Journal of Process Control 24 (2014) 272–285

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Review

A review on data-driven linear parameter-varying modeling approaches: A high-purity distillation column case study☆

A.A. Bachnas[a], R. Tóth[a,*], J.H.A. Ludlage[a], A. Mesbah[b]

[a] Control Systems Group, Electrical Engineering Department, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[b] Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 22 May 2013
Received in revised form
24 December 2013
Accepted 27 January 2014
Available online 12 March 2014

Keywords:
System identification
Linear parameter-varying systems
Support vector machine
Orthonormal basis functions
Prediction error minimization
High-purity distillation column

ABSTRACT

Model-based control strategies are widely used for optimal operation of chemical processes to respond to the increasing performance demands in the chemical industry. Yet, obtaining accurate models to describe the inherently nonlinear, time-varying dynamics of chemical processes remains a challenge in most model-based control applications. This paper reviews data-driven, Linear Parameter-Varying (LPV) modeling approaches for process systems by exploring and comparing various identification methods on a high-purity distillation column case study. Several LPV identification methods that utilize input–output and series expansion model structures are explored. Two LPV identification perspectives are adopted: (i) the local approach, which corresponds to the interpolation of Linear Time-Invariant (LTI) models identified at different steady-state operating points of the system and (ii) the global approach, where a parametrized LPV model structure is identified directly using a global data set with varying operating points. For the local approach, various model interpolation schemes are studied under an Output Error (OE) noise setting, whereas in the global case, a polynomial parametrization based OE prediction error minimization approach, an Orthonormal Basis Functions (OBFs) based model estimator and a Least-Square Support Vector Machine (LS-SVM) based non-parametric approach are investigated. Through extensive simulation studies, the aforementioned LPV identification approaches are analyzed in terms of the attainable model accuracy and local frequency response behavior of the obtained models. Recommendations are provided to achieve adequate choice between the methods for a particular process system at hand.

Contents

☆ This work was supported by the European Union's 7th Framework Program (FP7/2007-2013) under the grant Autoprofit (Grant No.: 257059); (www.fp7-autoprofit.eu).
* Corresponding author. Tel.: +31 638312635.
E-mail addresses: a.a.bachnas@tue.nl (A.A. Bachnas), r.toth@tue.nl, tothrola@ieee.org (R. Tóth), j.ludlage@tue.nl (J.H.A. Ludlage), amesbah@mit.edu (A. Mesbah).

0959-1524/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jprocont.2014.01.015

1. Introduction

Chemical processes often exhibit a significant nonlinear behavior when operated over a wide range of operating conditions. Despite the advances in model-based control technologies, it remains a challenge to realize high-performance operation of nonlinear chemical processes (in terms of product quality and process productivity) using a single *linear time-invariant* (LTI) model-based controller. For example, when a chemical process is operated under transient conditions (e.g., set-point changes and start-up procedures), the nonlinear behavior of the process becomes more dominant, which makes the use of a single LTI model inadequate to describe the system dynamics over the entire operating window. Distillation columns are a representative example of process systems that exhibit nonlinear dynamics and gain directionality when operated in the high-purity region [1–3].

To meet with the increasing performance demands in the process industry, model-based control strategies are commonly utilized for optimal process operation (e.g., see [4–8] and the citations therein for model-based control of distillation columns). These control approaches require accurate dynamic models to obtain satisfactory closed-loop performance and robustness with respect to process uncertainties. Describing the dynamics of chemical processes using first-principles laws often leads to complex, nonlinear models consisting of a large number of ordinary differential equations, which are typically computationally intractable for real-time control applications. For example, the dynamics of a distillation column can be described by a set of nonlinear, *differential algebraic equations* (DAEs), whose numerical complexity (in terms of the number of equations) can increase significantly with the number of theoretical trays, which can be in the order of hundreds in the high-purity case. Another common complication in first-principles modeling of chemical processes can arise from the need for rigorous descriptions of physicochemical phenomena governing the system dynamics (e.g., thermodynamics driving vapor–liquid equilibria in a distillation column) that may not be obtained straightforwardly.

Data-driven modeling (system identification) methods offer an alternative modeling approach to address the inherent challenges of first-principles modeling of process systems. These approaches enable arriving at a relatively simple, control-relevant description of the system dynamics. To achieve this objective, the primary question is which (possibly nonlinear) model structure(s) should be selected to capture the behavior of the process with a low complexity description.

In this paper, the concept of *linear parameter-varying* (LPV) identification [9] is exploited to describe the nonlinear dynamics of process systems over a wide range of operating conditions. As a generalization of the classical concept of *gain scheduling*, the LPV framework is able to describe nonlinear process dynamics in terms of a time-varying linear dynamical relation between the system input and output signals. The variations of the latter dynamical relation depend on a so-called *scheduling signal*, which represents the changes in the operating conditions of the process. Thus, LPV models preserve the advantageous properties of LTI models, while being able to represent a large class of nonlinear systems [10]. LPV model-based control has been applied in many application areas (e.g., aerospace engineering, automotive applications, high-tech systems) as it benefits from the extension of successful LTI control design strategies, such as *proportional-integral-derivative* (PID) control [11], *model predictive control* (MPC) [12], optimal control [13], and robust control [14] (see [15–25,10] and the citations therein).

This work explores the applicability of various LPV identification approaches for data-driven modeling of chemical processes using a high-purity distillation column case study. To give an overview of the available approaches, several methods that use *input–output* (IO) and *series-expansion* model structures are reviewed and compared. In view of these approaches, two different LPV identification perspectives are explored to describe the process dynamics. The first perspective, called the *local approach*, relies on the identification of multiple LTI models at several operating points of the process, which is followed by the interpolation of the resulting models over the entire operating range. Several widely used interpolation schemes and methods (see [19,26–30,21,17]) are studied. The other investigated LPV identification approach, called the *global approach*, is based on a single data set that spans over a large range of operating conditions. The data set is used to identify the functional dependencies of a linear model structure on the scheduling variable. To explore the performance of the global approach, a classical prediction error minimization approach using an input–output polynomial model structure with a priori chosen parametrization of the scheduling dependencies [31,20] is compared with an *orthonormal basis functions* (OBF's) based model estimator [17] and a *least-square support vector machine* (LS-SVM) based non-parametric approach [32]. The OBF method, which uses a series expansion based model structure, has been shown to exhibit high performance in the modeling of a tank reactor

[28], while the LS-SVM approach is a state-of-the-art method to describe large-scale dynamics with possibly unknown nonlinearities [33,34].

The primary motivation for investigating the above mentioned local and global approaches for data-driven modeling of a high-purity distillation column originates from [35,27], where significant difficulties are reported for the local LPV identification approaches. The objective of this work is twofold: (i) provide a complete overview of the local LPV approaches given in [27] and (ii) investigate the open questions related to the performance of local and global LPV approaches. It is observed that directionality and changing effective order of local models lead to degradation of the local estimates of the process dynamics using IO model structures, whereas series expansion model structures based approaches (e.g., the OBF method) and global estimation techniques tend to describe the process dynamics adequately.

