

# Perspectives of data-driven LPV modeling of high-purity distillation columns

A. A. Bachnas, R. Tóth, A. Mesbah and J. Ludlage

**Abstract**—This paper investigates data-driven, Linear-Parameter-Varying (LPV) modeling of a high-purity distillation column. Two LPV modeling approaches are studied: a local approach, corresponding to the interpolation of Linear Time-Invariant (LTI) models identified at steady-state purity levels, and a global Least-Square Support Vector Machine (LS-SVM) approach which offers non-parametric estimation of the system w.r.t. data with varying operating conditions. In an extensive simulation study, it is observed that the global LS-SVM approach outperforms the local methodology in capturing the dynamics of the high-purity distillation column under study. The simulation results suggest that the global LS-SVM approach provides a reliable modeling tool under realistic noise conditions.

**Index Terms**—system identification; linear parameter-varying systems; high-purity distillation column.

## I. INTRODUCTION

Chemical processes typically exhibit significant nonlinear behavior when operated over a wide range of operating conditions. Despite the advances of linear modeling and control technologies, it yet remains a challenge to provide high-performance operation of chemical processes, in terms of product quality and process productivity, using a single *linear time-invariant* (LTI) model based controller. In particular, when a chemical process is operated under transient conditions (e.g., set-point changes and start-up procedures) the nonlinear behavior of the process may become dominant, necessitating the use of dedicated nonlinear control solutions. Distillation columns are representative examples of process systems with a severe nonlinear behavior when operated in the high-purity region. These processes, which facilitate component separation of a mixture based on differences in their volatility (i.e., boiling points), are the most widely used separation technique in the chemical industry [1]. The phase change of components in a distillation column drives the most volatile components and the heavy components towards the top and the bottom of the column, respectively, creating a spatial distribution of the components.

To meet with the increasing performance demands of the chemical industry, often modern control design methods, such as model-based control and optimization strategies, e.g.,

*model predictive control* (MPC), are applied to regulate the column behavior. However, these approaches require accurate dynamic models to obtain satisfactory performance and robustness. Modeling of the column behavior based on first-principles commonly leads to a large scale and rather complex nonlinear model involving partial differential equations. Therefore, it appears to be attractive to apply a data-driven modeling procedure (system identification) to arrive at a relatively simple description of the column dynamics. In this problem, the principle question is which (nonlinear) model structures are to be used for such a modeling approach. Furthermore, it is well-known that high-purity distillation columns are particularly challenging systems due to their nonlinear behavior and gain directionality in the high-purity operating region [1]–[3].

In order to extend the validity of LTI models over a wide range of operating conditions of distillation columns and to cope with the nonlinear behavior in the high-purity region, the concept of *linear parameter-varying* (LPV) models appears attractive [4]. As a generalization of the classical concept of *gain scheduling*, this framework is able to model nonlinear process dynamics in a dedicated modeling framework, where a scheduling variable represents the varying operating conditions of the process. In this way, LPV models preserve the advantageous properties of LTI models, while being able to represent a large class of nonlinear systems [5]. Furthermore, the resulting models can be utilized to develop extensions of LTI control strategies, like PID [6], MPC [7], optimal [8] and robust control [9]. A few examples of existing approaches of LPV identification to provide models for these approaches are [10]–[14].

In this paper, we study the applicability of the LPV identification methodology to the data-driven modeling of a high-purity distillation column. We aim to follow two different philosophies to capture the plant dynamics. The first approach, called the *local approach*, is based on the identification of LTI models at several operating points of the process followed by the interpolation of the resulting models. To study the performance of this methodology, we aim to compare several interpolation structures and approaches. The other methodology, called *global approach*, is based on a single data set, exploring a large variety of operating conditions, which is then used to identify the functional dependencies of a linear model structure on the scheduling variable. To study performance of this concept, an LPV *least-square support vector machine* (LS-SVM) approach, the current state-of-the-art methodology to capture large-scale dynamics with unknown nonlinearities, is applied. The main motivation for such a comparative study on distillation

This work was supported by the European Union's 7<sup>th</sup> Framework Programme (FP7/2007-2013) under the grant Autoprofit (grant no.: 257059); ([www.fp7-autoprofit.eu](http://www.fp7-autoprofit.eu)).

A. A. Bachnas, Roland Tóth and J. Ludlage are with the Control Systems Group, Electrical Engineering Department, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, emails: [a.a.bachnas@tue.nl](mailto:a.a.bachnas@tue.nl), [r.toth@tue.nl](mailto:r.toth@tue.nl) and [j.ludlage@tue.nl](mailto:j.ludlage@tue.nl).

A. Mesbah is with the Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA, email: [amesbah@mit.edu](mailto:amesbah@mit.edu).

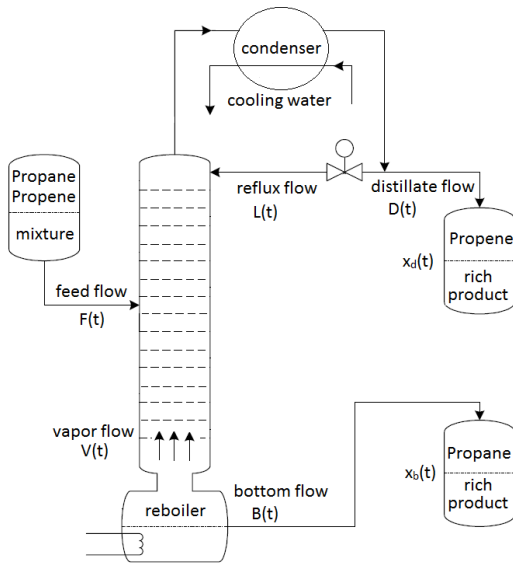


Fig. 1. Schematics of a typical PP-splitter (binary distillation column).

columns is based on [15], [16], where significant difficulties have been reported with the “local” LPV identification of this process operation. We will show why phenomena like changing local model order and directionality together lead to degradation of the estimates obtained in the local sense, while the global SVM approach captures the plant behavior in a highly accurate and efficient manner.

## II. DISTILLATION COLUMN DYNAMICS

The high-purity distillation column studied in this paper is a *propane-propene splitter* (PP-splitter). Based on the principle of boiling point differences of *propane* and *propene*, the PP-splitter is designed to separate the mixture of these substances into its components with a desired purity level. The high-purity products (distillates) are valuable for gasoline production or as raw material for further chemical synthesis. The anatomy of a PP-splitter is depicted in Fig. 1.

