A Bayesian Approach for LPV Model Identification and its Application to Complex Processes

Arash Golabi[1], Nader Meskin[1], Roland Tóth[2] and Javad Mohammadpour[3]

***Abstract*—Obtaining mathematical models that can accurately describe nonlinear dynamics of complex processes and be further used for model-based control design is a challenging task. In this paper, a Bayesian approach is introduced for data-driven identification of linear parameter-varying (LPV) regression models in an input-output (IO) dynamic representation form with an auto-regressive with exogenous variable (ARX) noise structure. The applicability of the proposed approach is then investigated for modeling of complex nonlinear process systems. In this approach, the dependency structure of the model on the scheduling variables is identified based on a Gaussian process (GP) formulation. The GP is used as a prior distribution to describe the possible realization of the scheduling dependent coefficient functions of the estimated model. Then, a posterior distribution of these functions is obtained given the measured data and the mean value of this distribution is used to determine the estimated model. The properties and performance of the proposed method are evaluated using an illustrative example of a chemical process, namely a distillation column, as well as an experimental case study with a three tank system.**

***Index Terms*—Linear parameter-varying (LPV) models; System identification; Bayesian method; Gaussian process; High-purity distillation column; Three tank system.**

I. Introduction

Today, complex industrial processes are requested to operate at a wide range of operational conditions on which their dynamic behavior exhibit significantly nonlinear characteristics. To meet with the performance expectations it is often inevitable to consider their nonlinear behavior for the design and development of appropriate control strategies. However, there are several challenges associated with nonlinear model-based control design for such complex processes including the need for accurate dynamical models to obtain a satisfactory performance and the complexity of the associated control problem to be solved, mostly in real time. The first principle model of a real-world industrial process usually contains a large number of parameters and heavy nonlinear relationships. Often system identification methods are seen as an attractive alternative to reduce the complexity of the models of these systems using measured or even simulated data. Moreover, often the understanding of the internal physical/chemical phenomena is not readily available. Therefore, it is imperative to use system identification methods capable of efficient exploration of an appropriate model structure that can captive these relationships directly from input-output data [1], [2]. To this end, data-driven approaches have been developed to address the above mentioned modeling challenges regarding different approaches such as nonlinear ARX methods [3], Wiener-Hammerstein models [4], neural networks [5], fuzzy systems [6] and least-squares support vector machine (LS-SVM) approaches [7] which have been applied to identify low complexity; yet accurate models.

To take advantage of the simplicity of linear time-invariant (LTI) control synthesis methods and, at the same time, accurately capture the dynamics of process systems over the whole operating regime, *linear parameter-varying* (LPV) identification methods have attracted considerable attention in the last decade (see, e.g., [8]–[13]). Different model structures including input-output [14], state-space [15] and orthonormal basis functions (OBF) representations [16] based approaches have been proposed for identification of LPV models of nonlinear systems. In particular, estimation of LPV models in an input-output (IO) setting has received a great deal of attention [17]–[19]. Most of the methods developed for LPV-IO model identification has been in the discrete-time domain using a linear regression form with static dependency of the model coefficients on the scheduling variables. However, in practice, these coefficient functions are not fully known and often their estimation is required from the measured data. The standard solution to this problem is to parameterize these functions using a combination of *a priori* known basis functions. However, inaccurate selection of the basis functions can lead to structural bias while the over-parameterization can increase the variance of the estimated model. Moreover, LPV modeling of nonlinear systems often requires that the nonlinear coefficients are dependent on the time-shifted versions of the scheduling variables referred to as dynamic dependency. Parameterization with such dynamic dependency often significantly increases the dimension of the parameterization space. Hence, it is appealing to estimate and learn the underlying dependencies of the LPV model coefficients directly from the measured data without a priori selection of a wide range of basis functions to ensure an adequate representation capability of the model structure. Some recent works have proposed the use of non-parametric approaches to obtain an efficient solution to this problem (see, e.g., [11], [20]–[22]).

In [20], a *least-square support vector machine* (LS-SVM) approach is proposed to reconstruct the dependency structure

This publication was made possible by the NPRP grant (No. 5-574-2-233) from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

[1]Department of Electrical Engineering, College of Engineering, Qatar University, P.O. Box 2713, Doha, Qatar (e-mail: arash.golabi@qu.edu.qa, nader.meskin@qu.edu.qa).

[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]Complex Systems Control Lab, College of Engineering, The University of Georgia, Athens, GA 30602, USA (e-mail: javadm@uga.edu).

for linear regression-based LPV models. The SVM approaches represent a class of supervised learning methods for efficient reconstruction of underlying functional relationships and structures in data with the LS-SVM being a particular subclass formulated for regression problems [23]. In the present paper, a different learning approach called the *Bayesian method* is proposed for the structural identification of a linear regression type of LPV-IO models. The difference between SVM approaches and Bayesian methods is that the former view the problem from a deterministic optimization point of view, while the latter start in a stochastic framework with a prior distribution of the parameters/functions to be estimated. Bayesian methods for system identification have been a subject of renewed interest in the last few years (see, e.g., [24]–[28]) due to the efficient characterization of *a priori* information in terms of kernels that encode structural properties of dynamical systems such as stability [29]–[31]. In this context, in [25], the impulse response of a linear system is modeled as the realization of a *Gaussian process* (GP) whose statistics include information not only on smoothness, but also on *bounded-input bounded-output* (BIBO) stability. Unlike classical learning methods, Bayesian approaches do not attempt to identify the best model from data or make the best guess in terms of predictions for new test inputs. Instead, they compute a posterior distribution over models or compute posterior predictive distributions for new test inputs [32]. These distributions offer a useful way to quantify the uncertainties of the estimated model, which can be exploited to achieve more robust predictions on new test points. A preliminary version of this work has appeared in [33] for single-input single-output systems in finite impulse response (FIR) form.

In comparison to Bayesian identification methods in the LTI framework, nonlinear dependency of the LPV model coefficients on the scheduling variables is considered to be the main source of complexity involved in LPV model identification. In fact in order to identify an LPV model of a nonlinear system using Bayesian approaches, the covariance function should be parameterized with a priori information to incorporate possible structural dependencies in the LPV model and asymptotic stability of the predictor. In this paper, we take advantage of the Bayesian method utilizing a Gaussian kernel to reconstruct the dependency structure of an LPV-IO model of a data-generating system. Then, the identified LPV-IO model will be used to predict the future behavior of the system followed by its utilization in controller synthesis in terms of an internal model control approach.