The paper is organized as follows. Section 2 introduces the general framework of LPV models, along with the commonly used model structures. In Sections 3 and 4, the local and global LPV identification approaches and the particular methods investigated in this paper are briefly reviewed. In Section 5, a first-principles model and operational settings of a high-purity distillation column are presented. The distillation column model is used as a benchmark system to investigate the properties of LPV identification approaches. In Section 6, a comparative analysis of the performance of the investigated approaches for the benchmark system at hand is given, followed by the conclusions in Section 7.

2. LPV models

The class of LPV systems 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* $p : \mathbb{Z} \to \mathbb{P}$ with a *scheduling "space"* $\mathbb{P} \subseteq \mathbb{R}^{n_\mathbb{P}}$. The scheduling variable is used to indicate the changes in the dynamical signal relations of the process at different operating conditions. This modeling concept can be understood as an embedding of the nonlinear behavior in the solution set of a linear structure, which can be achieved by following two different approaches called *local* and *global* modeling (discussed later). Converting first-principles nonlinear models to LPV form and selecting the scheduling variable is discussed in [10]. See Section 5.4 for an example how this scheduling variable is synthesized for the distillation column case study.

2.1. Representations and model structures

In discrete time, the dynamic description of an LPV system $\mathcal{S}$ can be formalized as a convolution in terms of p and the inputs u defined as

$$y(k) = \sum_{i=0}^{\infty} g_i(p, k)u(k-i), \tag{1}$$

where $u : \mathbb{Z} \to \mathbb{R}^{n_\mathbb{U}}$, $y : \mathbb{Z} \to \mathbb{R}^{n_\mathbb{Y}}$ denote the inputs and outputs of $\mathcal{S}$ with $n_\mathbb{U}, n_\mathbb{Y} \geq 1$, respectively, and $k \in \mathbb{Z}$ is the discrete time. The coefficients $g_i : \mathbb{P}^\mathbb{Z} \times \mathbb{Z} \to \mathbb{R}^{n_\mathbb{Y} \times n_\mathbb{U}}$ of (1) are functions of the scheduling variable and they define the varying linear dynamical relation between u and y. This description can also be seen as a *series expansion representation* of $\mathcal{S}$ in terms of the so called *pulse basis* $\{q^{-i}\}_{i=0}^{\infty}$, where q is the time-shift operator (i.e., $q^{-i}u(k) = u(k-i)$). It can be proven that for an asymptotically stable $\mathcal{S}$, the expansion (1) is convergent [10].

If the functions g_i only depend on the instantaneous value of the scheduling signal (i.e., $g_i(p(k))$) then the dependence is called *static*. Otherwise, the dependence is called *dynamic*, as the given coefficient also depends on time-shifted values of p. An important property of LPV systems is that for a constant scheduling signal (i.e., $p(k) = \bar{p}$ for all $k \in \mathbb{Z}$) (1) is equal to a convolution describing an LTI system as each $g_i(p, k)$ becomes constant.

Representation of $\mathcal{S}$ can be also formalized in terms of an *input–output* (IO) form (difference equation):

$$y(k) = -\sum_{i=1}^{n_\mathrm{a}} a_i(p(k))y(k-i) + \sum_{j=0}^{n_\mathrm{b}} b_j(p(k))u(k-j), \tag{2}$$

where $a_i : \mathbb{P} \to \mathbb{R}^{n_\mathbb{Y} \times n_\mathbb{Y}}$ and $b_j : \mathbb{P} \to \mathbb{R}^{n_\mathbb{Y} \times n_\mathbb{U}}$ are matrix coefficient functions dependent on $p(k)$. For the sake of simplicity, all coefficient functions in (2) and in the sequel are assumed to have *static dependency*. For more details on LPV representations and types of dependencies see [10].

In LPV identification, a dynamical model of the system is derived using measured data, which corresponds to the estimation of each a_i, b_j functions in (2) or each g_i in (1). Such an estimation can be accomplished using various parameterizations of these functions. A particular parametrization corresponds to a model structure represented model set in which the model that describes the best the relations of the observed data is searched for. Besides parameterizations of (2) (corresponding to the well-known LPV-IO model structures [10]), a particularly attractive model structure in LPV identification follows from the truncation of (1) to a finite number of expansion terms. Assuming static dependence of g_i on p, the resulting approximative model is

$$y(k) = \sum_{i=0}^{n_\mathrm{e}} g_i(p(k))u(k-i), \tag{3}$$

which is in fact the LPV version of the well known *finite impulse response* (FIR) models. In addition, noise or disturbances in the system can be modeled in terms of parametrized noise models. These choices and model structures will be discussed in the sequel.

3. LPV identification using the local approach

The local approach refers to the methodology to construct the LPV model as an 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 the model structure and the interpolation scheme as well as the number and location of the used operating points are application specific (e.g., see [10]). Some of the local approaches have been recently applied on distillation columns [35,27], on fermentation processes [36] and on a continuous tank reactor [28,36]. In the sequel, we intend to give an overview on all available methods for the LPV modeling of process systems in the local setting.

3.1. Choice of operating points

Every constant $\bar{p} \in \mathbb{P}$ in the operating envelop is associated with a $(\bar{u}, \bar{y})$ corresponding to the steady-state solutions of the first-principle based model at that purity level. To cover these operating points, often an adequate (equidistant) gridding of $\mathbb{P}$ (in terms of grid points $\mathcal{P} = \{\bar{p}^{(\tau)}\}_{\tau=1}^{N_\mathrm{loc}}$) is required to represent the variation of the local behaviors. Selection of these operating points not only defines the achievable accuracy of the resulting model, but also sets the number of experiments needed around these operating points to derive local estimates of the behavior. Note that optimized allocation of $\{\bar{p}^{(\tau)}\}_{\tau=1}^{N_\mathrm{loc}}$ can seriously lower the number of required LTI experiments [37] and hence the cost of the experimental campaign. Typically, in the process domain, adequate choice of grid

points is recommended to be 3–6 regarding each dimension of the scheduling variable.

3.2. 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 [38]. In practice, to gather measurement data at each operating point $\overline{p}^{(\tau)} \in \mathcal{P}$, the system is excited with an input signal $\tilde{u}$ selected to be a small amplitude white noise, PRBS or multisine around the corresponding $(\overline{u}^{(\tau)}, \overline{y}^{(\tau)})$ steady-state values. The response y is captured in terms of data sets $\mathcal{D}_\tau = \{(\tilde{u}(k)+\overline{u}^{(\tau)}, y(k)+v(k))\}_{k=1}^{N}$ with length $N>0$, where v denotes the cumulative effect of measurement and process noise. Denote by $y_\mathrm{m} \triangleq y+v$ the measured response and note that the excitation u is designed such that the system remains around the specified operating condition, but, at the same time, the data is informative enough to describe the dynamics of the system around $\overline{p}^{(\tau)}$. Such an experiment can be designed in a least costly manner and conducted even under closed-loop MPC control to minimize off-spec products (see [39,40] and the references therein). Furthermore, under the assumption that no significant deviation from the operating condition is introduced, the underlaying local behavior will remain to be LTI and hence v, in general, will have a rational spectra which can be modeled as a filtered white noise process. Respecting the resulting noise conditions, PEM can be applied on each $\mathcal{D}_\tau$ utilizing an LTI, e.g., *outputerror* (OE), model structure in the deviation variables $(\tilde{u}^{(\tau)} \triangleq u-\overline{u}^{(\tau)}, \tilde{y}^{(\tau)} \triangleq y-\overline{y}^{(\tau)})$:

