") and include them in the cost function associated with the underlying optimization problem. In conjunction with SVM, the proposed objective function finds the LPV model structure and the corresponding model coefficients in the presence of error in the variables. This is done using the so-called *kernel trick* approach instead of explicitly defining the feature maps (*i.e.*, basis functions) involved [6].

The rest of this paper is organized as follows: Section II describes the basic formulation for the LPV model identification problem studied here. Sections III presents the

*This work was made possible by the NPRP grant # NPRP 5-574-2-233 from Qatar National Research Fund (QNRF), a member of Qatar Foundation. The statements made here are solely the responsibility of the authors.

[1]Complex Systems Control Laboratory (CSCL), College of Engineering, The Univ. of Georgia, Athens, GA 30602, E-mail: `javadm@uga.edu`

[2]Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands, E-mail: `r.toth@tue.nl`

[3]Dept. of Electrical Engineering, Qatar University, Doha, Qatar, E-mail: `nader.meskin@qu.edu.qa`

proposed identification method (that we refer to as EIV-SVM). Simulation results are shown in Section IV and finally, concluding remarks will be made.

II. Identification of LPV Input/Output Models

We assume that the following SISO deffrence equation defines the behaviour of the data generating system,

$$y(t) = \sum_{i=1}^{n_a} a_i(p(t))y(t-i) + \sum_{j=0}^{n_b} b_j(p(t))u(t-j) + e(t), \tag{1}$$

where t represents the discrete time, y and u are the outputs and inputs of the system, and e represents a white stochastic noise process. We further assume that the coefficients a_i and b_j are dependent on the time-varying scheduling variable(s) $p(t)$. Note that (1) defines an auto regressive with exogenous input (ARX) dynamic structure. For identification of system (1), we will adapt the same model structure where the orders of n_a and n_b are assumed to be known. Commonly, in the LPV system identification, when the number of coefficient functions a_i and b_j is decided, then the dependence of the coefficients on $p(t)$ is parameterized as a linear combination of a finite number of basis functions with static dependence on p chosen *a priori*

$$a_i(p(t)) = \sum_{r=1}^{n_g} \alpha_{i,r}\psi_{i,r}(p(t)) \qquad i = 1, ..., n_a$$

$$b_j(p(t)) = \sum_{r=1}^{n_g} \beta_{j,r}\psi_{j,r}(p(t)) \qquad j = 0, ..., n_b,$$

where $\{\psi_{i,r}\}_{i=1,r=1}^{n_a,n_g}$ and $\{\psi_{j,r}\}_{j=0,r=1}^{n_b,n_g}$ are basis functions of the system coefficients. As described earlier in the paper, since improper selection of basis functions can cause structural bias, best choice of these functions is crucial. Our aim in this paper is to employ the so-called *kernel trick* in order to avoid the difficulties arising from choosing basis functions in a non-systematic way. As described later in the paper, tuning the kernel function parameters has a significant impact on the accuracy of the identified LPV model. In fact, the bias-variance trade-off is tuned, which means achieving a higher accuracy by tuning the parameters causes more sensitivity to noise. We next describe all the coefficients and basis functions in a compact LPV-ARX form and put them in a matrix form. To do so, we first define $x(t)$ as an $n_g = n_a + n_b + 1$ dimensional vector containing all the outputs and inputs as

$$x(t) = \begin{bmatrix} y(t-1) & \dots & y(t-n_a) & u(t) & \dots & u(t-n_b) \end{bmatrix}^\top,$$

and

$$\begin{bmatrix} a_1 & \dots & a_{n_a} & b_0 & \dots & b_{n_b} \end{bmatrix} = \begin{bmatrix} \rho_1^\top \phi_1(p(t)) & \dots & \rho_{n_g}^\top \phi_{n_g}(p(t)) \end{bmatrix},$$

where $\phi_i(p(t))$ is a nonlinear vector map from the scheduling signal space $\mathbb{P}$ to an n_H-dimensional space. ρ_i is a parameter in $\mathbb{R}^{n_H}$. Theoretically, n_H can be infinite, except in parametric LPV identification, where the number of basis functions is set *a priori*. Employing the aforementioned setup, the LPV-ARX model of (1) can be written in a compact form as

$$y(t) = \begin{bmatrix} \rho_1^\top \phi_1(p(t)) & \dots & \rho_{n_g}^\top \phi_{n_g}(p(t)) \end{bmatrix} x(t) + e(t),$$

or

$$y(t) = \rho^\top \Phi + e(t), \tag{2}$$

where

$$\Phi = \begin{bmatrix} \phi_1(p(t))x_1(t) & \dots & \phi_{n_g}(p(t))x_{n_g}(t) \end{bmatrix}^\top. \tag{3}$$