### A. First-principles based modeling

To investigate data-driven LPV modeling of the PP-splitter under study, a first-principles based model of the system is developed. This model, which is based on the modeling framework presented in [1] with 110 trays, is used in the sequel as the data-generating system. The primary modeling assumptions are as follows:

- 1) Phase equilibrium on each tray;
- 2) A binary mixture feed;
- 3) Constant relative volatility along the column;
- 4) Constant molar flows;
- 5) Top and bottom flows are ideally controlled (i.e., changes of the vapor and liquid flows have instantaneous effect on the top and bottom flows).

The operation of the distillation column is described in terms of constant liquid molar flows  $L(t)$  propagating from tray to tray towards the bottom re-boiler and constant vapor molar flows  $V(t)$  starting from the re-boiler and propagating through the trays towards the top condenser. The system

dynamics can be represented by a set of *differential algebraic equations* (DAEs) which describe the dynamics of component concentrations as well as the vapor and liquid flows along the 110 trays of the distillation column. A detailed description of the first-principles model and the system parameters can be found in [17], [18].

In the system at hand, the liquid and vapor flow rates in the column (manipulated via the re-boiler duty and reflux rate) are utilized as the manipulated variables to steer the operation of the PP-splitter (i.e.,  $u^{(1)} = V$ ,  $u^{(2)} = L$ ). The control objective is to regulate the bottom ( $x_b$ ) and top composition ( $x_d$ ) in terms of mole fraction of propane (the most volatile component) (i.e.,  $y^{(1)} = x_b$ ,  $y^{(2)} = x_d$ ). Therefore, the splitter is a large-scale (110<sup>th</sup> order) nonlinear  $2 \times 2$  *multi-input multi-output* (MIMO) system.

### B. Motivation for LPV modeling

High-purity distillation columns are well-known for their *nonlinear characteristics* and *directionality* problem which become more significant as the operating conditions approach the *high-purity* region ( $x_d \geq 95\%$  and  $x_b \leq 5\%$ ) [1], [19]. Due to directionality, the system response is dominated by the high-gain direction (only one product composition can be made purer), which significantly limits the performance of linear SISO controllers when trying to control both top and bottom compositions. To address this problem, a MIMO controller, such as MPC, is often used. Such a controller requires an accurate model in order to compute the correct control actions. However, because high-purity distillation columns exhibit strong nonlinear behavior, small deviation from the operating point typically results in different dynamics. This indicates that a linear model may be insufficient to describe the system dynamics over changing operating conditions. Hence, an alternative modeling solution is required which can preserve the simplicity of LTI control synthesis and, at the same time, it can accurately capture the system dynamics over the whole operating regime. As the LPV framework offers such a modeling paradigm, we investigate LPV identification of the PP-splitter and compare alternative LPV schemes for this application.

### C. Measurements

To generate realistic measurement records of the system, the first-principles model is simulated in continuous-time and discrete-time input/output data is collected with a sampling time  $T_s = 2.5$  min (chosen 10 times faster than the time-constant of the fastest step response w.r.t. the possible operating conditions). The inputs of the system are manipulated through zero-order-hold actuation synchronized with the sampling. Furthermore, the measurement data is assumed to be corrupted by an output additive zero mean white noise process  $v$  with a *signal-to-noise ratio* of 25dB.

## III. LPV MODELS

The LPV system class can be seen as an extension of LTI systems as the signal relations are considered to be linear, but the model parameters are assumed to be functions of a time-varying signal, the so-called *scheduling variable*  $p: \mathbb{Z} \rightarrow \mathbb{P}$

with a *scheduling “space”*  $\mathbb{P} \subseteq \mathbb{R}^n$ . This variable is used to indicate the changes in the dynamical signal relations of the plant at different operating conditions. In discrete time, the dynamic description of an LPV system  $\mathcal{S}$  can be formulated in terms of an input-output form (difference equation):

$$y(k) = - \sum_{i=1}^{n_a} a_i(p_k) y(k-i) + \sum_{j=0}^{n_b} b_j(p_k) u(k-j), \quad (1)$$

where  $p_k := p(k)$  and, in the  $2 \times 2$  MIMO case,  $a_i, b_j : \mathbb{P} \rightarrow \mathbb{R}^{2 \times 2}$  are matrix coefficient functions, dependent on  $p_k$ ,  $y(k) = [y^{(1)}(k) \ y^{(2)}(k)]^\top$  and  $u(k) = [u^{(1)}(k) \ u^{(2)}(k)]^\top$ . Note that for the sake of simplicity, all coefficient functions are assumed to have *static dependency*, i.e., being only dependent on the instantaneous value of  $p$  at a given time moment  $k$ . For more details on LPV representations and types of dependencies see [5].

For the PP-splitter, the bottom and top-product compositions are chosen as scheduling variables, i.e.,  $p_k = [x_b(k) \ x_d(k)]^\top$ . This choice is based on the capability of these variables to uniquely characterize the operating point of the system. Next, we identify the PP-splitter by either operating the column at different top and bottom purity levels resulting in data sets with constant  $p$  (local approach) or to vary the purity levels, i.e.,  $p$ , during operation and explore the changing dynamics of the system (global approach).

#### IV. LPV IDENTIFICATION VIA THE LOCAL APPROACH

The local approach refers to the methodology to construct the LPV model as the interpolation of LTI models identified around given operating points. Model interpolation is determined by the applied *interpolation scheme* in terms of the resulting model structure and the used *interpolation method*. Selection of these two items associated methodologies as well as the number and location of the used operating points are application specific (see [5]). The available choices in this respect are explained in the sequel:

##### A. Choice of operating points

The operating envelop of the system under study is described by the purity levels of the top and bottom distillates in terms of  $\mathbb{P} = [0.85, 0.99] \times [0.15, 0.01]$ . Every  $\bar{p} \in \mathbb{P}$  is associated with  $(\bar{y}, \bar{u})$  corresponding to the steady-state solutions of the first-principle model at that purity level. For the considered parameters of the PP-splitter, it has been found via analysis of the change of the local dynamical properties that  $10 \times 10$  equidistant gridding of  $\mathbb{P}$ , in terms of grid points  $\mathcal{P} = \{\bar{p}_i\}_{i=1}^{N_{\text{loc}}=100}$ , is required to represent the global behavior adequately [17]. Note that optimized allocation of  $\{\bar{p}_i\}_{i=1}^{N_{\text{loc}}}$  can seriously lower the number of required LTI experiments [20]. However, we will see that certain properties of the column behavior jeopardize the use of the local approach even in this data-rich setting.