$$\hat{y}^{(\tau)}(k) = -\sum_{i=1}^{n_\mathrm{a}} a_i^{(\tau)} \hat{y}^{(\tau)}(k-i) + \sum_{j=0}^{n_\mathrm{b}} b_j^{(\tau)} \tilde{u}^{(\tau)}(k-j), \tag{4a}$$

$$y_\mathrm{m}(k) = \hat{y}^{(\tau)}(k) + \overline{y}^{(\tau)} + \varepsilon(k), \tag{4b}$$

with $[\, a_1^{(\tau)} \;\; \dots \;\; a_{n_\mathrm{a}}^{(\tau)} \;\; b_0^{(\tau)} \;\; \dots b_{n_\mathrm{b}}^{(\tau)} \,]$ being the unknown parameters to be estimated with the objective that $\hat{y}^{(\tau)} \approx \tilde{y}^{(\tau)}$ and $\varepsilon \approx v$. The resulting parameter estimates correspond to a set of LTI-IO models $\{\mathcal{M}_\tau\}_{\tau=1}^{N_\mathrm{loc}}$. The orders $(n_\mathrm{a}, n_\mathrm{b})$ can be selected by cross-correlation based residual analysis, by BIC or AIC, or via regularization based sparse estimators [38].

Note that following this methodology, each local model is obtained in terms of local deviation variables $(\tilde{u}^{(\tau)}, \tilde{y}^{(\tau)})$ which prevents their direct interpolation for getting a global model of the plant behavior. Hence, it is important to write the resulting models in a compact form such that the steady-state values $(\overline{u}^{(\tau)}, \overline{y}^{(\tau)})$ are implicitly incorporated into the LTI descriptions. This is possible by introducing a new input signal $\tilde{u} = [\, u^\top \;\; 1 \,]^\top$ to write each identified local model as

$$\begin{aligned}\hat{y}^{(\tau)}(k) = &-\sum_{i=1}^{n_\mathrm{a}} \hat{a}_i^{(\tau)} \hat{y}^{(\tau)}(k-i) + \sum_{j=1}^{n_\mathrm{b}} \underbrace{\begin{bmatrix} \hat{b}_j^{(\tau)} & -\hat{b}_j^{(\tau)} \overline{u}^{(\tau)} \end{bmatrix}}_{\tilde{b}_j^{(\tau)}} \tilde{u}(k-j) \\ &+ \underbrace{\begin{bmatrix} \hat{b}_0^{(\tau)} & \left(I + \sum_{i=1}^{n_\mathrm{a}} \hat{a}_i^{(\tau)} \right) \overline{y}^{(\tau)} - \hat{b}_0^{(\tau)} \overline{u}^{(\tau)} \end{bmatrix}}_{\tilde{b}_0^{(\tau)}} \tilde{u}(k),\end{aligned} \tag{5}$$

where now $\hat{y}^{(\tau)}$, the output of the τ-th local model, provides approximation of the true output y, and $\hat{a}_i^{(\tau)}$ and $\hat{b}_j^{(\tau)}$ are the estimated model parameters. In this way, each estimated local model is transformed via (5) from local deviation variables to the nominal input u and output signals y.

3.3. Interpolation methods

Interpolation of a given (indexed) set of discrete data points $\mathcal{Z} = \{z_\tau \in \mathbb{R}\}_{\tau=1}^{N_\mathrm{loc}}$ 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 (e.g., see [41]). 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 3.4), the following "standard" approaches are commonly used [10,19,27,36].

3.3.1. RBF interpolation

Radial basis functions, like *Gaussian functions*, centered at the operating points $\overline{p}^{(\tau)}$ can be used as a simple methodology to interpolate $\mathcal{Z} = \{z_\tau\}_{i=1}^{N_\mathrm{loc}}$. The spread of the RBFs is optimized in terms of an objective function $\mathcal{V}$ to arrive at a smooth interpolation of the given data points in $\mathcal{Z}$. As the interpolation space $\mathbb{P}$ has dimensions $n_\mathbb{P}$ equal to the dimension of p, hence essentially often a multidimensional interpolation problem is required to be solved. To set up such a problem, the following paths can be taken:

- Choose $n_\mathbb{P}$-dimensional parametrized RBFs, like Gaussian functions

$$w^{(\tau)}(\,\cdot\,, \theta) = \exp(-\|p-\overline{p}^{(\tau)}\|_{\theta,2}), \tag{6}$$

  with $\|x\|_{\theta,2}^2 = x^\top \theta x$ and the positive symmetric $\theta \in \mathbb{P}^{n_\mathbb{P} \times n_\mathbb{P}}$ being the parameters of the basis
- Formulate the multidimensional 'basis' as the multiplication of 1D basis, corresponding to each scheduling variable separately:

$$w^{(\tau)}(\,\cdot\,, \theta) = \prod_{i=1}^{n_\mathbb{P}} w_i^{(\tau)}(\,\cdot\,, \theta), \quad \text{where} \quad w_i^{(\tau)}(\,\cdot\,, \theta) = \exp(-\theta_i |p_i - \overline{p}_i^{(\tau)}|). \tag{7}$$

Optimization of multi-dimensional RBFs represents a significantly more complex problem than in the 1D case, therefore only the multiple 1D-RBF interpolation method is investigated here. As an example for the $n_\mathbb{P} = 2$ case, the following trapezoidal RBF is introduced for each $\overline{p}^{(\tau)} = [\, \overline{p}_1^{(\tau)} \;\; \overline{p}_2^{(\tau)} \,]^\top$

$$w_i^{(\tau)}(\,\cdot\,, \theta) = \max\left(\min\left(\frac{\cdot - \theta_{i,1}^{(\tau)}}{\theta_{i,2}^{(\tau)} - \theta_{i,1}^{(\tau)}}, 1, \frac{\theta_{i,4}^{(\tau)} - \cdot}{\theta_{i,4}^{(\tau)} - \theta_{i,3}^{(\tau)}} \right), 0 \right), \tag{8}$$

with $i=1, 2$ and $\{\theta_{i,j}^{(\tau)}\}_{j=1}^{4} \in \mathbb{R}^4$ being the parameters of (8) chosen such that the trapezoid function is centered at $\overline{p}_i^{(\tau)}$ (i.e., $w_i^{(\tau)}(\overline{p}_i^{(\tau)}, \theta) = 1$). The resulting interpolated function in the $n_\mathbb{P} = 2$ case is

$$w(\overline{p}, \theta, \mathcal{Z}) = \sum_{\tau=1}^{N_\mathrm{loc}} z_\tau w_1^{(\tau)}(\overline{p}_1, \theta) w_2^{(\tau)}(\overline{p}_2, \theta). \tag{9}$$