III. LPV Model Identification Using LS-SVM

Least-squares (LS)-based algorithms have been widely utilized for system identification of linear and nonlinear systems in a regression form [9]. In addition, they have been applied for LPV I/O model identification with linear predictors using *a priori* specified parametrization of the dependencies [10], [11]. With the use of LS-SVM for LPV model identification, first proposed in [6], the dependence of the basis functions on the LPV parameters is assumed to be unspecified. The idea behind the work by Tóth *et al.* [6], [7] is that the time-varying coefficients of the LPV model described in an input/output form can be estimated using the so-called *kernel trick* method without assigning specific basis functions. In fact, inherent nonlinearity of the coefficient dependencies can be “learned” efficiently in a projected high-dimensional feature space [6].

A. An LS-SVM Estimator under Uncertain/Noisy Scheduling

In this paper, we extend the work in [6], [7] to develop an SVM-based identification method that can cope with observation/measurement errors in the scheduling variable $p(t)$. To this purpose, we represent the LPV model in a regression form that is appropriate within the SVM setting, as follows

$$y(t) = \sum_{i=1}^{n_g} [\rho_i^\top \phi_i(p(t)) + \Delta v_i(p(t))]x_i(t) + e(t), \tag{4}$$

where Δv_i represents the uncertainties in the i^{th} coefficient function caused by errors by, *e.g.*, the measurement process. The purity ratio of distillation columns that is used as the scheduling parameter in the LPV identification of the process, is an example of roughly measured scheduling variables that always contain some observation/measurement error. We note that Δv_i is naturally different than the environmental noise $e(t)$ that is directly added to the system output. The error-in-variable terms Δv_i, captured in (4), can negatively affect the LPV system identification since the data collected from the system, *i.e.*, $x_i(t)$, are based on noise-free scheduling trajectory actually influencing the system, while the measured scheduling trajectory obtained for the model identification purposes is noisy. To model the impact of error in variables in the SVM formulation, we add these uncertainties directly to the coefficient functions to be identified as

$$y(t) = [a_1 + \Delta a_1 \ \dots \ a_{n_a} + \Delta a_{n_a} \ \ b_0 + \Delta b_0 \ \dots \\ b_{n_b} + \Delta b_{n_b}]x(t) + e(t).$$

Using the basis function formulation of the model coefficients, we have

$$y(t) = [\rho_1^\top \phi_1(p(t)) + \Delta v_1 \ \dots \ \rho_{n_g}^\top \phi_{n_g}(p(t)) + \Delta v_{n_g}]x(t) + e(t) \quad (5)$$

or

$$y(t) = \rho^\top \Phi + \Delta V^\top x(t) + e(t),$$

where Φ was defined by (3) and the error in variables are lumped into a vector ΔV defined by

$$\Delta V^\top = \begin{bmatrix} \Delta v_1 & \dots & \Delta v_{n_g} \end{bmatrix}.$$

Note that ΔV is considered to be stochastic with $\mathbb{E}\{\Delta V\} = 0$.

B. SVM Regression with Error in Variables

To characterize an estimate for the model presented in (4), we propose the following cost function

$$J(\rho, e, \Delta V) = \frac{\gamma}{2} \| \begin{bmatrix} \Delta V & e \end{bmatrix} \|_F + \frac{1}{2} \sum_{i=1}^{n_g} \rho_i^\top \rho_i, \quad (6)$$

which is inspired by the standard cost function used in the total least-squares (TLS) method, that can cope with both error-in-variables and measurement noise [12], [13], [14]. In the cost function above, γ is the regularization parameter. We then expand the matrices and the Frobenius norm and assign different weights (regularization parameters) to ΔV and e resulting in

$$J(\rho, e, \Delta V) = \frac{\gamma_1}{2} \sum_{t=1}^{N} \sum_{i=1}^{n_g} \Delta v_i^\top(t) \Delta v_i(t) + \frac{\gamma_2}{2} \sum_{t=1}^{N} e^2(t) + \frac{1}{2} \sum_{i=1}^{n_g} \rho_i^\top \rho_i. \quad (7)$$

where γ_1 and γ_2 expresses the trade-off between te l_2-loss (prediction error) and l_2-coefficient deviation and regularization in this multi-objective cost function. Note that Δv_i decouples from $e(t)$ due to its correlation with γ in (6).