##### B. System identification at the operating points

The strength of the local LPV identification approach is based on the fact that the LTI models can be identified using *prediction error minimization* (PEM) based system

identification which is a well-established theory in the field of process modeling and control [21]. To gather data for identification, at each operating point  $\bar{p}_i \in \mathcal{P}$ , the system is excited by a white noise input with variance 1 around the corresponding  $(\bar{u}_i, \bar{y}_i)$  steady-state values, and, according to the experimental conditions detailed in Section II-C, the response is captured in terms of data sets  $\mathcal{D}_i = \{(u(k), y(k) + v(k))\}_{k=1}^N$  with length  $N = 500$ . This corresponds to 86 days of total experimentation time due to the slow sampling time and equidistant gridding of  $\mathbb{P}$ . Respecting the noise conditions (see Section II-C), the PEM is applied on each  $\mathcal{D}_i$  utilizing an *output error* (OE) model structure. The order of the OE model structure has been chosen to be (2,2) based on correlation analysis conducted on validation data sets gathered with the same experimental conditions.

##### C. Interpolation methods

Interpolation of given discrete data points  $\{\delta_i\}_{i=1}^{N_{\text{loc}}}$  in terms of functions has a long history in numerical analysis with a huge variety of available approaches like linear, polynomial, spline, trigonometric and *radial basis function* (RBF) methods, see, e.g., [22]. In the LPV modeling framework, where the interpolated points can correspond to model outputs or locally identified samples of the coefficient functions  $a_i$  and  $b_j$  (see Section IV-D), the following “standard” approaches are commonly used [5], [14], [16], [23]:

1) *RBF interpolation*: Radial basis function, like Gaussian functions, centered at the operating points  $\bar{p}_i$  can be used as a simple methodology to interpolate  $\{\delta_i\}_{i=1}^{N_{\text{loc}}}$ . The spread of the RBFs is optimized to arrive to a smooth interpolation of the given data points.

2) *Polynomial interpolation*: In polynomial interpolation,  $n^{\text{th}}$ -order monomials (on two indeterminants in this case) are employed as basis functions to obtain an interpolated form of the data points  $\{\delta_i\}_{i=1}^{N_{\text{loc}}}$  in terms of their linear combination. The weights of the basis in the linear combination are optimized in an  $\ell_2$ -sense.

3) *Bilinear interpolation*: This approach corresponds to piece-wise linear interpolation of two variables on a 2D grid. (see [16], [24] for implementation details on these interpolation methods). Note that these approaches can be trivially extended for any scheduling and parameter dimensions.

##### D. Interpolation schemes

Three interpolation schemes are regularly applied to develop LPV models: (i) coefficient, (ii) output and (iii) input interpolation. Due to the space restrictions, here we only investigate scheme (i) and (ii) as scheme (iii) is outperformed by these approaches (see [17] for performance analysis of scheme (iii)). Since these schemes accommodate interpolation of LTI models which operate in deviation variables, the steady-state values  $(\bar{y}_i, \bar{u}_i)$  needs to be incorporated as *trimming values* in these schemes.

(i) *Coefficient interpolation*: By this scheme, the local models are interpolated based on their model coefficients and their input and output trimming values. This corresponds to the interpolation of  $\{\bar{u}_i\}_{i=1}^{N_{\text{loc}}}$ ,  $\{\bar{y}_i\}_{i=1}^{N_{\text{loc}}}$  and of the locally identified values of  $a_i$  and  $b_j$  in terms of  $\{\hat{a}_{i,l}\}_{i=1,l=1}^{n_a, N_{\text{loc}}}$  and

$\{\hat{b}_{j,l}\}_{j=0,l=1}^{n_b,N_{loc}}$  using the previously introduced interpolation approaches. This is possible by introducing a new input signal  $\tilde{u} = [u^\top \ 1]^\top$ . Then each identified local model is written as a  $2 \times 3$  MIMO model:

$$\hat{y}_l(k) = -\sum_{i=1}^{n_a} \hat{a}_{i,l} \hat{y}_l(k-i) + \underbrace{\sum_{j=1}^{n_b} [\hat{b}_{j,l} \quad -\hat{b}_{j,l} \tilde{u}_l]}_{\tilde{b}_{j,l}} \tilde{u}(k-j) + \underbrace{\left[ \hat{b}_0^{(\tau)} \quad (I + \sum_{i=1}^{n_a} \hat{a}_{i,l}) \tilde{y}_l - \hat{b}_{0,l} \tilde{u}_l \right]}_{\tilde{b}_{0,l}} \tilde{u}(k) \quad (2)$$

where  $\hat{y}_l$  is the model provided approximation of the true output  $y$ , and  $\hat{a}_i^{(\tau)}$  and  $\hat{b}_j^{(\tau)}$  are the estimated model parameters. Based on these, the resulting LPV model is given by

$$\hat{y}(k) = -\sum_{i=1}^{n_a} a_i(p_k) \hat{y}(k-i) + \sum_{j=0}^{n_b} b_j(p_k) \tilde{u}(k-j), \quad (3)$$

with  $a_i(\cdot)$  and  $b_j(\cdot)$  being the element-wise interpolation of the matrix sequences  $\{\hat{a}_{i,l}\}_{i=1,l=1}^{n_a,N_{loc}}$  and  $\{\hat{b}_{j,l}\}_{j=0,l=1}^{n_b,N_{loc}}$ .