The parameters collected in θ can be for example optimized based on inter-sample data $(z_{\overline{p}}, \overline{p}) \in \mathcal{Z}_\mathrm{int}$ (obtained from other experiments or by dividing the data set to an interpolation set and inter-sample set) via the minimization of $\mathcal{V}(\theta, \mathcal{Z}_\mathrm{int}) \triangleq \max_{(z_{\overline{p}}, \overline{p}) \in \mathcal{Z}_\mathrm{int}} |z_{\overline{p}} - w(\overline{p}, \theta)|$. This implies that to have a well-defined optimization problem, the following constraint should be satisfied:

$$\sum_{\tau=1}^{N_\mathrm{loc}} w_1^{(\tau)}(\overline{p}_1, \theta) w_2^{(\tau)}(\overline{p}_2, \theta) = 1, \quad \forall \overline{p} \in \mathbb{P}. \tag{10}$$

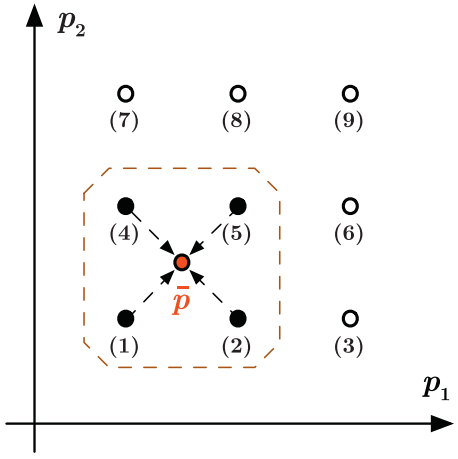

Fig. 1. Illustration for bilinear interpolation.

Note that in case of Gaussian functions, the above constraint is not directly feasible and hence it is enforced by re-normalization: $\tilde{w}(\bar{p},\theta,\mathcal{Z}) = w(\bar{p},\theta,\mathcal{Z})/\sum_{\tau=1}^{N_{\text{loc}}}\prod_{i=1}^{n_{\mathbb{P}}} w_i^{(\tau)}(\bar{p}_i,\theta)$.

3.3.2. Polynomial interpolation

In polynomial interpolation, n-th-order monomials are employed as basis functions. Hence, in the $n_{\mathbb{P}} = 2$ case, a polynomial function is formulated as:

$$w(\bar{p},\theta,\mathcal{Z}) = \sum_{i=0}^{n}\sum_{j=0}^{n}\theta_{i,j}\ \ \bar{p}_1^i\ \ \bar{p}_2^j, \tag{11}$$

with $\theta_{i,j} \in \mathbb{R}$ being the parameters of the polynomials collected in θ. For example, these parameters are chosen such that the ℓ_2 approximation error $\mathcal{V}(\theta,\mathcal{Z}) \triangleq \sum_{\tau=1}^{N_{\text{loc}}}(z_\tau - w(\bar{p}^{(\tau)},\theta))^2$ is minimized, but other objective functions based on different choice of the error measure ℓ_1, ℓ_∞, etc, are also applicable. Note that if $n < N_{\text{loc}}$, then inter-sample data is not required to optimize θ nor it is ensured that $w(\bar{p}^{(\tau)},\theta,\mathcal{Z}) = z^{(\tau)}$.

3.3.3. Bilinear interpolation

This approach corresponds to the piece-wise linear interpolation of variables on a multidimensional grid. In case of $n_{\mathbb{P}} = 2$ and an equidistant grid $r \times r$ with size $r = 3$, let $\mathbb{I}$ be the set of indices of all the black points given in Fig. 1, i.e., $\mathbb{I} = \{1, 2, 4, 5\}$. Let $\bar{p} \in \mathbb{P}$ such that $\bar{p}_1^{(\tau)} \le \bar{p}_1 \le \bar{p}_1^{(\tau+1)}$ and $\bar{p}_2^{(\tau)} \le \bar{p}_2 \le \bar{p}_2^{(\tau+r)}$ for a $\tau \in \mathbb{I}$ as exemplified in Fig. 1 for $\tau = 1$. Note that, due to the equidistant gridding, $\bar{p}_2^{(\tau)} = \bar{p}_2^{(\tau+1)}$, $\bar{p}_1^{(\tau)} = \bar{p}_1^{(\tau+r)}$ and $\Delta_{\bar{p}} = |\bar{p}_1^{(\tau+1)} - \bar{p}_2^{(\tau)}||\bar{p}_2^{(\tau+r)} - \bar{p}_2^{(\tau)}|$ is constant. The interpolated function with bilinear basis can be written as

$$\begin{aligned} w(\bar{p},\mathcal{Z}) = {} & \frac{z_\tau}{\Delta_{\bar{p}}}|\bar{p}_1^{(\tau+1)} - \bar{p}_1||\bar{p}_2^{(\tau+r)} - \bar{p}_2| + \frac{z_{\tau+1}}{\Delta_{\bar{p}}}|\bar{p}_1^{(\tau+1)} - \bar{p}_1||\bar{p}_2 - \bar{p}_2^{(\tau)}| \\ & + \frac{z_{\tau+r}}{\Delta_{\bar{p}}}|\bar{p}_1 - \bar{p}_1^{(\tau)}||\bar{p}_2^{(\tau+r)} - \bar{p}_2| + \frac{z_{\tau+r+1}}{\Delta_{\bar{p}}}|\bar{p}_1 - \bar{p}_1^{(\tau)}||\bar{p}_2 - \bar{p}_2^{(\tau)}|, \end{aligned} \tag{12}$$

where each term corresponds to a sum of the z_τ's in the neighborhood of $\bar{p}$ (dashed line in Fig. 1) weighted by the relative distance of $\bar{p}$ from the corresponding $\bar{p}^{(\tau)}$'s. At the boundaries of the grid (circles in Fig. 1), the interpolated function simplifies to a 1D interpolation. Note that these approaches can be trivially extended for any scheduling and parameter dimensions or even non-equidistant gridding.

3.4. Interpolation schemes

There are four interpolation schemes which are regularly applied to develop LPV models: (i) coefficient, (ii) output, (iii) input and (iv) series-expansion based interpolation schemes. Due to space restrictions, only schemes (i), (ii) and (iv) are investigated. In our case study, scheme (iii) is outperformed by these approaches (see [42] for performance analysis of scheme (iii)) and use of this scheme is only suggested for Hammerstein type of systems. The above mentioned schemes can be used with any of the previously discussed interpolation methods to construct the LPV model.