C. Constrained Optimization Problem

The optimization problem described earlier in this section is solved using the Lagrangian method considering the LPV model in the regression form as the problem constraint. The overall objective is now to solve the following problem

$$\min_{\{\rho, e, \Delta V\}} J(\rho, e, \Delta V) = \frac{\gamma_1}{2} \sum_{t=1}^{N} \sum_{i=1}^{n_g} \Delta v_i^\top(t) \Delta v_i(t) + \frac{\gamma_2}{2} \sum_{t=1}^{N} e^2(t) + \frac{1}{2} \sum_{i=1}^{n_g} \rho_i^\top \rho_i$$

$$s.t. \ \ y(t) = \sum_{i=1}^{n_g} [\rho_i^\top \phi_i(p(t)) + \Delta v_i(p(t))] x_i(t) + e(t).$$

The error function variables can be determined by setting the Lagrangian for this constrained optimization problem as

$$L(\rho, e, \Delta V, \alpha) = J(\rho, e, \Delta V) - \Big[\sum_{t=1}^{N} \alpha_t \sum_{i=1}^{n_g} [\rho_i^\top \phi_i(p(t)) + \Delta v_i(p(t))] x_i(t) + e(t) - y(t) \Big],$$

where α_t's are the Lagrangian multipliers. We then employ the Karush-Kuhn-Tucker (KKT) condition to find the saddle point of L which under the zero-duality gap corresponds also to the optimum of J,

$$\begin{aligned} \frac{\partial L}{\partial \Delta v_i} = 0 \ &\rightarrow \ \Delta v_i(t) = \frac{\alpha_t}{\gamma_1} x_i(t) \\ \frac{\partial L}{\partial e(t)} = 0 \ &\rightarrow \ e_i(t) = \frac{\alpha_t}{\gamma_2} \\ \frac{\partial L}{\partial \rho_i} = 0 \ &\rightarrow \ \rho_i = \sum_{i=1}^{n_g} \alpha_t \phi_i(t) x_i(t) \\ \frac{\partial L}{\partial \alpha_t} = 0 \ &\rightarrow \ e(t) = y(t) - \sum_{i=1}^{n_g} [\rho_i^\top \phi_i(t) + \Delta v_i(t)] x_i(t). \end{aligned} \quad (8)$$

Substituting the obtained variables back into (8) results in

$$y(t) = \sum_{i=1}^{n_g} \big\{ [\sum_{j=1}^{N} \alpha_j x_i(j) \phi_i^\top(j)] \phi_i(t) + \alpha_t \gamma_1^{-1} x_i(t) \big\} x_i(t) + \gamma_2^{-1} \alpha_t.$$

By collecting the related terms together, we have

$$y(t) = \sum_{i=1}^{n_g} \sum_{t=1}^{N} \alpha_j \underbrace{x_i(j) \phi_i^\top(j) \phi_i(t) x_i(t)}_{[\Omega]_{j,t}^i} + \gamma_1^{-1} \alpha_t x_i(t) x_i(t) + \gamma_2^{-1} \alpha_t, \quad (9)$$

where we then define

$$[\Omega]_{j,t} = \sum_{i=1}^{n_g} [\Omega]_{j,t}^i = \sum_{i=1}^{n_g} x_i(j) \phi_i^\top(j) \phi_i(t) x_i(t). \quad (10)$$

that can allow us to write (9) in the matrix form considering the discrete time instants $t = 1, \dots, N$. This leads to the following expression

$$Y = (\Omega + \gamma_1^{-1} \mathrm{diag}(\sum_{i=1}^{n_g} x_i^2(1), \ \dots, \sum_{i=1}^{n_g} x_i^2(N)) + \gamma_2^{-1} I_N) \alpha. \quad (11)$$

Writing the first term of (11) in the kernel form as in [6] yields a systematic way to cope with the basis functions complexity. In fact, this new formulation is based on the *kernel trick* that estimates the inner product of the feature maps in a lower dimensional space without any need to directly define these functions. The elements of the matrix Ω are defined by

$$\begin{aligned} [\Omega]_{j,t}^i &= x_i(j) \phi_i^\top(j) \phi_i(t) x_i(t) \\ &= x_i(j) \langle \phi_i^\top(j), \phi_i(t) \rangle x_i(t) \\ &= x_i(j) (K^i(p(j), p(t))) x_i(t). \end{aligned}$$