(ii) *Output Interpolation*: By this scheme, the LPV model is obtained as an interpolated (weighted) function of the output of the local LTI models:

$$\hat{y}(k) = c(p(k), \hat{y}_1(k), \dots, \hat{y}_{N_{loc}}(k)), \quad (4)$$

with  $\hat{y}_i$  being the output of the  $i^{\text{th}}$  local LTI model (with incorporated trimming). Note that in this case, the interpolation points  $\{\delta_i\}_{i=1}^{N_{loc}}$  are described by the sequences  $\hat{y}_i$  and for the RBF and the bilinear case the interpolation of each element of  $c$  is carried out as discussed previously. In the polynomial case, the interpolation decomposes as

$$\hat{y}(k) = c(k) = \sum_{i=1}^{N_{loc}} \sum_{j=1}^{n_w} \theta_{i,j} \psi_j(p_k) (\hat{y}_i(k) + \bar{y}_i), \quad (5)$$

where each  $\psi_j: \mathbb{P} \rightarrow \mathbb{R}$  is a monomial basis in  $p$  and  $\theta_{i,j}$  are the weighting parameters. Note that a particular difference w.r.t. scheme (i) is that here the objective is not to interpolate a given set of discrete data, but to provide a smooth transition from one model output to the other as the operating point changes. Therefore, to optimize such transitions in terms of  $c$ , additional information (data set) about the transient (“inter-sample”) behavior of the system is required. Thus, a data set  $\mathcal{D}_{\text{global}} = \{(u(k), y(k) + v(k), p_k)\}_{k=1}^N$  is assumed to be available with a varying scheduling trajectory. If such a data-set is not available, then  $c$  can be synthesized in terms of bilinear interpolation or with RBF’s with equal spread. On the other hand, using  $\mathcal{D}_{\text{global}}$  requires to minimize an output error criterion  $\mathcal{V}(\mathcal{D}_{\text{global}}, \theta) = \|y(k) - c(k)\|_2^2$  and optimize the parameters of the interpolation function (e.g.,  $\theta_{i,j}$  in (5)) accordingly. Such an interpolation approach, can be considered as a “glocal” approach [25], which borders the set of global LPV identification approaches.

#### E. Identification results by the local approach

The discussed  $2 \times 3$  local modeling approaches have been used to arrive to an LPV model of the distillation column based on locally identified LTI models at the operating points

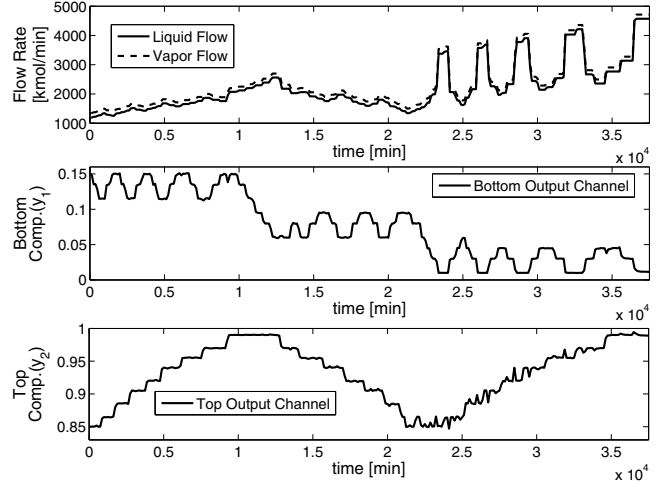


Fig. 2. Applied excitation signals and measured response in  $\mathcal{D}_{\text{global}}$ .

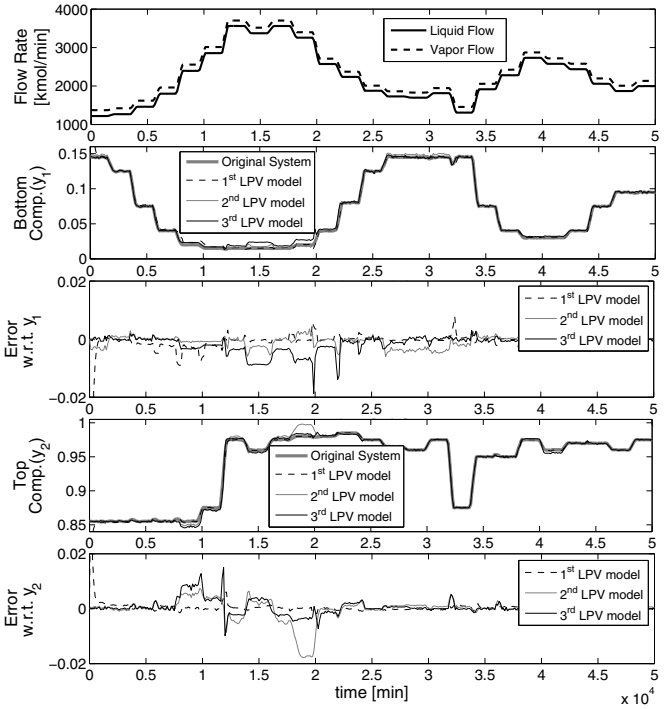


Fig. 3. The noise-free validation data  $\mathcal{D}_{\text{val}}$  based comparison of the true system output and the simulated output of the LPV models estimated via the local approaches (the three best results: output scheme with polynomial interpolation (1<sup>st</sup> model), coefficient scheme with polynomial interpolation (2<sup>nd</sup> model), coefficient scheme with RBF interpolation (3<sup>rd</sup> model)).

$\mathcal{P}$ . The LTI models have been identified as discussed in Section IV-B. For the “glocal” approaches (output interpolation with polynomial or RBF interpolation), a data set  $\mathcal{D}_{\text{global}}$  with  $N = 15000$  is collected from the PP-splitter model. This corresponds to 26 days of experimentation time.  $\mathcal{D}_{\text{global}}$  is generated using an input signal, see Fig. 2, which is able to excite the transient and dynamical behavior of the system over its operating regime. This input signal is a combination of a deterministic component added to a uniform noise with distribution  $\mathcal{U}(-1, 1)$ . The same characteristics for the measurement noise considered in Section II-C are used to collect the local and global data sets. After solving the

TABLE I

$\mathcal{D}_{\text{val}}$  BASED COMPARISON OF THE SIMULATED OUTPUT ERROR OF THE LPV MODELS ESTIMATED VIA THE LOCAL APPROACHES.