(i) Coefficient interpolation: In this scheme, the local models $\{\mathcal{M}_\tau\}_{\tau=1}^{N_{\text{loc}}}$ (in the form of (5)) are interpolated based on their model coefficients $\{\hat{a}_i^{(\tau)}\}_{i=1,\tau=1}^{n_{\text{a}},N_{\text{loc}}}$ and $\{\tilde{b}_j^{(\tau)}\}_{j=0,\tau=1}^{n_{\text{b}},N_{\text{loc}}}$ to obtain an LPV model in an IO form. Subsequently, the resulting LPV model in the variables $(\hat{y},\tilde{u})$ is given by

$$\hat{y}(k) = -\sum_{i=1}^{n_{\text{a}}} a_i(p(k))\hat{y}(k-i) + \sum_{j=0}^{n_{\text{b}}} b_j(p(k))\tilde{u}(k-j), \tag{13}$$

where $\hat{y}$ is the approximation of the true system output y and $a_i(\cdot)$ and $b_j(\cdot)$ are the element-wise interpolation of the matrix sequences $\{\hat{a}_i^{(\tau)}\}_{i=1,\tau=1}^{n_{\text{a}},N_{\text{loc}}}$ and $\{\tilde{b}_j^{(\tau)}\}_{j=0,\tau=1}^{n_{\text{b}},N_{\text{loc}}}$.

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

$$\hat{y}(k) = c(p(k), \hat{y}^{(1)}(k), \ldots, \hat{y}^{(N_{\text{loc}})}(k)), \tag{14}$$

with $\hat{y}^{(\tau)}$ being the output response of the local LTI model $\mathcal{M}_\tau$ (in the form of (5)) for the augmented input signal $\tilde{u}$. Note that, in this scheme, the interpolation points $\{z_\tau\}_{\tau=1}^{N_{\text{loc}}}$ are described by the sequence $\{\bar{y}^{(\tau)}\}_{\tau=1}^{N_{\text{loc}}}$ for the RBF and the bilinear case, but after obtaining the interpolation function (9) or (12) for $\hat{y}$ it is evaluated as

$$\hat{y}(k) = w(p(k),\theta,\mathcal{Z}_k), \quad \mathcal{Z}_k = \{\hat{y}^{(\tau)}(k)\}_{\tau=1}^{N_{\text{loc}}}. \tag{15}$$

In the polynomial case, the interpolation decomposes as

$$\hat{y}(k) = c(k) = \sum_{\tau=1}^{N_{\text{loc}}} w^{(\tau)}(p(k),\theta)\ \ \hat{y}^{(\tau)}(k), \tag{16}$$

where each $w^{(\tau)}$ is a separate matrix polynomial function given in the form of (11). A particular difference of the output interpolation scheme 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. To optimize such transitions in terms of c, additional information (data set) about the transient ("inter-sample") behavior of the system is therefore required. Hence, 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 designed in terms of bilinear interpolation or RBFs with equal spread. On the other hand, using $\mathcal{D}_{\text{global}}$ requires modifying the objective function to $\mathcal{V}(\theta,\mathcal{D}_{\text{global}}) = \sum_{k=1}^{N}\|y(k) - c(k)\|_2^2$ and optimizing the parameters accordingly. This approach can also be used with RBF interpolation. Such an interpolation approach, can be considered as a "glocal" approach [43,19,21], which borders the set of global LPV identification approaches.

(iv) Series-expansion scheme: So far, interpolation schemes have been considered that assemble the global model in an IO representation form. However, given the convolution form (1), interpolation of the local models can be performed in a series expansion form as well. To illustrate why this approach is attractive, it is important to note that the transfer function $F \in \mathcal{RH}_2^{n\times m}$ of any local (augmented)

LTI model, where $\mathcal{RH}_2$ is the *Hardy space* of square integrable, real, rational, and stable transfer functions, can be written as

$$F(q) = w_0 + \sum_{i=1}^{\infty} w_i \phi_i(q), \tag{17}$$

where $\{\phi_i\}_{i=1}^{\infty}$ is a basis for $\mathcal{RH}_2$ characterized in terms of a set of stable pole locations $\lambda_1, \ldots, \lambda_{n_r}$, q is the time-shift operator and $w_i \in \mathbb{R}^{n_\mathbb{Y} \times n_\mathbb{U}}$ [44,45]. This gives not only a uniform representation of all $\mathcal{M}_\tau$, but due to the infinite impulse response characteristics of each $\phi_i(z)$, a faster convergence rate of the expansion than by FIR representations. By using a truncated expansion in (17), an attractive model structure for LTI identification results with a well-defined stochastic framework (see [44] for OBFs and their properties).

The series expansion (17) can be extended to LPV systems [10], such that for a given basis $\{\phi_i\}_{i=1}^{\infty}$ an asymptotically stable LPV system can be written in the time domain as

$$y(k) = w_0(p,k)\tilde{u} + \sum_{i=1}^{\infty} w_i(p,k)\phi_i(q)\tilde{u}, \tag{18}$$

where w_i are matrix functions with appropriate dimension and with dynamic dependence on p. An obvious choice of model structure is to use a truncated expansion (i.e., truncating (18) to a finite sum in terms of $\{\phi_i\}_{i=1}^{n_e}$) and to assume static dependence of the coefficients similar to the FIR case (3):

$$\hat{y}(k) = w_0(p(k))\tilde{u} + \sum_{i=1}^{n_e} w_i(p(k))\phi_i(q)\tilde{u}. \tag{19}$$

As in the FIR case, this structure is linear in the coefficients $\{w_i\}_{i=1}^{n_e}$ and therefore gives a simple form to develop interpolation of the local models by first expanding all $\mathcal{M}_\tau$ in terms of the basis $\{\phi_i\}_{i=1}^{n_e}$ and then interpolating the local expansion coefficients by any previously discussed interpolation method (to arrive at the interpolated expansion coefficients $w_i(p(k))$ defining the LPV form (19)). This form not only provides a simple interpolation recipe, but also ensures asymptotic stability of the resulting LPV model. In addition, it can be applied to cases where the local model order changes.

To have an efficient model structure in terms of (19) with a minimal number of interpolated coefficients, a fast convergence rate of (18) is required. This corresponds to an optimal selection of an OBF set $\{\phi_i\}_{i=1}^{n_e}$ such that the approximation error of (19) is minimal w.r.t. the system. In terms of the Kolmogorov theory for OBF models [46], this relates to the optimization of the pole locations $\lambda_1, \ldots, \lambda_{n_r}$ of the OBFs and also an optimal choice of n_r w.r.t. the set of all possible pole locations associated with the local LTI aspects. In practice, this is accomplished with a so-called *Fuzzy Kolmogorov c-Max* (FK*c*M) algorithm which, based on samples of the local pole locations (obtained through LTI identification of the system at some operating points), is capable of efficiently solving the optimal OBF selection problem [17].