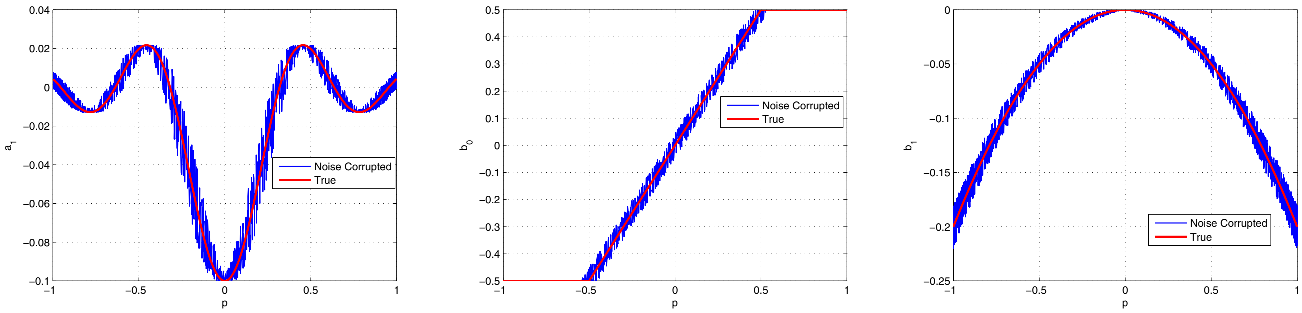

Fig. 1. LPV system coefficients corrupted by noise

where K^i is a positive definite kernel function that satisfies Mercer's conditions in the inner product $\langle \phi_i(j), \phi_i(t) \rangle$ space without explicitly calculating the mapping. In fact, the *kernel trick* only requires the calculation of the modified inner product using every pair of data points and the kernel function's value instead of knowing the basis functions. Although, choosing the most appropriate kernel highly depends on the problem at hand and fine tuning of its parameters can easily become a tedious and cumbersome task, the choice of a particular kernel can be very intuitive and straightforward depending on what kind of information we are expecting to extract from the data. Among various possible choices for kernel functions, the use of *radial basis function (RBF)*, *polynomial*, and *sigmoid function* is appealing due to their ability to represent the nonlinearities in different types of data [15]. In this paper, we use the above three kernel functions and their performance is compared in the next section. The following equation represents the RBF kernel function

$$K^i(p(j), p(t)) = \exp\big(- \frac{\|p(j) - p(t)\|_2^2}{2\sigma_i^2} \big), \qquad (12)$$

where σ_i is an adjustable parameter. The *polynomial* kernels are represented by

$$K^i(p(j), p(t)) = \big(1 + \frac{p(t)^\top p(j)}{c}\big)^d, \qquad (13)$$

where adjustable parameters are the slope c and the polynomial degree d. Finally, the implemented *sigmoid* kernel function is

$$K^i(p(j), p(t)) = \tanh\big(\lambda p(t)^\top p(j) + \beta\big), \qquad (14)$$

where λ and β are the tuning parameters.

After substitution of the chosen kernel function into Eq. (11), the solution to this linear equation is given by

$$\alpha = (\Omega + \gamma_1^{-1}\mathrm{diag}(\sum_{i=1}^{n_g} x_i^2(1), \; \dots, \sum_{i=1}^{n_g} x_i^2(N)) + \gamma_2^{-1} I_N)^{-1} Y.$$

Using the obtained expression for α and the *kernel trick* approach, coefficients of the LPV-ARX model estimate are calculated as

$$a_i(\cdot) = \rho_i^\top \phi_i(\cdot) + \Delta v_i(t) = \sum_{t=1}^{N} \alpha_t x_i(t) K^i(p(t), \cdot) + \frac{\alpha_t}{\gamma_1} x_i(t),$$
$$b_j(\cdot) = \rho_j^\top \phi_j(\cdot) + \Delta v_j(t) = \sum_{t=1}^{N} \alpha_t x_j(t) K^j(p(t), \cdot) + \frac{\alpha_t}{\gamma_1} x_j(t).$$

where N is the number of the measurements.