The applicability of the proposed Bayesian-based method for LPV model identification is demonstrated in identifying the LPV model of a laboratory setup of a three tank system and a simulation model of a high-purity distillation column. The three tank system has been extensively used in control literature to investigate linear and nonlinear multivariable feedback control. Next to that, high-purity distillation columns are well known for their nonlinear characteristics and directionality problem making them a changeling problem for LTI control pointing for a nonlinear or LPV solution. The results on these systems show that the proposed LPV identification approach can provide a promising tool to tackle the above stated challenges by offering an accurate model identification method with an attractive computational load.

This paper is organized as follows. In Section II, the problem of LPV model identification based on the Bayesian approach is presented and the proposed method for reconstructing the dependency structure of the estimated model is described. In Section III, properties and performance of the proposed approach are studied through both on a numerical and experimental study in chemical processes. Moreover, a control application of the proposed method provided model is studied. Finally, concluding remarks are made in Section IV.

II. LPV model Identification Based on a Bayesian Approach

In this section, the Bayesian formulation is employed to reconstruct the dependency structure of an LPV data-generating system with an auto-regressive model with exogenous input (ARX) noise structure.

A. The LPV Bayesian estimation problem

For a multi-input multi-output (MIMO) LPV model with $n_{\mathbb{Y}}$ outputs and $n_{\mathbb{U}}$ inputs, the ARX model structure for the *j*-th output is described as

$$y^{(j)}(k) = -\sum_{l=1}^{n_{\mathbb{Y}}}\sum_{i=1}^{n_{\mathrm{a}}} a_{l,i}^{(j)}\left(p(k)\right) y^{(l)}(k-i) + \sum_{s=1}^{n_{\mathbb{U}}}\sum_{i=0}^{n_{\mathrm{b}}} b_{s,i}^{(j)}(p(k)) u^{(s)}(k-i) + e^{(j)}(k), \qquad (1)$$

where $k \in \mathbb{Z}$ is the discrete time, $u^{(s)}(k)$ and $y^{(l)}(k)$ denote the *s*-th input and the *l*-th output signals, respectively, $a_{l,i}^{(j)} : \mathbb{P} \to \mathbb{R}$, $b_{s,i}^{(j)} : \mathbb{P} \to \mathbb{R}$ are model coefficients dependent on $p(k)$ for the j-th output channel, $p : \mathbb{Z} \to \mathbb{P}$ is the so-called scheduling variable with compact range $\mathbb{P} \subseteq \mathbb{R}^{n_p}$, $e^{(j)}(k)$ is an *independent and identically distributed* (i.i.d.) white stochastic noise process with distribution $\mathcal{N}(0, \sigma^2_{e^{(j)}})$ (also being independent of the input signal $u^{(s)}$), n_{a} is the number of time-shifted instances of the past outputs, and n_{b} is the number of time-shifted instances of the past inputs used to predict the current output. Moreover, $\mathbb{R}$ and $\mathbb{Z}$ are the set of real numbers and integer numbers, respectively. The process model (1) is fully characterized by the nonlinear coefficients $\{a_{l,i}^{(j)}\}_{l=1,i=1}^{n_{\mathbb{Y}},n_{\mathrm{a}}}$ and $\{b_{s,i}^{(j)}\}_{s=1,i=0}^{n_{\mathbb{U}},n_{\mathrm{b}}}$ for the j-th channel. The model structure (1) can be also represented as

$$y^{(j)}(k) = \sum_{i=1}^{n_{\mathrm{g}}} g_i^{(j)}(p(k)) x_i(k) + e^{(j)}(k), \qquad (2)$$

where

$$x_{\vartheta_y(l,i)}(k) = y^{(l)}(k-i) \qquad \text{for } i \in \mathbb{I}_1^{n_{\mathrm{a}}}, l \in \mathbb{I}_1^{n_{\mathrm{Y}}}, \qquad (3)$$

$$x_{\vartheta_u(s,i)}(k) = u^{(s)}(k-i) \qquad \text{for } i \in \mathbb{I}_0^{n_{\mathrm{b}}}, s \in \mathbb{I}_1^{n_{\mathrm{U}}}, \qquad (4)$$

with $\mathbb{I}_{\nu_1}^{\nu_2} := \{i \in \mathbb{Z} \mid \nu_1 \le i \le \nu_2\}$ being an index set, $\vartheta_y(l,i) = n_{\mathrm{a}}(l-1)+i, \vartheta_u(s,i) =$

$n_{\mathbb{Y}}n_{\mathrm{a}} + (s-1)(n_{\mathrm{b}}+1) + 1 + i$, $n_{\mathrm{g}} = n_{\mathbb{Y}}n_{\mathrm{a}} + n_{\mathbb{U}}(n_{\mathrm{b}}+1)$, and

$$[g_1^{(j)} \quad \dots \quad g_{n_{\mathrm{g}}}^{(j)}] = [a_{1,1}^{(j)} \dots a_{1,n_{\mathrm{a}}}^{(j)} \dots a_{n_{\mathbb{Y}},1}^{(j)} \dots a_{n_{\mathbb{Y}},n_{\mathrm{a}}}^{(j)} \quad b_{1,0}^{(j)} \dots b_{1,n_{\mathrm{b}}}^{(j)} \dots b_{n_{\mathbb{U}},0}^{(j)} \dots b_{n_{\mathbb{U}},n_{\mathrm{b}}}^{(j)}]. \tag{5}$$