4. LPV identification using the global approach

Compared to the previously discussed local methodology, the global identification concept corresponds to a drastically different approach of constructing an LPV of the system from data. By the global identification concept, an LPV model is estimated in one step based only on a single data set $\mathcal{D}_{\text{global}}$ with a varying scheduling trajectory. A commonly applied method to implement this approach is to use the LPV extension of the PEM framework for LPV-IO models in the form of (2) (see [10,15]):

$$y(k) = -\sum_{i=1}^{n_a} a_i(p(k))y(k-i) + \sum_{j=0}^{n_b} b_j(p(k))u(k-j) + v(k), \tag{20}$$

where $v(k)$ is a zero-mean stochastic noise process. Almost all global LPV identification methods, including subspace schemes, require a linear parametrization of each model coefficient function a_i and b_j in terms of an a priori chosen set of basis functions $\{\psi_s\}_{s=1}^{n_\psi}$, $\psi_s : \mathbb{P} \to \mathbb{R}$

$$\zeta_i(\cdot, \theta) = \theta_{i,0} + \sum_{s=1}^{n_\psi} \theta_{i,s}\psi_s(\cdot), \tag{21}$$

with $\zeta_i = a_i$ for $1 \le i \le n_a$, $\zeta_{n_a+1+j} = b_j$ for $0 \le j \le n_b$ and $\theta_{i,s} \in \mathbb{R}^{n_\mathbb{Y} \times n_\mathbb{U}}$ being the unknown parameters to be estimated. This approach requires an efficient selection of $\{\psi_s\}_{s=1}^{n_\psi}$ such that (21) can capture the underlaying nonlinearities of the process adequately. However, to perform the selection of basis functions for process systems, a complicated analysis of a first-principle model is often required where the economical benefits of data-driven modeling can be easily lost. This has led to an increasing interest in the so-called *non-parametric* methods that provide direct estimation of the coefficient functions [47,23]. Furthermore, most process models, like the distillation column studied in Section 5, is described by a large-scale first-principles model which can be drastically reduced to a much lower order model (2nd-order model in case of the distillation column) without significant loss of accuracy (see Section 6.1). However, in most cases, the original nonlinearities of the model translate to heavy nonlinear p-dependencies in a low order LPV approximation. Hence, besides investigating two classical parametrization based approaches (i.e., the OE estimation scheme and an OBF based global estimator), a recently introduced non-parametric LPV identification method known as the LPV *least-square support vector machine* (LPV LS-SVM) method [47] is also studied. The latter method can approximate the nonlinear functional dependencies of the model directly from data. Note that all the presented approaches can be applied to MIMO systems with any finite $n_\mathbb{P}$.

4.1. The prediction error minimization approach

Due to the linearity of the considered model (20), the one-step ahead predictor in the case of a generalized polynomial noise model has a similar form as in the LTI case under the commonly taken assumption of noise free measurement of the scheduling variable p. This facilitates the extension of the classical prediction error minimization approaches as discussed in [10,21,48].

Provided that, in the process under study, the noise $v(k)$ is assumed to be output additive and white (i.e., OE type), the one-step-ahead prediction $\hat{y}(k|\theta, k-1)$ in terms of (20) is equivalent to the simulated response $\hat{y}(k)$ of

$$\hat{y}(k) = \sum_{i=1}^{n_g} \zeta(p(k),\theta)x_i(k) = \sum_{i=1}^{n_g} \theta_{i,0}x_i(k) + \sum_{i=1}^{n_g}\sum_{s=1}^{n_\psi} \theta_{i,s}\psi_s(p(k))x_i(k), \tag{22}$$

with $x_i(k) = \hat{y}(k-i)$ for $i \in \mathbb{I}_1^{n_a}$ and $x_i(k) = u(k-j)$ with $j = i - n_a - 1$ for $i \in \mathbb{I}_{n_a+1}^{n_g=n_a+n_b+1}$. Here $\mathbb{I}_{s_1}^{s_2} = \{i \in \mathbb{Z} | s_1 \le i \le s_2\}$ denotes an index set. Minimizing the ℓ_2-loss of the prediction:

$$V(\theta, \mathcal{D}_{\text{global}}) = \frac{1}{N}\sum_{k=1}^{N} \|y(k) - \hat{y}(k)\|_2^2, \tag{23}$$

w.r.t. the parameter vector θ under the constraint (22) leads to a classical nonlinear optimization problem that can be solved by gradient descent or pseudo-linear regression [21,48]. Note that other noise structures such as ARX, ARMAX, BJ, etc. can also be handled by the LPV-PEM methods [48,21].

4.2. The OBF approach

An other popular approach leads through the use of the OBF's based expansion model structures. If the parametrization of the expansion coefficients w_i in (19) is formulated as in (21), then the following predictor form is obtained

$$\hat{y}(k|\theta, k-1) = \hat{y}(k) = \zeta_0(p(k), \theta)u(k) + \sum_{i=1}^{n_e} \zeta_i(p(k), \theta)\phi_i(q)u(k), \tag{24}$$

for an OE noise model case. Note that this expression is in u and not $\tilde{u}$. In fact, (24) can be written as

$$\hat{y}^\top(k) = \varphi(k)\theta, \tag{25}$$

with

$$\varphi(k) = \begin{bmatrix} u^\top(k) & \psi_1(p(k))u^\top(k) & \dots & \psi_{n_\psi}(p(k))u^\top(k) \\ & \phi_1(q)u^\top(k) & \dots & \psi_{n_\psi}(p(k))\phi_{n_e}(q)u^\top(k) \end{bmatrix}, \tag{26}$$

and $\theta \in \mathbb{R}^{n_\mathbb{Y}(n_e+1)(n_\psi+1)\times n_\mathbb{U}}$ that contains all the parameters of $\{\zeta_i\}_{i=0}^{n_e}$ collected into a matrix according to the structure of (26). In this case, minimization of (23) corresponds to a least squares solution (see [17] for the stochastic properties of the OBF estimator).

4.3. The LS-SVM approach

The SVM approach exploits the reproducing kernel theory to provide non-parametric approximation of functions based on a series expansion with kernel functions evaluated on the observed data points. This is made possible by exploiting the dual solution of general least-squares problems. To explain the underlying estimation mechanism, let us start with the assumed model structure of the LS-SVM approach

$$y(k) = \sum_{i=1}^{n_g} \psi_i(p(k))\theta_i x_i(k) + e(k), \tag{27}$$

where each row of $\psi_i : \mathbb{R} \to \mathbb{R}^{n_\mathbb{Y}\times n_\psi}$ denotes an undefined, potentially infinite dimensional vector of basis functions, $\theta_i \in \mathbb{R}^{n_\psi \times n_\mathbb{U}}$ is the i-th parameter matrix, $e(k)$ is the equation error and $x_i(k)$ is defined as in Section 4.1. The estimates of the coefficient functions in the form of $\psi_i(p(k))\theta_i$ are obtained using the solution of the following optimization problem

$$\begin{aligned} \min_{\theta,e} \quad & \mathcal{J}(\theta, e) = \frac{1}{2}\sum_{i=1}^{n_g} \|\theta_i\|_F^2 + \frac{1}{2N}\sum_{k=1}^{N} \|e(k)\|_{\Gamma,2}^2, \\ \text{s.t.} \quad & e(k) = y(k) - \sum_{i=1}^{n_g} \psi_i(p_k)\theta_i x_i(k), \end{aligned}$$