IV. Simulation Results

In order to evaluate the efficiency of the proposed LS-SVM-based LPV model identification method (that we hereby refer to as "EIV-SVM") when the LPV parameters are corrupted by noise, we apply it to the example in [6]. The following LPV model, in an input/output form is considered:

$$y(t) = a_1(p(t))y(t-1)) + \sum_{i=0}^{1} b_i(p(t))u(t-i) + e_0(t), \qquad (15)$$

where $p(t) \in [-1 \; 1]$. To generate data for identifying the system described by (15), $N = 1500$ samples of data points have been simulated using $u(t) = \sin(\frac{\pi}{2}t)$, $p(t) = \sin(\frac{t}{4})$ and independent and identically distributed (i.i.d.) e_0 with $e_0 \sim \mathcal{U}(-1, 1)$. We also assume that instead of $p(t)$ only $p^*(t) = p(t) + w(t)$ is available to be measured in the system where w is also i.i.d. and $w(t) \sim \eta \times \mathcal{U}(-1, 1)$ where η is a coefficient to control the noise level in the scheduling variable. In the following example, η is assigned 0.05 and 0.1 for the first and second cases, respectively. It should be noted that to avoid clipping of the distribution of η here $p^*(t)$ is allowed to deviate from $[-1 \; 1]$.

The coefficients of the LPV system above are considered to have the following nonlinear dependencies on the scheduling variable

$$b_0(p(t)) = \begin{cases} +0.5 & \text{if } p(t) > 0.5 \\ p(t) & \text{if } -0.5 < p(t) < 0.5 \\ -0.5 & \text{if } p(t) < -0.5 \end{cases}$$

$$b_1(p(t)) = -0.2 \times p_t^2$$
$$a_1(p(t)) = -0.1 \times \frac{\sin(\pi^2 p(t))}{\pi^2 p(t)}.$$

TABLE I

THE MSE AND BFR OF THE EIV-SVM AND LS-BASED SVM METHODS OVER 100 RUNS

		LS-SVM (RBF)		EIV-SVM (RBF)		EIV-SVM (polynomial)		EIV-SVM (sigmoid)	
Noise Level		MSE	BFR	MSE	BFR	MSE	BFR	MSE	BFR
Case I	Mean	5.4905e-04	0.8326	4.9695e-04	0.8867	5.4534e-04	0.8785	6.1577e-04	0.8723
	Std	7.4846e-06	0.0031	6.9000e-06	0.0032	6.5217e-06	0.0035	7.1842e-6	0.0042
Case II	Mean	7.7970e-04	0.7660	6.2417e-04	0.8128	6.6442e-4	0.8046	7.1246e-4	0.8026
	Std	7.3813e-05	0.0068	5.2667e-05	0.0056	6.3196e-4	0.0064	5.9834e-6	0.0062

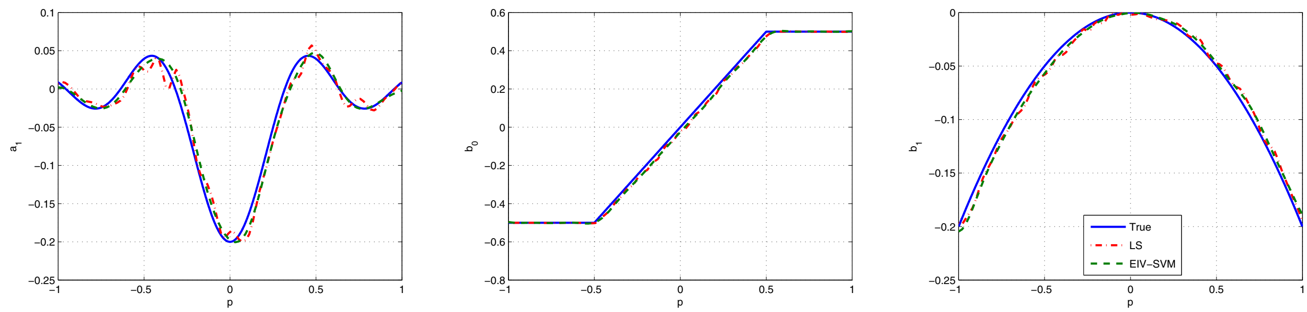

Fig. 2. Comparison of LPV model identification using the proposed method in this paper and that in [6]: LS and EIV-SVM, respectively, represent the LS-SVM based method in [6], and the LS-SVM based method proposed in this paper to cope with the error in variables.