In the context of the Bayesian setting for coefficient estimation, a Gaussian process (GP) is used to describe a distribution of nonlinear functions representing the coefficients of an LPV model and the posterior distributions of these functions are obtained, given the available observations. Formally, a GP is a stochastic process such that any finite subcollection of random variables has a multivariate Gaussian distribution [32]. Consider a standard scalar Gaussian process regression model [32]

$$\mathcal{Y}(k) = \mathcal{F}(\phi(k)) + e(k), \tag{6}$$

where $\phi(k)$ is the "input vector", $\mathcal{Y}(k)$ is the observed target value and $e(k)$ is an i.i.d. noise process with a zero mean and variance of σ_e^2, i.e., $e(k) \sim \mathcal{N}(0, \sigma_e^2)$. It is also assumed that the function $\mathcal{F}(.)$ is a particular realization of a multivariate process with zero-mean Gaussian distribution as

$$\mathcal{F}(.) \sim \mathcal{GP}(0, \mathcal{K}(.,.)), \tag{7}$$

with a symmetric positive semidefinite covariance function $\mathcal{K}(.,.)$ where $\mathcal{GP}(.,.)$ denotes a Gaussian process distribution. The following result on the joint distribution of the observed target value and the function value $\mathcal{F}^*$ at a test point ϕ^*, i.e., $\mathcal{F}^* = \mathcal{F}(\phi^*)$, is used in Bayesian calculations

$$\begin{bmatrix} \mathcal{Y} \\ \mathcal{F}^* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathcal{K}(\Phi,\Phi) + \sigma_e^2 I & \mathcal{K}(\Phi,\phi^*) \\ \mathcal{K}(\phi^*,\Phi) & \mathcal{K}(\phi^*,\phi^*) \end{bmatrix}\right), \tag{8}$$

where Φ is a vector of "input training data" and $\mathcal{Y}$ is a vector of observed target values. Note that if there are n training points and one test point ϕ^*, then $\mathcal{K}(\Phi,\phi^*)$ is an $n \times 1$ vector. The posterior distribution for Gaussian process regression can then be obtained as [32]

$$\mathbb{E}\{\mathcal{F}^*\} = \bar{\mathcal{F}}^* = \kappa^{*\top}(\mathcal{K}(\Phi,\Phi) + \sigma_e^2 I)^{-1}\mathcal{Y}, \tag{9}$$

$$\mathrm{cov}\{\mathcal{F}^*\} = \mathcal{K}(\phi^*,\phi^*) - \kappa^{*\top}(\mathcal{K}(\Phi,\Phi) + \sigma^2 I)^{-1}\kappa^*, \tag{10}$$

where $\kappa^* = \mathcal{K}(\phi^*,\Phi)$ and $\mathbb{E}$ denotes the expectation operator.

Next, the above Bayesian formulation is used for the identification of the LPV IO model (1). For each $g_i^{(j)}(.)$ in Equation (2), the covariance function for all $p(s), p(k) \in \mathbb{P}$ is defined as

$$\begin{aligned} \mathrm{cov}\{g_i^{(j)}(p(k)), g_i^{(j)}(p(s))\} &= \mathbb{E}\{g_i^{(j)}(p(k)) g_i^{(j)}(p(s))\} \\ &= \mathcal{K}_i^{(j)}(p(k), p(s)), \end{aligned} \tag{11}$$

where $\mathcal{K}_i^{(j)}(.,.)$ is a positive kernel function. The Gaussian kernel commonly used in non-parametric estimation is defined as [34]

$$\mathcal{K}_i^{(j)}(p(k), p(s)) = \lambda_i^{(j)} \exp\left(-\frac{||p(k) - p(s)||_2^2}{(\sigma_i^{(j)})^2}\right), \tag{12}$$

which is also referred to as *Radial Basis Function* (RBF). In (12), $\sigma_i^{(j)}$ and $\lambda_i^{(j)}$ are the hyper-parameters of the prior distribution which specify the width and scale of the RBF. We note that the kernel function in (12) is expressed as a function of the scheduling variable p and describes the prior assumption on how observed values of $g_i^{(j)}$ relate to each other depending on the distance between $p(k)$ and $p(s)$. It is assumed that a training data set are given as $\mathcal{D} = \{y(k), u(k), p(k)\}_{k=1}^N$. The joint distribution of the observed value and coefficient function values $g_i^{(j)}(.)$ of the LPV model structure are given by

$$\begin{bmatrix} y^{(j)} \\ g_i^{(j)}(.) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K_{y^{(j)}} + \sigma_{e^{(j)}}^2 I & \kappa_i^{*(j)}(.) \\ \kappa_i^{*(j)\top}(.) & \mathcal{K}_i^{(j)}(.,.) \end{bmatrix}\right), \tag{13}$$

where (.) denotes any value of the scheduling parameter and

$$\kappa_i^{*(j)}(.) = \begin{bmatrix} x_i(1)\mathcal{K}_i^{(j)}(p(1),.) \\ \vdots \\ x_i(N)\mathcal{K}_i^{(j)}(p(N),.) \end{bmatrix}, \tag{14}$$

$$[K_{y^{(j)}}]_{s,k} = \sum_{i=1}^{n_{\mathrm{g}}} [K_{y^{(j)}}]_{s,k}^i, \tag{15}$$

$$[K_{y^{(j)}}]_{s,k}^i = x_i(s)\left(\mathcal{K}_i^{(j)}(p(s),p(k))\right)x_i(k). \tag{16}$$

The property of linear combinations of mutually independent normal random variables [35] is employed here to compute *covariance* and *mean* of $y^{(j)}(.)$ based on (2) as

$$\begin{aligned} &\mathrm{cov}\{y^{(j)}(k), y^{(j)}(s)\} = \\ &\sum_{i=1}^{n_{\mathrm{g}}} x_i(k)\mathrm{cov}\{g_i^{(j)}(p(k)), g_i^{(j)}(p(s))\}x_i(s) + \sigma_{e^{(j)}}^2, \\ &\qquad = \sum_{i=1}^{n_{\mathrm{g}}} x_i(k)\mathcal{K}_i^{(j)}(p(k),p(s))\, x_i(s) + \sigma_{e^{(j)}}^2, \\ &\qquad = \sum_{i=1}^{n_{\mathrm{g}}} [K_{y^{(j)}}]_{s,k}^i + \sigma_{e^{(j)}}^2. \end{aligned} \tag{17}$$