		output interpolation		coefficient interpolation	
		MSE	BFR	MSE	BFR
Polynomial interpolation	$y_1$	$3.11 \cdot 10^{-6}$	96.24%	$3.70 \cdot 10^{-6}$	95.89%
	$y_2$	$6.23 \cdot 10^{-7}$	98.28%	$1.58 \cdot 10^{-5}$	91.33%
RBF interpolation	$y_1$	$4.50 \cdot 10^{-5}$	85.68%	$1.10 \cdot 10^{-5}$	92.91%
	$y_2$	$8.07 \cdot 10^{-5}$	80.42%	$8.32 \cdot 10^{-6}$	93.72%
Bilinear interpolation	$y_1$	$4.49 \cdot 10^{-5}$	85.69%	$1.43 \cdot 10^{-5}$	91.92%
	$y_2$	$7.96 \cdot 10^{-5}$	80.55%	$1.09 \cdot 10^{-5}$	92.81%

interpolation problems via direct computation (scheme (i) with the RBF and scheme (i), (ii) with the bilinear approach) quadratic (polynomial approach for scheme (i) and (ii)) and nonlinear optimization (scheme (ii) with the RBF method) the accuracy of the resulting LPV models have been compared in a simulation study using an independent noise-free data set  $\mathcal{D}_{\text{val}} = \{(u(k), y(k), p(k))\}_{k=1}^N$  with  $N = 20000$ , generated by the PP-splitter model with a different varying  $p$  and different realization of  $u$  than in the estimation data set(s). The achieved simulation results are presented in Fig. 4 and in Table I in terms of the *mean squared error* (MSE)

$$\text{MSE} = \|y(k) - \hat{y}(k)\|_2^2. \quad (6)$$

and the *best fit rate* (BFR) :

$$\text{BFR} = 100\% \cdot \max\left(1 - \frac{\|y(k) - \hat{y}(k)\|_2}{\|y(k) - \bar{y}\|_2}, 0\right), \quad (7)$$

where  $\bar{y}$  is the mean of the noise-free output  $y$  of the original system and  $\hat{y}$  is the simulated output of the LPV model calculated on  $\mathcal{D}_{\text{val}}$ . The mean BFR of the identified local models w.r.t. step-response data generated by the linearization of the PP-splitter model at the operating points  $\mathcal{P}$  has been 90.42% and 91.44% for  $y_1$  and  $y_2$  respectively.

Based on these results, the output scheme with polynomial interpolation, has produced the highest BFR on  $\mathcal{D}_{\text{val}}$ . This can be explained by the "glocal" nature of the modeling approach as the interpolation parameters has been optimized w.r.t. both the local and the transient behavior (using  $\mathcal{D}_{\text{global}}$  data set). On the other hand, the output scheme with RBF interpolation, which uses a similar "glocal" method, has produce a lower BFR, which has turned out to be a result of finding the local optimum of the interpolation problem. Significant performance increase has not been obtained by using different initializations in solving the optimization which clearly indicates the complexity of this scheme compared to the quadratic problem of polynomial interpolation. Regarding the coefficient scheme based approaches, the obtained performance has been relatively the same as the local accuracy of the obtained LTI models. Among the interpolation methods, the RBF approach achieved a slightly better result which is due to the fact that this interpolation is capable to ensure that the LPV model behavior is only determined by a single local LTI model if  $p$  is close to the operating point where this model was identified. On the contrary, the polynomial interpolation composes the behavior from all identified local model parameters over the operating regime, resulting in

TABLE II

THE LOCAL  $\mathcal{H}_{\infty}$  ERROR OF THE ESTIMATED LPV MODELS W.R.T. THE TRANSFER FUNCTIONS OF THE LINEARIZED AND DISCRETIZED SYSTEM (COMPUTED ON A  $19 \times 19$  GRIDDING OF  $\mathbb{P}$ ).

	output interpolation		coefficient interpolation	
	Mean	Max	Mean	Max
Poly. int.	$7.62 \cdot 10^{-3}$	$1.25 \cdot 10^{-2}$	$6.21 \cdot 10^{-4}$	$2.29 \cdot 10^{-2}$
RBF int.	$4.79 \cdot 10^{-4}$	$9.79 \cdot 10^{-4}$	$4.71 \cdot 10^{-4}$	$8.74 \cdot 10^{-4}$
Bilin. int.	$3.37 \cdot 10^{-3}$	$1.11 \cdot 10^{-2}$	$4.91 \cdot 10^{-4}$	$1.55 \cdot 10^{-3}$

a smooth, but almost arbitrary transitions of the model parameters from one local operating point to the other. This can even result in an unstable LPV behavior. While bilinear interpolation does not suffer from this problem (like the RBF approach), it introduces no-smooth transitions which produce a worse interpolated behavior than by the RBF method.

However, comparison of the simulated output response w.r.t. a particular trajectory of  $p$  might not fully reveal the quality of the model. To test, how the models can generalize the local dynamics of the system w.r.t. other operating points, the  $\mathcal{H}_{\infty}$  error:

$$\varepsilon_{\bar{p}} \triangleq \|G_{o,\bar{p}}(z) - \hat{G}_{\bar{p}}(z)\|_{\infty}, \quad (8)$$

i.e., the  $\mathcal{H}_{\infty}$  norm of the difference of the transfer function  $G_{o,\bar{p}}(z)$  of the linearized and discretized PP-splitter model at operating condition  $\bar{p} \in \mathbb{P}$  and the (frozen) transfer function of  $\hat{G}_{\bar{p}}(z)$  of the LPV model for a constant scheduling trajectory  $p(k) \equiv \bar{p}$ , is computed on a dense  $19 \times 19$  gridding of  $\mathbb{P}$ . The mean and maximum of  $\varepsilon_{\bar{p}}$  on this grid is displayed in Table II.

This reveals that the coefficient interpolation based approaches result in a significantly smaller local  $\mathcal{H}_{\infty}$  error if compared to the output interpolation based approach. This means that the local properties of the model including stability are better preserved by these approaches. The lower BFR results of the coefficient schemes in Table I, can now be interpreted as the result of the lower nominal performance of the identified local models and not the result of the interpolation algorithm. Hence it is reasonable that the glocal approaches have better BFR, since they can sacrifice the overall "dynamical" quality of the identified model to achieve a higher fit. However, this results in a certain kind of structural over-fitting which distorts the local accuracy of the model and can cause significant problems if a controller is designed on it. Additionally, the glocal method using RBF approach, is also able to preserve the local properties of the model reasonably well, as the basis functions have a region where only a single model is active, and the influence of the other local models is limited unlike in the polynomial approach. This explains the contradictory results reported in [15], [16] with the polynomial interpolation method.

The quality of the resulting LPV models heavily depends on the quality of the local LTI estimates. One way to achieve better results is by increasing the model order in the OE identification. However, it turns out that this results in a significant increase of the parameter variance of the model estimates in the low separation region as higher order (local) modes vanish (become unidentifiable) in this region.

This is undesirable as the interpolation methods will be significantly influenced by the noise and result in random interpolated behavior of the estimated LTI models. This property, called *changing local model order*, jeopardizes any local identification approach which is based on interpolation of IO or state-space models. This also explains why the approach reported in [15], [16] has been applied successfully on other applications without this property. To overcome this problem, the use of a flexible model structure, which is resilient to changing local model order, the so-called LPV *orthonormal basis functions* approach [12] is suggested as a future research direction.