We illustrate two sets of simulation results. First, we compare the accuracy of the LPV model identification approach in this paper with that in [6] considering an RBF kernel function for both cases. As described earlier, in addition to a white Gaussian noise added to the system output with signal to noise ratio of 30dB, another white Gaussian noise is directly added to the scheduling parameter that affects the three LPV model coefficients, as depicted in Figure 1. The results of one run of simulations using the noisy scheduling parameter $p^*(t)$ are illustrated in Figure 2. As observed from the three subplots, the proposed method in this paper outperforms the LS-SVM method in [6] in identifying the three parameter-varying coefficients b_0, b_1 and a_1. It is noted that the hyperparameters γ_1, γ_2, used for the model learning have been tuned with a trial-and-error.

In the second set of simulation results, we compare the three kernel functions described in the previous section to evaluate the performance of the proposed SVM-based model identification approach in the presence of error in variable. To examine the accuracy of the proposed identification method and compare it with the previous work of Tóth *et al.* [6], we consider two error measures of mean square error (MSE) and best fit rate (BFR) defined as

$$\mathrm{MSE} = \frac{1}{N}\sum_{t=1}^{N}(y(t) - \hat{y}(t))^2,$$

$$\mathrm{BFR} = \max\{0, 1 - \frac{\| y(t) - \hat{y}(t) \|_2}{\| y(t) - \bar{y} \|_2}\},$$

where $\bar{y}$ is the mean of the output in the validation data set, $y(t)$, and $\hat{y}(t)$ is the simulated output. Similar to the first simulation, measurements are corrupted by white Gaussian noise and also a white Gaussian noise directly added to the scheduling variable that affects the three LPV model coefficients (as shown in Figure 1). The comparative analysis is done for two different noise levels added to the scheduling parameter $p(t)$. In the two cases examined, noise signals are generated by $\mathcal{U}(-1, 1)$ multiplied by 0.05 (case I) and 0.1 (case II), respectively. A Monte-Carlo simulation study is performed for a numerical illustration of the identification algorithms through changing random white Gaussian noise in the scheduling variable. In addition, we employ three kernels described in the previous section to evaluate the performance of the proposed SVM-based model identification approach. The regularization parameters are selected through trial and error as $\gamma_1 = 1200$ and $\gamma_2 = 6000$. Also, the parameters associated with each one of the three kernel functions were tuned by cross-validation.

The results of 100 runs are analyzed and the mean and standard variations of the BFR and MSE values are shown in Table I indicating that the proposed EIV-SVM method of this paper leads to a better approximation of the LPV model coefficients. In addition, the subplots in Figure 3 illustrate the estimates of the three LPV model coefficients as a function of the LPV parameter $p(t)$ for three kernel functions with the proposed EIV-SVM method. We note that the same error in variable approach, as in the first set of simulations, is considered here. Also, the presented results in Table I and Figure 3 indicate that the RBF kernel (with the tuned parameters $\sigma_1 = \sigma_2 = \sigma_3 = 0.5$) outperforms the other two kernels due to its capability to characterize nonlinearities in the collected data from the LPV model.

To summarize the simulation results, the plots demonstrate that, in the presence of noise in the scheduling variables, the proposed EIV-SVM method exhibits an improved capability of identify the structure of the coefficient functions compared to the LS-SVM method proposed first in [6]. The Monte-