From (9) and (10), the posterior distribution of $g_i^{(j)}(.)$, i.e., the LPV model coefficient functions, based on the prior distribution (13), are chosen to be estimated as the following conditional mean

$$\mathbb{E}\left\{g_i^{(j)}(.) \mid \mathcal{D}\right\} = \bar{g_i}^{(j)}(.) = \sum_{k=1}^{N} \alpha_k^{(j)} x_i(k)\mathcal{K}_i^{(j)}(p(k),.), \tag{18}$$

while their conditional covariance is characterized by

$$\mathrm{cov}\left\{g_i^{(j)}(.) \mid \mathcal{D}\right\} = \mathcal{K}_i^{(j)}(.,.) - \sum_{k=1}^{N} \beta_{i,k}^{(j)}(.) x_i(k)\mathcal{K}_i^{(j)}(p(k),.), \tag{19}$$

in which $\alpha_k^{(j)}$ and $\beta_{i,k}^{(j)}$ are the k-th elements of N-dimensional vectors $\alpha^{(j)}$ and $\beta_i^{(j)}$, respectively, defined by

$$\alpha^{(j)} = \left(K_{y^{(j)}} + (\sigma_e^{(j)})^2 I\right)^{-1} Y^{(j)}, \tag{20}$$

with

$$Y^{(j)} = \begin{bmatrix} y^{(j)}(1) & \dots & y^{(j)}(N) \end{bmatrix}^\top,$$

and

$$\beta_{i,k}^{(j)}(.) = (K_{y^{(j)}} + (\sigma_e^{(j)})^2 I)^{-1} \kappa_i^{*(j)}(.),$$

where $\kappa_i^{*(j)}(.)$ is defined in (14).

It should be noted that, in principle, the LS-SVM approach [20] and the proposed Bayesian method are similar. The main difference is the availability of the estimation of the *covariance* function (19) based on which the Bayesian method computes a posterior distribution over the model set. Computation of the covariance function provides a useful way to quantify the uncertainties of the estimated LPV model of the underlying system.

Remark 1: Although, the assumption of an ARX model structure with white noise equation error has been commonly made in system identification with Bayesian methods (see, e.g., [24]–[27], [31], [34]), it may constrain the effectiveness and applicability of the method. As an extension of this work, instrumental variable (IV) approaches can be developed for the Bayesian approach to handle more realistic noise scenarios without specifying a direct noise parametrization or structure [36].

B. Estimation of the hyper-parameters

To estimate the parameters $\lambda_i^{(j)}$ and $\sigma_i^{(j)}$ in the kernel function (12), an exhaustive search can be performed over a discrete grid of values, using an independent validation data set and by minimizing the ℓ_2-loss over the prediction error, but this can be quite slow and time consuming [35]. As an alternative, the empirical Bayesian approach can be used, which offers a direct estimation of these parameters via maximizing the marginalized likelihood of the model estimate. Assume that $\theta_i^{(j)}$ denotes the vector whose components are the hyper-parameters describing the kernel function (12), i.e., $\theta_i^{(j)} = [\lambda_i^{(j)} \sigma_i^{(j)}]^\top$. Based on (17) and the Gaussian assumption on the distribution of $g_i^{(j)}$ and $e^{(j)}$, the logarithm of the marginal likelihood for the j-th output channel is given by

$$\begin{aligned}\log \mathscr{P}\left(y^{(j)}|\mathcal{D},\theta^{(j)}\right) &= \log \mathcal{N}\left(y^{(j)}\middle|0,\sum_{i=1}^{n_\mathrm{g}}\left([K_{y^{(j)}}]^i\right)\right)\\ &= -\frac{1}{2}y^{(j)\top}\sum_{i=1}^{n_\mathrm{g}}\left([K_{y^{(j)}}]^i\right)^{-1}y^{(j)} - \\ &\quad \frac{1}{2}\log\left|\sum_{i=1}^{n_\mathrm{g}}\left([K_{y^{(j)}}]^i\right)\right| - \frac{N}{2}\log(2\pi), \quad (21)\end{aligned}$$

where $\theta^{(j)} = [\theta_1^{(j)} \ ... \ \theta_{n_\mathrm{g}}^{(j)}]$, $\mathscr{P}$ denotes the probability density function and $[K_{y^{(j)}}]^i$ is the matrix, whose elements are determined by (16). Therefore, the hyper-parameters $\theta_i^{(j)}$ can be estimated to maximize the marginal likelihood function in (21) as

$$\hat{\theta}^{(j)} = \arg\max_{\theta_1^{(j)},...,\theta_{n_\mathrm{g}}^{(j)}} J(\theta_1^{(j)},...,\theta_{n_\mathrm{g}}^{(j)}), \quad (22)$$

and

$$J(\theta^{(j)}) := \log \mathscr{P}(y^{(j)}|\mathcal{D},\theta_1^{(j)},...,\theta_{n_\mathrm{g}}^{(j)}). \quad (23)$$

Then, the estimated hyper-parameters $\hat{\theta}_i^{(j)}$ are substituted in (12) to obtain the estimates $g_i^{(j)}(.)$ in (18).

C. Summary of the proposed Bayesian method

Based on the data set $\mathcal{D} = \{y(k), u(k), p(k)\}_{k=1}^N$, the proposed LPV identification algorithm is summarized below.

1) Determine the order of the LPV model, i.e., the value of n_a and n_b in (2) by using cross-validation [35] based order selection.
2) Estimate the hyper-parameters in (12) based on (22) and $\mathcal{D}$ to obtain $\hat{\theta}_i^{(j)}$.
3) Obtain the function estimate $g_i^{(j)}(.)$ based on (18) using $\mathcal{D}$.

III. Simulation and Experimental Studies

To illustrate the performance of the proposed LPV model identification method, we use the model of a high-purity distillation column and a laboratory three tank system for which the Bayesian method is utilized based on the proposed algorithm in the Section II-C to estimate a non-parametric LPV-IO model. Then, using a validation data set $\mathcal{D}_{N^*}^{\mathrm{val}} = \{(u^*(k), p^*(k), y^*(k))\}_{k=1}^{N^*}$ compute the one-step-ahead prediction of the j-th output channel at the time instant k as follows

$$\begin{aligned}\hat{y}^{(j)}(k) = &\sum_{i=1}^{n_\mathrm{a}} \bar{g}_i^{(j)}(p^*(k))\hat{x}_i(k) + \\ &\sum_{s=1}^{n_\mathrm{U}}\sum_{i=1}^{n_\mathrm{b}+1} \bar{g}_{\vartheta(s,i)}^{(j)}(p^*(k))x_{\vartheta(s,i)}^*(k),\end{aligned}$$

where

$$\begin{aligned}\hat{x}_{\vartheta_y(l,i)}(k) &= \hat{y}^{(l)}(k-i) \ \text{ for } i \in \mathbb{I}_1^{n_\mathrm{a}}, l \in \mathbb{I}_1^{n_\mathrm{Y}},\\ x_{\vartheta(s,i)}^*(k) &= u^{*(s)}(k-i+1) \ \text{ for } i \in \mathbb{I}_1^{n_\mathrm{b}+1}, s \in \mathbb{I}_1^{n_\mathrm{U}},\end{aligned}$$

in which $u^{*(s)}$ is the s-th input in the validation data set and $\vartheta(s,i) = (s-1)(n_b+1)+i$.