#### V. LPV IDENTIFICATION VIA THE GLOBAL APPROACH

By the global identification concept, an LPV model of the system is estimated in one step based only on a single data set  $\mathcal{D}_{\text{global}}$  with a varying scheduling trajectory. A commonly applied approach to implement this scheme is to use the LPV extension of the PEM framework for LPV-IO models in the form of (1), see [5], [10]. Almost all approaches in that context, including subspace schemes, require linear parameterization of each model coefficient function  $a_i$  and  $b_j$  in terms of an a priori chosen set of basis functions  $\{\psi_j\}_{j=1}^{n_w}$ ,  $\psi_j : \mathbb{P} \rightarrow \mathbb{R}$ :

$$\phi_i = \theta_{i,0} + \sum_{j=1}^{n_w} \theta_{i,j} \psi_j, \quad (9)$$

with  $\phi_i = a_i$  for  $1 \leq i \leq n_a$ ,  $\phi_{n_a+1+i} = b_i$  for  $0 \leq i \leq n_b$  and  $\theta_{i,j}$  being the unknown parameters to be estimated. This method however, requires an efficient selection of  $\{\psi_j\}_{j=1}^{n_w}$  such that (9) is able to capture the underlying nonlinearities of the system adequately. For process systems, to accomplish such a basis function selection procedure, a complicated analysis of the first-principle model is required where the economical benefits of data-driven modeling can be easily lost. Besides, it is not guaranteed that no unexpected nonlinearities show up in the system behavior due to non-ideal operation of actuators or actual morphology of the installation. This has led to an increasing interest for methods capable to provide direct, so-called *non-parametric*, estimation of the coefficient functions [26], [27]. Due to the fact that the original PP-splitter model is described by a large-scale first principle model which, in the “local sense”, can be reduced to the a  $2^{\text{nd}}$ -order model without a significant loss of accuracy (see Section IV-E), it is expected that the original nonlinearities of the model translate to heavy nonlinear  $p$ -dependencies in a low order LPV model. Furthermore, the property of changing local model order also indicates that an a priori choice of basis functions in (9) can seriously influence the model accuracy. Therefore, we apply a recently introduced non-parametric LPV identification method, called the LPV *least-square support vector machine* (LPV LS-SVM) method [26], which can approximate the nonlinear functional dependencies of the model directly from data.

##### A. The LPV LS-SVM approach

To avoid the use of complicated notation, the LPV LS-SVM method is introduced in the sequel using a *multiple*

*input single-output* (MISO) setting. This form of the estimator accommodates identification of the original system for each output channels  $y^{(1)}$  and  $y^{(2)}$  separately. The utilized model structure is

$$y^{(j)}(k) = \sum_{i=1}^{n_g} (\theta_i^{(j)})^\top \psi_i^{(j)}(p_k) x_i^{(j)}(k) + e_k^{(j)}, \quad (10)$$

where  $j \in \mathbb{I}_1^2$ , with  $\mathbb{I}_{s_1}^{s_2} := \{i \in \mathbb{Z} \mid s_1 \leq i \leq s_2\}$  being an index set,  $\psi_i^{(j)} : \mathbb{R} \rightarrow \mathbb{R}^{n_w}$  denotes an undefined, potentially infinite dimensional vector of basis functions,  $\theta_i^{(j)} \in \mathbb{R}^{n_w}$  is the  $i^{\text{th}}$  parameter vector,  $e_k^{(j)}$  is the prediction error, and  $x_i^{(j)}(k)$  is defined as:

$$x_i^{(j)}(k) = y_j(k-i), \quad \text{for } i \in \mathbb{I}_1^{n_a}, \quad (11a)$$

$$x_{n_a+1+i}^{(j)}(k) = u_1(k-i), \quad \text{for } i \in \mathbb{I}_0^{n_b}, \quad (11b)$$

$$x_{n_a+n_b+2+i}^{(j)}(k) = u_2(k-i), \quad \text{for } i \in \mathbb{I}_0^{n_b}. \quad (11c)$$

The estimates of the coefficient functions in the form of  $(\theta_i^{(j)})^\top \psi_i^{(j)}(p_k)$  are obtained via the solution of the following optimization problem(s)

$$\min_{\theta^{(j)}, e^{(j)}} \mathcal{J}(\theta^{(j)}, e^{(j)}) = \frac{1}{2} \sum_{i=1}^{n_g} \|\theta_i^{(j)}\|_2^2 + \frac{\gamma^{(j)}}{2} \sum_{k=1}^N (e_k^{(j)})^2$$

$$\text{s.t.} \quad e_k^{(j)} = y^{(j)}(k) - \sum_{i=1}^{n_g} (\theta_i^{(j)})^\top \psi_i^{(j)}(p_k) x_i^{(j)}(k)$$

with  $\gamma^{(j)} \in \mathbb{R}_0^+$  as the *regularization parameter*, influencing the complexity of the model in terms of the  $\ell_2$ -norm of  $\theta^{(j)}$ . An interesting feature of this optimization problem that it can be solved without specifying the functions  $\psi_i^{(j)}$  or estimating the parameter vectors  $\theta_i^{(j)}$ . What makes this possible is deriving the solution in the dual space of this optimization problem and instead of specifying  $\psi_i^{(j)}$ , its inner product  $(\psi_i^{(j)})^\top \psi_i^{(j)}$  is defined by an a priori chosen kernel function  $K_i^{(j)}$ . This is called the *kernel trick*, see [26]. Hence, the underlying choice of nonlinear basis functions  $\psi_i^{(j)}$  is avoided by specifying a series expansion structure of the unknown coefficient function  $\phi_i^{(j)}$  described by the kernel  $K_i^{(j)}$ . There are many possible kernel functions that can be used in LPV identification, such as linear, polynomial, RBF, etc. For the sake of simplicity, we apply a gaussian RBF kernel which is commonly used in nonlinear function estimation. Hence, the kernel is formulated as:

$$K_i^{(j)}(p_k, p_l) = \exp\left(-\frac{\|p_k - p_l\|_2^2}{(\sigma_i^{(j)})^2}\right) \quad (12)$$

with  $\sigma_i^{(j)} > 0$  specifying the width of the RBF. Together with the regularization parameter  $\gamma^{(j)}$ , the kernel widths  $\sigma_i^{(j)}$  are the tuning parameters of the estimator determining the complexity of the estimated model. Hence, they can be seen as trade-off parameters between bias and variance of the estimate. Based on  $K_i^{(j)}$ , the following matrices are constructed

$$[\Omega^{(j)}]_{s,k} = \sum_{i=1}^{n_a+2n_b+2} [\Omega_i^{(j)}]_{s,k}$$

$$[\Omega_i^{(j)}]_{s,k} = x_i^{(j)}(s) (K_i^{(j)}(p(s), p(k))) x_i^{(j)}(k),$$

TABLE III

$\mathcal{D}_{\text{val}}$  BASED COMPARISON OF THE SIMULATED OUTPUT ERROR OF THE LPV MODEL ESTIMATED VIA THE LS-SVM APPROACH.