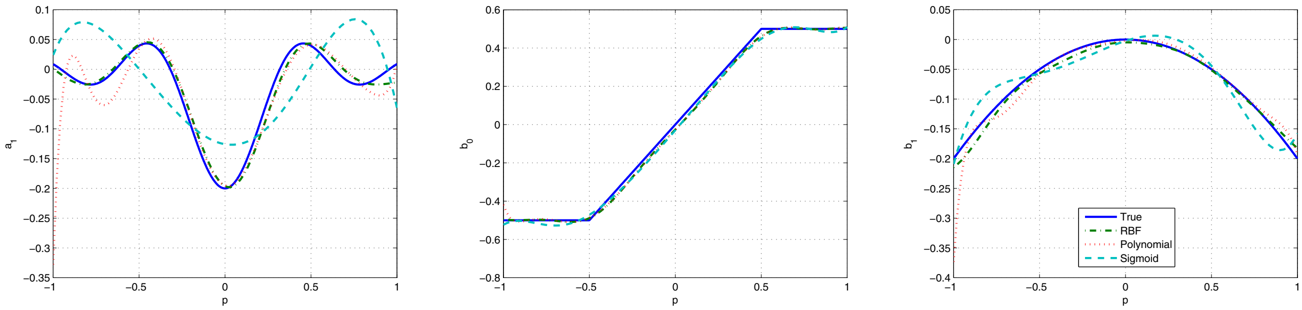

Fig. 3. Estimated LPV model coefficients using RBF, polynomial, and sigmoid kernels: The results illustrate that for the data generated from the given LPV system, the RBF kernel outperforms the other two kernel functions in terms of accurately calculating the coefficient functions.

Carlo simulation results also showed that the proposed EIV-SVM method not only increases the BFR of the estimated output, but also lowers the standard deviation.

CONCLUDING REMARKS

We presented in this paper new results on the extension of LS-SVM as a powerful machine learning tool for model identification of LPV systems in input/output form. The problem was formulated in a way to yield a solution that can handle errors in the scheduling variables. The cost function we defined in the SVM setting included an additional term associated with the errors in variables. This allowed the kernel-based identification method to partially compensate for the error in p to avoid misestimating of the system parameters and lead to a set of new expressions (compared to [6]) for LPV model coefficients by changing the basis functions.

REFERENCES

[1] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*. Berlin: Springer, 2010.

[2] B. Bamieh and L. Giarre, "Identification of linear parameter varying models," *Int. Journal of Robust and Nonlinear Contro*l, vol. 12, pp. 841-853, 2002.

[3] M. Sznaier, C. Mazzaro, and T. Inanc, "An LMI approach to control oriented identification of LPV systems," *Proc. of the American Control Conference*, pp. 3682-3686, July 2000.

[4] I. Goethals, K. Pelckmans, J. Suykens, and B. De Moor, "Identification of MIMO Hammerstein models using least squares support vector machines," *Automatica*, vol. 41, no. 7, pp. 1263-1272, 2005.

[5] V. Laurain, W. Zheng, and R. Tóth, "Introducing instrumental variables in the LS-SVM based identification framework," *in Proc. of the 50th IEEE Conf. on Decision and Control*, pp. 3198-3203, Dec. 2011.

[6] R. Tóth, V. Laurain, W.X. Zheng, and K. Poolla, "Model structure learning: A support vector machine approach for LPV linear-regression models," *in Proc. of the 50th IEEE Conf. on Decision and Control*, pp. 3192-3197, Dec. 2011.

[7] V. Laurain, R. Tóth, W-X. Zheng, and M. Gilson, "Nonparametric identification of LPV models under general noise conditions: an LS-SVM based approach," *Proc. of the 16th IFAC Symposium on System Identification*, pp. 1761-1766, July 2012.

[8] Y.W. Chang, C.J. Hsieh, K.W. Chang, M. Ringgaard, and C.J. Lin., "Training and testing low-degree polynomial data mappings via linear SVM," *J. Machine Learning Research* vol. 11, pp. 1471-1490, 2010.

[9] L. Ljung, *System identification*, Wiley Online Library, 1999.

[10] S. Chena, S. A. Billingsa, and W. Luoa, "Orthogonal least squares methods and their application to nonlinear system identification," *International Journal of Control*, vol. 50, no. 5, pp. 1873-1896, 1989.

[11] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Systems and Control Letters*, vol. 56, no. 5, pp. 373-380, 2007.

[12] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal Processing*, vol. 87, no. 10, pp. 2283-2302, 2007.

[13] J.B. Rosen, H. Park, and J. Glick, "Total least norm formulation and solution for structured problems," *SIAM J. Matrix Analysis and Application*, vol. 17, pp. 110-126, 1996.

[14] P. Lemmerling, N. Mastronardi, and S. Van Huffel, "Fast algorithm for solving the Hankel/Toeplitz structured total least squares problem," *Numerical Algorithms*, vol. 23, pp. 371-392, 2000.

[15] B. Schölkopf and A. Smola, *Learning with kernels*, Cambridge, MA, MIT Press, 2002.

[16] V. Vapnik, *Statistical Learning Theory*, New York: Wiley, 1998.

[17] K. R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Networks*, vol. 12, pp. 181-201, Mar. 2001.

[18] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Knowledge Discovery and Data Mining*, vol. 2, pp. 121-167, June 1998.