Moreover, a control application of the developed LPV-IO model is studied through the liquid level control in a single tank system.

A. LPV modeling of a high-purity distillation column

The high-purity distillation column considered in this paper is a *propane-propene splitter* (PP-splitter). Based on the principle of boiling point difference of propane and propene, the PP-splitter is designed to separate the mixture of these substances to its components with a desired purity level. The high-purity products (distillates) are valuable for the gasoline production or as a raw material for further chemical synthesis [37]. The schematic of a PP-splitter is depicted in Figure 1. High-purity distillation columns are well known for their nonlinear dynamics and directionality problem. The response of these systems is dominated by the high-gain direction which limits the performance of linear SISO controllers when controling both top and bottom compositions. To tackle this issue, a MIMO controller is often used which requires an accurate

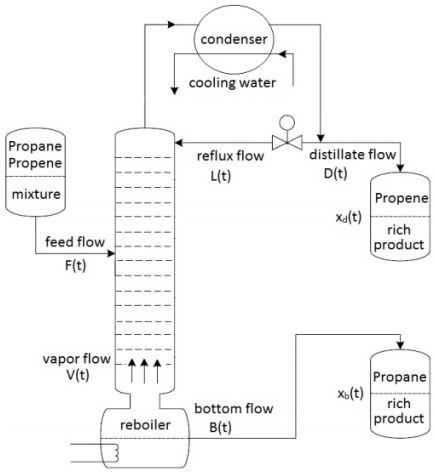

Fig. 1: Schematic of a typical PP-splitter [38].

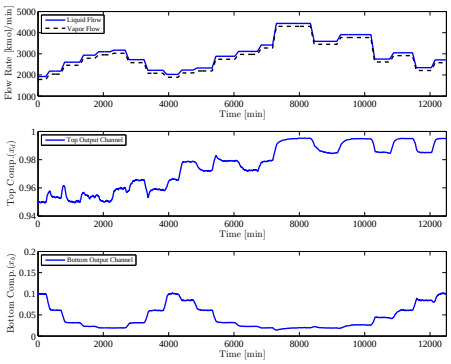

Fig. 2: Measurement data used for LPV model identification of the distillation column.

model in order to compute the correct control action. However, due to the strong nonlinear behavior of high-purity distillation columns, a single LTI model may be insufficient to describe the system dynamics. Hence, an alternative modeling solution is required which can preserve the simplicity of LTI control synthesis while accurately capturing the system dynamics over the whole operating regime. Toward this goal, the proposed Bayesian identification method is utilized to obtain an LPV-IO model of the high-purity distillation column.

To generate realistic measurement records of the system, a first-principle model of a typical PP-splitter is used [37]. In the process, the liquid flow rate L and the vapor flow rate V are used as the manipulated variables to control the operation of the high-purity column. The system outputs consist of the top composition x_d (Propene) and bottom x_b (Propane) products. The resulting distillation column model description is a large-scale (110^{th} order), nonlinear, 2-input 2-output system. The bottom and top product compositions are chosen as scheduling variables, i.e., $p(k) = [x_b(k)\ x_d(k)]^\top$ since they uniquely characterize the operating point of the system and are available via measurements. The scheduling range is selected such that it covers a large set of local operating points described by the top and bottom compositions in the region $\mathbb{P} = [0.95, 0.995] \times [0.02, 0.1]$. The first principle model is simulated in continuous time and input/output data is collected with the sampling rate of $T_s = 5$ min, which is 20 times faster than the time constant of the fastest step response with respect to all the possible operating conditions. The inputs of the system are manipulated through a zero-order-hold actuation synchronized with the data sampling. Moreover, the measurement data is assumed to be corrupted by an output additive zero-mean white noise process with *signal-to-noise ratio* (SNR) of 15dB.

To identify the LPV model, we use the data set $\mathcal{D} = \{(u(k), y(k), p(k))\}_{k=1}^{N}$ with $N = 2500$, which corresponds to 8.7 days of experimentation time. The data set is generated using an input signal which is a combination of a deterministic component added to a white noise with uniform distribution $\mathcal{U}(-0.5, 0.5)$ corresponding to a *standard deviation* (std) of $1/\sqrt{12}$ (see Figure 2). Cross-validation based order selection has been used to choose the structure of the LPV model leading to $n_\text{a} = 3$ and $n_\text{b} = 3$. The hyper-parameters have been obtained as $\lambda_i^{(1)} = 0.01$, $\lambda_i^{(2)} = 0.0095$, $\sigma_i^{(1)} = 0.16$ and $\sigma_i^{(2)} = 0.75$ for $i \in \mathbb{I}_1^{n_\text{g}}$ based on the algorithm proposed in the Section II-B. Figure 3 shows the true system output, the LPV model output and the error between the true output and the simulated output of the identified LPV model w.r.t. $\mathcal{D}_{30\text{k}}^{\text{val}}$, which is the validation data set selected to be different from the training data set. To show the dependency, Figure 4 depicts four coefficients ($a_{1,1}^{(1)}, b_{1,0}^{(1)}, a_{1,1}^{(2)}$ and $b_{1,0}^{(2)}$ in (2)) of the identified LPV model of the distillation column as a function of scheduling variables. The *mean squared error* (MSE) and the *best fit rate* (BFR) have been used to evaluate the performance of the Bayesian approach as

$$\text{MSE} = ||y(k) - \hat{y}(k)||_{\ell_2}^2, \tag{24}$$

$$\text{BFR} = 100\% \times \max\left(1 - \frac{||y(k) - \hat{y}(k)||_{\ell_2}}{||y(k) - \bar{y}||_{\ell_2}}\right), \tag{25}$$

where $\bar{y}$ denotes the mean of y and $\hat{y}$ is the simulated model response for the input in $\mathcal{D}_{30\text{k}}^{\text{val}}$.

The MSE for the estimation error of y_1 is calculated to be 1.76×10^{-7} while that of y_2 is 5.27×10^{-9}. The BFR for the estimation of y_1 is 99.84% while that of y_2 is 99.76%. Based on the obtained results, it can be concluded that the Bayesian method has led to a highly accurate LPV model for high-purity distillation column with an outstanding BFR and MSE.