Output	BFR	STD w.r.t BFR	MSE	STD w.r.t MSE
$y_1$	99.82%	0.02%	$7.67 \cdot 10^{-9}$	$1.47 \cdot 10^{-9}$
$y_2$	98.23%	0.19%	$7.50 \cdot 10^{-7}$	$1.72 \cdot 10^{-7}$

with  $s, k \in \mathbb{I}_1^N$ . Using these kernel matrices, the dual solution of the SVM estimation problem is computed in terms of the dual parameters  $\beta^{(j)} \in \mathbb{R}^N$ :

$$\hat{\beta}^{(j)} = (\Omega^{(j)} + \gamma^{-1} I_N)^{-1} Y^{(j)}. \quad (13)$$

with  $Y^{(j)} = [y^{(j)}(1) \dots y^{(j)}(N)]^\top$ . Based on  $\hat{\beta}^{(j)}$ , the estimates of the coefficient functions are given as

$$\hat{\phi}_i^{(j)}(\bullet) = (\theta_i^{(j)})^\top \psi_i^{(j)}(\bullet) = \sum_{s=1}^N \hat{\beta}_s^{(j)} x_i(s) K_i^{(j)}(p(s), \bullet). \quad (14)$$

Note that this approach can also be applied to the general MIMO case and any finite scheduling dimensions.

### B. Results of the LPV LS-SVM approach

Compared to the local approaches, the LS-SVM approach provides a direct estimation of LPV model from data, hence it makes sense to assess its stochastic performance. For this reason, the LPV LS-SVM estimator has been tested in a Monte-Carlo study using 100 global data sets generated with the same input and noise conditions as  $\mathcal{D}_{\text{global}}$  (same deterministic input trajectory, but new realizations of the random input component and the noise). In this study, the LS-SVM has been applied on each of these data sets separately to identify 2<sup>nd</sup>-order LPV models for each output channel. To tune the hyper parameters  $\{\gamma^{(j)}\}_{j=1}^2$  and  $\{\sigma_i^{(j)}\}_{i=1, j=1}^{n_a+2n_b+2, 2}$ , an other data set  $\mathcal{D}_{\text{tune}} = \{(u(k), y(k) + v(k), p_k)\}_{k=1}^N$ , with  $N = 15000$ , independently generated from  $\mathcal{D}_{\text{global}}$  and  $\mathcal{D}_{\text{val}}$ , has been recorded from the PP-splitter model using the same input and noise settings. Based on  $\mathcal{D}_{\text{tune}}$ , auto-tuning of the hyper parameters via non-linear optimization has been conducted using the predictor form of the LS-SVM estimator. The resulting optimized values are  $\gamma^{(1)} = 303.91$ ,  $\gamma^{(2)} = 5.74 \cdot 10^4$  and  $\sigma_i^{(1)} = 9.53$ ,  $\sigma_i^{(2)} = 15.84$  using equal width for all kernels. The simulation error of one of the identified LPV models w.r.t.  $\mathcal{D}_{\text{val}}$  in the considered Monte-Carlo study is depicted in Fig. 4. The mean and *standard deviation* (STD) of the BFR and the MSE of this simulation error over the Monte-Carlo study are presented in Table III.

Based on the results, it can be concluded that LS-SVM has obtained models with high average BFR and a low MSE. Furthermore, the STD of these error measures have been also reasonably low over the 100 Monte-Carlo runs, indicating low variance of the model estimates. However, it is important to note that under the OE noise setting, estimation by the basic LPV LS-SVM (which is based on LPV-ARX model structure) will always result in biased estimates even if in this case optimal tuning of the hyper parameters can partly compensate for the bias. The resulting models can be further refined, e.g., by using the instrumental-variable SVM approach introduced in [28] or possible extension of results reported in [29], [30]. However, due the already adequate

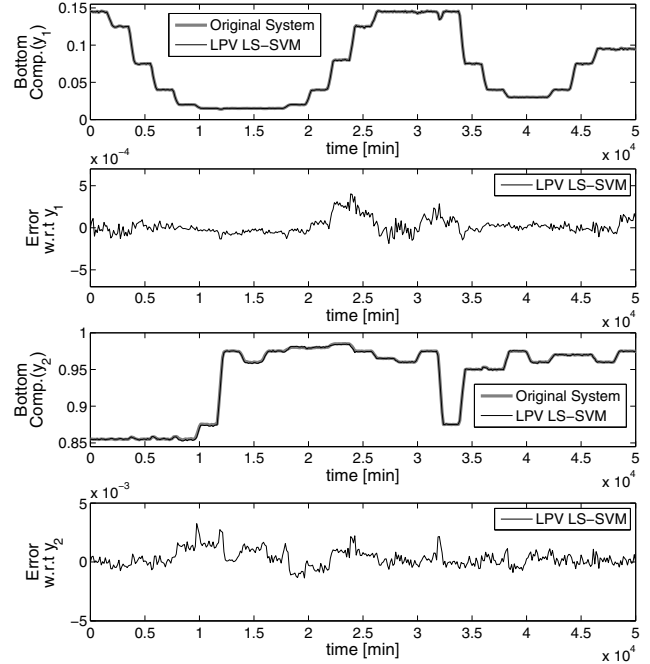


Fig. 4.  $\mathcal{D}_{\text{val}}$  based comparison of the true system output and the simulated output of the LPV model estimated via the global LS-SVM approach.

fit of the model on independently generated validation data, further refinement w.r.t. possible noise bias is not necessary.