with $\|\theta_i\|_F = \sqrt{\text{trace}(\theta_i^\top \theta_i)}$ being the Frobenius matrix norm and $\|e(k)\|_{\Gamma,2}^2 = e(k)^\top \Gamma e(k)$ being the weighted Euclidian norm where the symmetric, commonly block diagonal $\Gamma = \text{diag}(\gamma_1, \dots, \gamma_{n_\mathbb{Y}})$, $\Gamma \in \mathbb{R}_+^{n_\mathbb{Y}\times n_\mathbb{Y}}$, serves as a *regularization parameter*, influencing the complexity of the model in terms of the Frobenius-norm of θ. An interesting feature of this optimization problem is that it can be solved without specifying the functions ψ_i or estimating the parameters θ_i. The key step is to derive the solution in the dual space of this optimization problem. Instead of specifying ψ_i, its inner product $\psi_i\psi_i^\top$ is defined by an a priori chosen kernel function K_i, which is known as the *kernel trick* [47]. Hence, the underlying choice of nonlinear basis functions ψ_i is avoided by specifying a series expansion structure of the unknown coefficient function ϕ_i described by the kernel K_i. 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 can apply a diagonal Gaussian RBF kernel which is commonly used in nonlinear function estimation. Hence, the kernel is formulated as

$$K_i(p(k), p(l)) = \exp\left(-(p(k) - p(l))^\top \Sigma_i (p(k) - p(l))\right), \tag{28}$$

where $\Sigma_i = \text{diag}(\sigma_{i,1}, \dots, \sigma_{i,n_\mathbb{P}})$ consist of $\sigma_{i,j} > 0$ which specify the widths of the RBFs. Together with the regularization parameters $\{\gamma_j\}_{j=1}^{n_\mathbb{Y}}$, the kernel widths $\{\sigma_{i,j}\}_{i=1,j=1}^{n_g,n_\mathbb{P}}$ 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 estimates. Based on K_i, the following matrices are constructed

$$\begin{aligned} [\Omega]_{s,k} &= \sum_{i=1}^{n_g} [\Omega_i]_{s,k}, \\ [\Omega_i]_{s,k} &= x_i^\top(s)(K_i(p(s), p(k)))x_i(k), \end{aligned}$$

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 $\beta \in \mathbb{R}^{N\times n_\mathbb{Y}}$

$$\hat{\beta} = \left(\frac{1}{N}\Omega + \Gamma^{-1} I_N\right)^{-1} \frac{1}{N} Y, \tag{29}$$

where $Y = [\, y(1) \quad \dots \quad y(N)\,]^\top$. Based on $\hat{\beta} = [\, \hat{\beta}_1^\top \quad \dots \quad \hat{\beta}_N^\top\,]^\top$, the estimates of the coefficient functions are given as

$$\hat{\zeta}_i(\,\cdot\,) = \psi_i(\,\cdot\,)\theta_i = \sum_{s=1}^{N} \hat{\beta}_s^\top x_i^\top(s) K_i(p(s),\,\cdot\,). \tag{30}$$

Hence, a non-parametric estimate of the coefficient functions directly follows from the series expansion form of (30) by computing the dual estimate (29) (see [32] for the properties of this estimator). Note that (27) corresponds to an ARX noise model which might not be a valid assumption on the data-generating system (consider just the previously mentioned typically OE nature of the measurement noise). To avoid bias resulting from this improper noise model, an instrumental variable based form of the LS-SVM has been developed in [47] and other modifications of the LS-SVM were proposed in [49,50]. These extensions make possible to apply the LS-SVM approach for general noise scenarios.

5. Case study: high-purity distillation column

To show the capabilities of data-driven LPV modeling of process systems, the performance of the LPV identification methods introduced in Sections 3 and 4 are investigated for data-driven modeling of a high-purity distillation column.

Table 1
Settings of the high-purity distillation column.

N_T	Number of trays	110
N_F	Feed stage location	39
α	Relative volatility	1.12
F	Molar feed flow	215 kmol/min
x_f	Mole fraction of the feed	0.65
q_f	Liquid phase fraction of the feed	1.0
M	Molar holdup (all trays)	30 kmol

5.1. Motivation for LPV identification

High-purity distillation columns are well-known for their *nonlinear dynamics* and *directionality characteristics*, which become more significant as the operating conditions approach the *high-purity* region [51],[52] (a thorough analysis of the nonlinear behavior of the distillation column under study is given in [42]). Due to directionality effects, the system response is dominated by the high-gain direction (i.e., only one product composition can be controlled effectively), which significantly limits the performance of linear single-input single-output controllers to regulate the top and bottom compositions. To address this problem, model predictive control strategies can be used to control the multivariable system dynamics. Such control strategies require an accurate model to adequately predict the system dynamics, based on which the optimal process operation is sought. However, because high-purity distillation columns exhibit a strong nonlinear behavior, small deviations from the operating point typically result in different dynamics. This indicates that a linear model may be insufficient to describe the system dynamics over varying operating conditions. Hence, an alternative modeling solution is required to simultaneously preserve the simplicity of LTI models for controller design and to ensure the attainment of an adequate description of the process dynamics over a wide range of operating conditions. As the LPV framework offers such a data-driven modeling paradigm, the applicability of the LPV identification approaches to the high-purity distillation column is investigated and compared in terms of their achieved simulation error and representation of the local frequency response.

5.2. First-principles modeling

Under the assumptions of constant relative volatility, 100% tray efficiency, constant molar flows, and constant vapor and liquid holdups, the first-principles modeling framework in [1] is adopted to describe the dynamics of a binary distillation column. The system dynamics are represented by a set of DAEs, which describe the component concentrations, as well as the vapor and liquid flows along the distillation column. The settings of the high purity distillation column under study are given in Table 1. A detailed description of the first-principles model and the system parameters can be found in [42].

The first-principles model is utilized as the data-generating system for the identification study. In the process at hand, the liquid flow rate L and the vapor flow rate V (manipulated through the reflux rate and re-boiler duty, respectively) are used as the manipulated variables to control the operation of the high-purity distillation column (i.e., $u_1 = L$ and $u_2 = V$). The system outputs consist of the composition of top (x_d) and bottom (x_b) products (i.e., $y_1 = x_d$ and $y_2 = x_b$). The resulting distillation column model describes a large-scale (110th order), nonlinear, 2×2 multi-input multi-output system.

5.3. Measurements

To generate realistic measurement records for system identification, the first-principles model is simulated in continuous-time, while discrete-time input/output data is collected with a sampling time $T_s = 5$ min. The sampling time is chosen 20 times faster than the time constant of the fastest step response of the system for the operating region under investigation. The inputs of the system are manipulated through zero-order-hold actuation synchronized with respect to the sampling time. The generated measurements are corrupted by a discrete-time output additive white noise v with a *signal-to-noise ratio*[1] (SNR) of 15 dB.

5.4. Choice of the scheduling variables

An important step in LPV identification is the choice of the scheduling variable p, which governs how an LPV representation of the dynamical behavior is defined. For the high-purity distillation column, a natural choice for the scheduling variables is the top and bottom product compositions (i.e., $p(k) = [\, x_d(k) \quad x_b(k) \,]^\top$). This is because the latter variables can uniquely characterize different operating points of the system at steady state. The operating region is selected such that it entails a large set of local operating points described by the top and bottom compositions in the region of $\mathbb{P} = [0.95, \quad 0.995] \times [0.02, \quad 0.1]$. The dynamics of the distillation column are identified either by operating the column at different top and bottom purity levels resulting in data sets with constant p (local approach), or by varying the purity levels (i.e., p) during operation and exploring the changing dynamics of the process (global approach).