It is important to note that in this paper, the LPV-ARX model structure is used for a system with an *output additive* (OE) noise structure which results in a biased estimation. However, the optimal tuning of the hyper-parameters can partly compensate for the bias. The proposed approach can be refined by using the instrumental-variable approach introduced in [36]. The comparison of the mean values of the MSE and BFR of the LPV Bayesian approach with the available global LPV model identification approaches is shown in Table I with the results for the global approaches borrowed from [37]. In comparison with the LPV LS-SVM and the parametric global

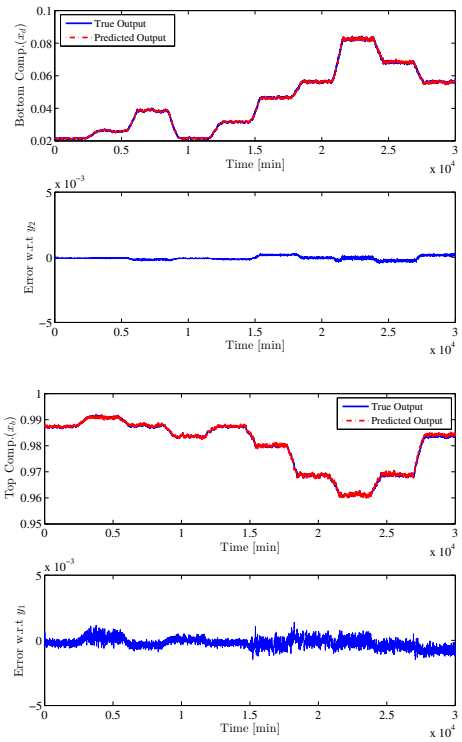

Fig. 3: Comparison of the true system output x_b, x_d and the output of the estimated LPV distillation column model based on the validation data set.

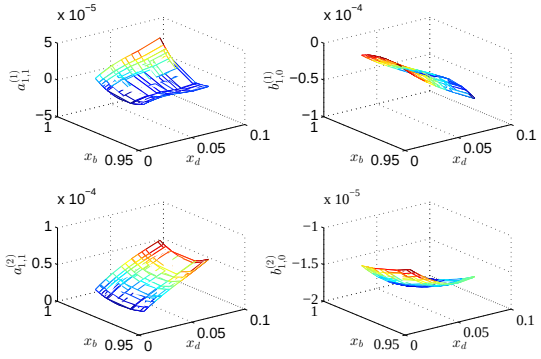

Fig. 4: Four coefficient functions of the identified LPV model of the distillation column.

LPV-OBF and OE methods, the means for both BFR and MSE are slightly worse with the proposed LPV Bayesian approach. Bayesian approaches do not attempt to identify the best model from data or result in the best guess in terms of predictions for the new test inputs. However, they are able to compute a posterior distribution over models and posterior predictive distributions for the new test inputs. These distributions can be used to quantify the uncertainties of the estimated model, which can be exploited to achieve more robust predictions on new test points. For example, in Figure 5 the uncertainty can be described as 95% confidence region for coefficient functions $a_{1,1}^{(1)}$ and $a_{1,1}^{(2)}$. Here, the confidence region is defined based on the covariance function in (19). It should be emphasized

TABLE I: The comparison of mean of BFR and MSE for the LPV models estimated via different approaches.

	LPV Bayesian	LPV LS-SVM	LPV OE	Global LPV-OBF
Avg. BFR of y_1	96.84%	97.02%	92.49%	97.68%
Avg. MSE of y_1	1.76×10^{-7}	8.27×10^{-8}	5.03×10^{-7}	5.14×10^{-8}
Avg. BFR of y_2	99.76%	99.89%	99.77%	99.90%
Avg. MSE of y_2	5.27×10^{-9}	5.14×10^{-10}	2.02×10^{-9}	3.79×10^{-10}

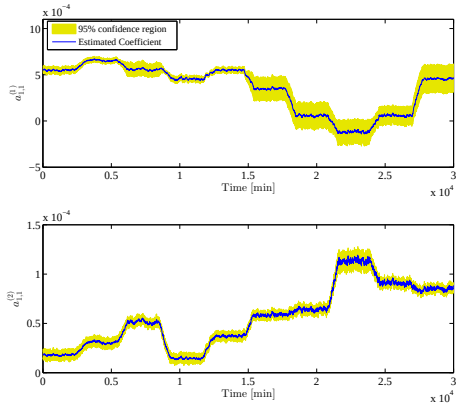

Fig. 5: Mean and 95% confidence region of the estimated coefficient functions $a_{1,1}^{(1)}$ and $a_{1,1}^{(2)}$ evaluated along a given scheduling trajectory.

that none of the previously reported identification methods for LPV systems can provide such analytical tool to quantify the uncertainty in the model coefficients. Moreover, the OBF model requires a very high-order polynomial dependency for the associated coefficient functions; hence, this makes the LPV Bayesian approach truly promising from the computational point of view.

B. LPV modeling of a three tank system

In this section, the proposed LPV model identification approach is examined using an experimental three tank liquid level system. The three tank system has been widely used as a benchmark to investigate linear and nonlinear multivariable feedback control. As shown in Figure 6, the system under consideration consists of three plexiglas cylinders T1, T3 and T2. The pump flow rates Q_1 and Q_2 denote the input signals. The liquid levels of h_1 in cylinder T1 and h_2 in cylinder T2 denote the output signals. Therefore, the model to be identified is a 2-input, 2-output LPV model of three-tank system. The liquid levels of h_1 and h_2 (in cm) are chosen as the scheduling variables, i.e., $p(k) = \begin{bmatrix} h_1(k) & h_2(k) \end{bmatrix}^\top$ since they uniquely characterize the operating point of the system. The scheduling range is selected such that it contains a large set of local operating points described by the system outputs in the region $\mathbb{P} = [0, 40] \times [0, 40]$. The system is excited with a random multilevel signal and the resulting output signals are depicted in Figure 7. A training data set $\mathcal{D}$ with $N = 450$ samples is collected with a sampling rate of 1 sec from the three tank system. Cross-validation based order selection has been used to choose the structure of LPV model with $n_a = 3$ and $n_b = 3$. The hyper-parameters have been obtained as $\lambda_i^{(1)} = 0.025$,

Fig. 6: The three tank system TTS20.