## VI. CONCLUSION

Local and global approaches of *linear parameter-varying* (LPV) identification have been compared in the data-driven modeling of a high-purity distillation column. It has been shown that the local, especially output interpolation schemes utilizing data describing transient dynamics of the system, can achieve adequate approximation of the underlying input-output behavior. However, regarding the generalization capability of the resulting models, the coefficient interpolation schemes have shown better performance than the output schemes. It has been also revealed that due to changing local model order of the system, the local approaches encounter with significant difficulties if an LPV input-output or state-space representation based model structure is used. On the other hand, the investigated non-parametric global approach, the least-squares support vector machine, has shown high performance in capturing the input-output behavior. Moreover, it has also shown adequate generalization capabilities in terms of the local dynamical properties of the true system despite the inaccurate noise model introduced bias. With respect to both the local and global approaches, several points of improvement have been pointed out, serving as objectives for further research in studying LPV modeling of process systems.

## REFERENCES

- [1] S. Skogestad and M. Morari, "Control of ill-conditioned plants: High purity distillation," in *Proc. of the AIChE Annual Meeting*, Nov. 1986.
- [2] —, "LV-control of a high-purity distillation column," *Chemical Engineering Science*, vol. 43, no. 1, pp. 33–48, 1988.

- [3] M. V. Finco, W. L. Luyben, and R. E. Polleck, "Control of distillation columns with low relative volatilities," *Industrial & Engineering Chemistry Research*, vol. 28, pp. 75–83, 1989.
- [4] W. Rugh and J. S. Shamma, "Research on gain scheduling," *Automatica*, vol. 36, no. 10, pp. 1401–1425, 2000.
- [5] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, ser. Lecture Notes in Control and Information Sciences, Vol. 403. Heidelberg: Springer, 2010.
- [6] A. Kwiatkowski, H. Werner, J. P. Blath, A. Ali, and M. Schultalbers, "Linear parameter varying PID controller design for charge control of a spark-ignited engine," *Control Engineering Practice*, vol. 17, no. 11, pp. 1307–1317, 2009.
- [7] T. Besselmann, J. Löfberg, and M. Morari, "Explicit model predictive control for linear parameter-varying systems," in *Proc. of the 47th IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 3848–3853.
- [8] A. Packard, "Gain scheduling via linear fractional transformations," *Systems & Control Letters*, vol. 22, no. 2, pp. 79–92, 1994.
- [9] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice-Hall, 1998.
- [10] L. Giarré, D. Bauso, P. Falugi, and B. Bamieh, "LPV model identification for gain scheduling control: An application to rotating stall and surge control problem," *Control Engineering Practice*, vol. 14, no. 4, pp. 351–361, 2006.
- [11] J. W. van Wingerden and M. Verhaegen, "Subspace identification of bilinear and LPV systems for open- and closed-loop data," *Automatica*, vol. 45, no. 2, pp. 372–381, 2009.
- [12] R. Tóth, P. S. C. Heuberger, and P. M. J. Van den Hof, "Asymptotically optimal orthonormal basis functions for LPV system identification," *Automatica*, vol. 45, no. 6, pp. 1359–1370, 2009.
- [13] —, "An LPV identification framework based on orthonormal basis functions," in *Proc. of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009, pp. 1328–1333.
- [14] Y. Zhu and G. Ji, "LPV model identification using blended linear models with given weightings," in *Proc. of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009, pp. 1674–1679.
- [15] J. Huang, G. Ji, and Y. Zhu, "Some study on the identification of multi-model LPV models with two scheduling variables," in *Proc. of the 16th IFAC Symposium on System Identification*, Brussels, Belgium, July 2012.
- [16] J. Huang, G. Ji, Y. Zhu, and P. van den Bosch, "Identification of multi-model LPV models with two scheduling variables," *Journal of Process Control*, vol. 22, no. 7, pp. 1198–1208, 2012.
- [17] A. A. Bachnas, "Linear parameter-varying modelling of a high-purity distillation column," Master's thesis, Delft University of Technology, 2012.
- [18] I. J. Halvorsen and S. Skogestad, *Encyclopedia of Separation Science, Distillation*. Academic Press, 2000.
- [19] E. Jacobsen, P. Lundström, and S. Skogestad, "Modelling and identification for robust control of ill-conditioned plants - a distillation case study," in *Proc. of the American Control Conf.*, Boston, USA, June 1991, pp. 242–248.
- [20] A. A. Khalate, X. Bombois, R. Tóth, and R. Babuška, "Optimal experimental design for LPV identification using a local approach," in *Proc. of the 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009, pp. 162–167.
- [21] L. Ljung, *System Identification, theory for the user*, 2nd ed. Prentice-Hall, 1999.
- [22] K. E. Atkinson, *An Introduction to Numerical Analysis*. John Wiley and Sons, 1989.
- [23] X. Jin, B. Huang, and D. S. Shook, "Multiple model LPV approach to nonlinear process identification with EM algorithm," *Journal of Process Control*, vol. 21, no. 1, pp. 182–193, 2011.
- [24] A. A. Bachnas, R. Tóth, A. Mesbah, and J. Ludlage, "Data-driven LPV modeling of high-purity distillation columns," *Submitted to Journal of Process Control*, 2013.
- [25] G. Mercère, M. Lovera, and E. Laroche, "Identification of a flexible robot manipulator using a linear parameter-varying descriptor state-space structure," in *Proc. of the 50th IEEE Conf. on Decision and Control*, Orlando, Florida, USA, Dec. 2011, pp. 783–790.
- [26] V. Laurain, R. Tóth, W. Zheng, and M. Gilson, "Nonparametric identification of LPV models under general noise conditions. an LS-SVM based approach," in *Proc. of the 16th IFAC Symposium on System Identification*, Brussels, Belgium, July 2012, pp. 1761–1766.
- [27] K. Hsu, T. L. Vincent, and K. Poolla, "Nonparametric methods for the identification of linear parameter varying systems," in *Proc. of the Int. Symposium on Computer-Aided Control System Design*, San Antonio, Texas, USA, Sept. 2008, pp. 846–851.
- [28] 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*, Orlando, Florida, USA, Dec. 2011, pp. 3198–3203.
- [29] J. Suykens and J. Vandewalle, "Recurrent least squares support vector machines," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, no. 7, pp. 1109–1114, 2000.
- [30] T. Falck, J. Suykens, and B. De Moor, "Linear parametric noise models for least squares support vector machines," in *Proc. of the 49th IEEE Conf. on Decision and Control*, Atlanta, USA, Dec. 2010, pp. 6389–6394.