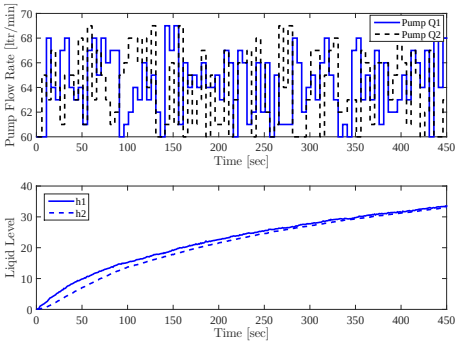

Fig. 7: Measurement data used for LPV model identification of the three tank system.

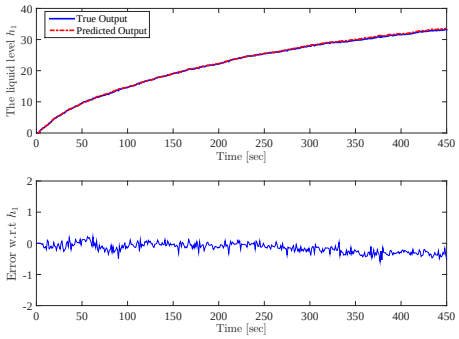

Fig. 8: The liquid level in the first cylinder: the real output versus the predicted one.

$\lambda_i^{(2)} = 0.02$, $\sigma_i^{(1)} = 40.1$, and $\sigma_i^{(2)} = 52.2$ for $i \in \mathbb{I}_1^{n_g}$ based on the algorithm proposed in Section II-B. Figures 8 and 9 show the real system output of the three tank system and the simulated output of the estimated LPV model and the corresponding output error w.r.t. $\mathcal{D}_{900}^{\text{val}}$, which is the validation data set selected to be different from the training data set. The MSE for the estimation error of h_1 is calculated to be 0.01 and h_2 is calculated to be 0.05. The BFR for the estimation of h_1 is 99.89% and of h_2 is 99.91%. Based on the obtained results, it can be concluded that the Bayesian method has led to a highly accurate LPV model for the three tank system with a high BFR.

C. Control application of the proposed LPV modeling approach

In this section, a control application of the proposed LPV model identification approach is given. To this aim, first an LPV model of a single tank system is identified using the proposed algorithm, and then, an internal model control (IMC) feedback structure is employed such that the tank liquid level tracks the step reference signal $r(t)$ in the presence of disturbances. In this framework, a model of the process is used internally to simulate the response of the system. In the IMC structure, the internal model is the identified LPV-IO model and the controller is an LTI one. The configuration of the IMC structure is shown in Figure 10. The LPV-IO model obtained via the proposed algorithm is used to simulate the output of the single tank system. A band-limited white noise with the power of 0.0001 is chosen to model the disturbance input $d(s)$. The controller transfer function $C(s)$ is obtained as $C(s) = \frac{4.02(13.8s+1)}{4s+1}$. In Figure 11, the tank liquid level and its corresponding reference signal are shown under the IMC feedback structure. The presented IMC controller clearly demonstrates the usefulness of the obtained LPV-IO model for control design purposes. It is noted that a parameter-dependent controller can also be designed for achieving an improved performance, which is beyond the scope of this paper.

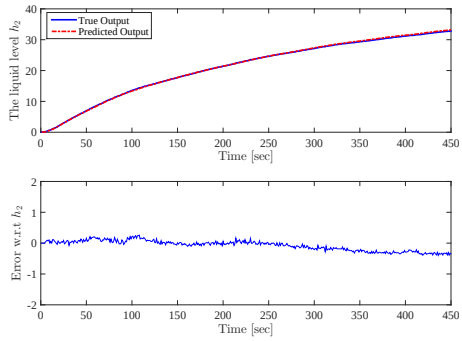

Fig. 9: The liquid level in the second cylinder: the real output versus the predicted one.

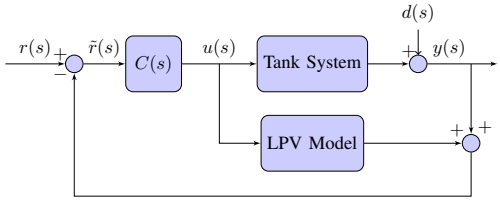

Fig. 10: Internal model control feedback structure.

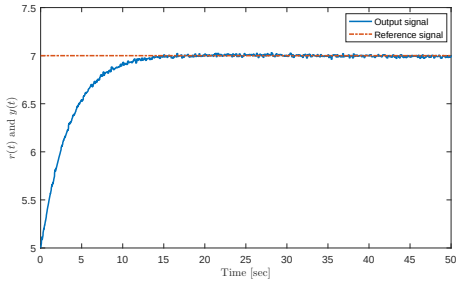

Fig. 11: The output signal $y(t)$ and the reference signal $r(t)$.

IV. Conclusion

In this paper, a Bayesian formulation for identification of LPV-IO models has been introduced. A Gaussian process (GP) is used to describe the distribution over nonlinear coefficient functions. Without a prior information about the parametrization of the underlying coefficient functions, the proposed approach is capable of reconstructing the dependency structure of the LPV model based on the posterior distribution. The applicability of the proposed approach has been investigated through an extensive simulation study using a complex chemical process, namely a high-purity distillation column and an experimental case study using a three tank system laboratory setup. By comparing the output of the identified LPV models and the original systems, it is shown that the proposed LPV Bayesian approach achieves high accuracy in capturing the input-output behavior of such nonlinear systems.

References

[1] R. K. Pearson, "Nonlinear input/output modeling," *Journal of Process Control*, vol. 5, no. 4, pp. 197 – 211, 1995.

[2] Y. Lu and B. Huang, "Robust multiple-model LPV approach to nonlinear process identification using mixture t-distributions," *Journal of Process Control*, vol. 24, no. 9, pp. 1472 – 1488, 2014.

[3] S. Chen, X. Wang, and C. Harris, "NARX-based nonlinear system identification using orthogonal least squares basis hunting," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 1, pp. 78–84, Jan 2008.

[4] G. van der Veen, J.-W. van Wingerden, and M. Verhaegen, "Global identification of wind turbines using a Hammerstein identification method," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 4, pp. 1471–1478, July 2013.

[5] S.-I. Moon, A. Keyhani, and S. Pillutla, "Nonlinear neural-network modeling of an induction machine," *IEEE Transactions on Control Systems Technology*, vol. 7, no. 2, pp. 203–211, Mar 1999.

[6] W. Yu, "A novel fuzzy-neural-network modeling approach to crude-oil blending," *IEEE Transactions on Control Systems Technology*, vol. 17, no. 6, pp. 1424–1431, Nov 2009.

[7] M. Al-Dhaifllah and D. Westwick, "Identification of auto-regressive exogenous Hammerstein models based on support vector machine regression," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 6, pp. 2083–2090, Nov 2013.

[8] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, ser. Lecture Notes in Control and Information Sciences. Berlin, Germany: Springer, 2010, vol. 403.

[9] P. Gebraad, J. van Wingerden, P. Fleming, and A. Wright, "LPV identification of wind turbine rotor vibrational dynamics using periodic disturbance basis functions," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 4, pp. 1183–1190, July 2013.

[10] W. Larimore and H. Javaherian, "Identification of linear parameter-varying engine models," in *American Control Conference (ACC), 2013*, June 2013, pp. 2274–2279.

[11] K. Hsu, T. L. Vincent, and K. Poola, "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, 2008, pp. 846–851.

[12] S. Wattamwar, S. Weiland, and T. Backx, "Identification of low-order parameter-varying models for large-scale systems," *Journal of Process Control*, vol. 20, no. 1, pp. 158–172, 2010.

[13] R. Tóth, P. S. Heuberger, and P. M. J. Van Den Hof, "Order and structural dependence selection of LPV-ARX models using a nonnegative Garrote approach," in *Proc. of the 48th IEEE Conf. on Decision and Control*, Shanghai, China, 2009, pp. 7406–7411.

[14] B. Bamieh and L. Giarre, "Identification of linear parameter varying models," *Int. Journal of Robust and Nonlinear Control*, vol. 12, no. 1, pp. 841–853, Apr. 2002.

[15] J. W. van Wingerden and M. Verhaegen, "Subspace identification of Bilinear and LPV systems for open- and closed-loop data," *Automatica*, vol. 45, pp. 372–381, 2009.

[16] 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.

[17] X. Wei and L. D. Re, "On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque," *IFAC Proceedings Volumes*, vol. 39, no. 1, pp. 517 – 522, 2006.

[18] R. Tóth, V. Laurain, M. Gilson, and H. Garnier, "Instrumental variable scheme for closed-loop LPV model identification," *Automatica*, vol. 48, no. 9, pp. 2314 – 2320, 2012.

[19] S. Boonto and H. Werner, "Closed-loop system identification of LPV input-output models - application to an arm-driven inverted pendulum," in *Proc. of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec 2008, pp. 2606–2611.

[20] 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 Conference on Decision and Control and European Control Conference*, Orlando, Florida, USA, Dec 2011, pp. 3192–3197.

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

[22] D. Piga and R. Toth, "LPV model order selection in an LS-SVM setting," in *Proc. of the 52nd IEEE Annual Conference on Decision and Control*, Firenze, Italy, Dec 2013, pp. 4128–4133.

[23] B. Scholkopf and A. Smola, *Learning with kernels*, ser. Lecture Notes in Control and Information Sciences. Berlin, Germany: Cambridge MA:MIT press, 1998.

[24] A. Chiuso and G. Pillonetto, "A Bayesian approach to sparse dynamic network identification," *Automatica*, vol. 48, no. 8, pp. 1553 – 1565, 2012.

[25] G. Pillonetto and G. D. Nicolao, "A new kernel-based approach for linear system identification," *Automatica*, vol. 46, no. 1, pp. 81 – 93, 2010.

[26] T. Chen, H. Ohlsson, and L. Ljung, "On the estimation of transfer functions, regularizations and Gaussian processes-revisited," *Automatica*, vol. 48, no. 8, pp. 1525 – 1535, 2012.

[27] V. Peterka, "Bayesian approach to system identification," in *Trends and Progress in System Identification*. Pergamon Press, 1981, pp. 239–304.

[28] A. Chiuso, G. Pillonetto, and G. De Nicolao, "Subspace identification using predictor estimation via Gaussian regression," in *Proc. of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec 2008, pp. 3299–3304.

[29] G. Pillonetto and A. Chiuso, "Tuning complexity in kernel-based linear system identification: The robustness of the marginal likelihood estimator," in *Proc. of the European Control Conference*, Strasbourg, France, June 2014, pp. 2386–2391.

[30] A. Chiuso and G. Pillonetto, "Bayesian and nonparametric methods for system identification and model selection," in *Proc. of the European Control Conference*, Strasbourg, France, June 2014, pp. 2376–2381.

[31] G. Pillonetto, F. Dinuzzo, T. Chen, G. D. Nicolao, and L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey," *Automatica*, vol. 50, no. 3, pp. 657 – 682, 2014.

[32] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. London, England: Massachusetts Institute of Technology, 2006, vol. 403.

[33] A. Golabi, N. Meskin, R. Tóth, and J. Mohammadpour, "A Bayesian approach for estimation of linear-regression LPV models," in *Proc. of the 2014 IEEE 53rd Annual Conference on Decision and Control*, Los Angeles, California, USA, Dec 2014, pp. 2555–2560.

[34] G. Pillonetto, M. H. Quang, and A. Chiuso, "A new kernel-based approach for nonlinear system identification," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2825–2840, December 2011.

[35] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, ser. Adaptive computation and machine learning series. United states of America: Cambridge MA:MIT press, 2012.

[36] V. Laurain, R. Tóth, D. Piga, and W. X. Zheng, "An instrumental least squares support vector machine for nonlinear system identification," *Automatica*, vol. 54, pp. 340 – 347, 2015.

[37] A. Bachnas, R. Tóth, J. Ludlage, and A. Mesbah, "A review on data-driven linear parameter-varying modeling approaches: A high-purity distillation column case study," *Journal of Process Control*, vol. 24, no. 4, pp. 272 – 285, 2014.

[38] A. Bachnas, R. Tóth, A. Mesbah, and J. Ludlage, "Perspectives of data-driven LPV modeling of high-purity distillation columns," in *Proc. of the European Control Conference*, Zürich, Switzerland, July 2013, pp. 3776–3783.