5.5. Gathering data for local LPV identification

For the scheduling parameters of the high-purity distillation column, an extensive analysis of the change of the local dynamical properties revealed that a 6×6 equidistant gridding of $\mathbb{P}$ on a log scale (in terms of the grid points $\mathcal{P} = \{\overline{p}^{(\tau)}\}_{\tau=1}^{N_{\text{loc}}=36}$) is required to adequately represent the variation of the local behavior on the operating region $\mathbb{P}$ (see [42]).

To collect measurement data at each operating point $\overline{p}^{(\tau)} \in \mathcal{P}$, the system is excited with a *random binary noise* signal $\breve{u}(k)$ (of amplitude 5×10^{-3} divided by the local DC gain w.r.t. u_1 and u_2 averaged over the output directions to ensure that the operating condition is only slightly perturbed $\approx \pm 0.01$ %) around the respective $(\overline{u}^{(\tau)}, \overline{y}^{(\tau)})$ steady-state values and the system response is recorded in terms of data sets $\mathcal{D}_\tau = \{(\breve{u}(k) + \overline{u}^{(\tau)}, y(k) + v(k))\}_{k=1}^{N}$ with length $N = 80$. This corresponds to 10 days of total experimentation time. Note that as the data is gathered in batches of approximately 6.6 h, it is not required to upset the column behavior continuously during the experimental campaign. In addition, the experiments can be performed under closed-loop circumstances, where an operational MPC controller and least-costly experiment design can minimize the loss in production (see [39,40]). However, to preserve the simplicity of the discussion, closed-loop identification is not considered in this work.

5.6. Gathering data for global LPV identification

To gather data for the global and "glocal" approaches (output interpolation with polynomial or RBF interpolation), a noisy data set $\mathcal{D}_{\text{global}}$ with $N = 2500$ samples and SNR = 15 is collected from the distillation column model. This corresponds to 8.7 days of experimentation time which can either be obtained in a continuous manner or in terms of batches of data, which can be then used to formulate the data matrices required for the discussed approaches.

[1] The signal-to-noise ratio is defined as $\text{SNR} := 10 \times \log_{10}\left(\frac{\|y-v\|_2^2}{\|v\|_2^2}\right)$.

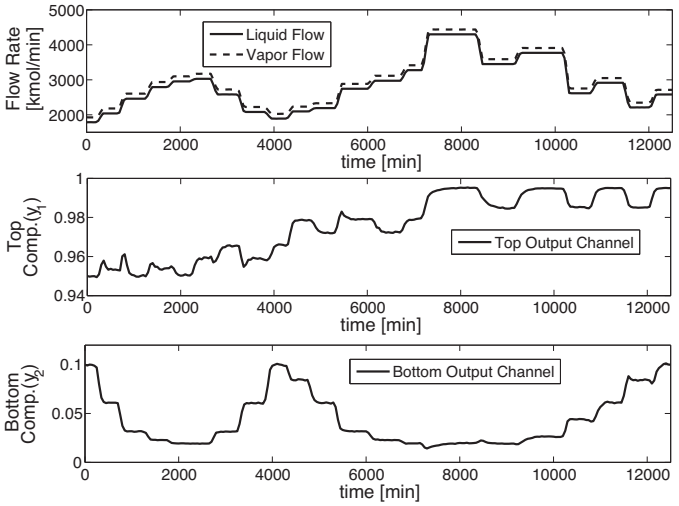

Fig. 2. Applied excitation signals and measured responses in $\mathcal{D}_{\text{global}}$.

In this study, $\mathcal{D}_{\text{global}}$ is continuously generated using an input signal, see Fig. 2, which is able to excite the transient behavior of the system over its operating regime. This input signal is a combination of a deterministic component added to a white noise with uniform distribution $\mathcal{U}(-0.5, 0.5)$ corresponding to an *standard deviation* (std) of $1/\sqrt{12}$.

6. Results and discussion

In the sequel, the previously introduced LPV data-driven modeling methodologies, summarized on Fig. 3, are tested on the identification problem of high-purity distillation column model discussed in Section 5.

6.1. Local identification

First the results of the local approaches are discussed together with the estimation of the local models, synthesis of basis functions for the OBF method and analysis of the obtained model performance based on validation data and frequency domain comparison of the captured local dynamics.

6.2. Estimation of local LTI models

Using the local data set $\mathcal{D}_\tau$ gathered from the system, LTI models of the local system behavior have been estimated using an *output-error* (OE) noise model structure. In particular, the models have been estimated using *canonical variate analysis* (CVA) based subspace identification (using the `n4sid` implementation in MATLAB) and the orders (n_a, n_b) of the OE models have been chosen to be (2, 2) based on BIC-based cross-correlation analysis conducted on the available local data sets [38].

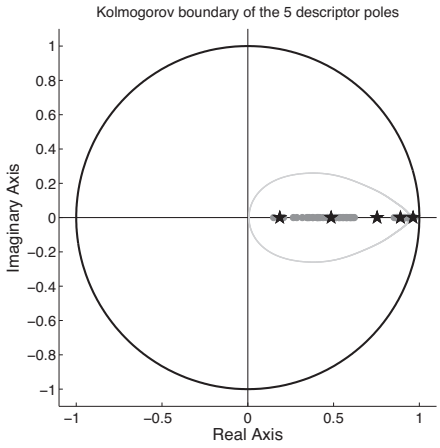

Fig. 4. Results of the FK*c*M clustering: sampled poles (∘), resulting pole locations of the basis (⋆), and Kolmogorov error bound (bold grey line).

The resulting models have been evaluated in terms of the *best fit rate* (BFR) [53]

$$\text{BFR} = 100\% \cdot \max\left(1 - \frac{||y(k) - \hat{y}(k)||_2}{||y(k) - \bar{y}||_2}, 0\right), \tag{31}$$

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 model calculated on validation data. The mean BFR of the identified local models, computed on a 80 samples long step-response (separate step test for each transfer channel with amplitude 5×10^{-3} divided by the local DC gain) data generated by the linearization of the column model at the corresponding operating points, has been (94.95 %, 93.47 %) for $u_1 \to (y_1, y_2)$ and (95.14 %, 93.33 %) for $u_2 \to (y_1, y_2)$.

6.2.1. Basis selection for the OBF scheme

To synthesize the basis functions for the OBF scheme, basis selection is performed by applying the *Fuzzy-Kolmogorov c-Max* (FKcM) approach [17] to the pole locations of the identified local models displayed with ∘ in Fig. 4. Based on trial and error, the number of basis has been selected to $n_e = 5$ to adequately represent the sampled pole locations. The optimized pole locations of the OBFs are depicted in Fig. 4 by ⋆. The typical error measure, the best achievable *Kolmogorov bound* [17] is given in grey in Fig. 4 and the resulting Kolmogorov n_e-width −25.09 dB computed w.r.t. the pole locations indicate that the basis functions exhibit an adequate convergence rate. The reasonable size of the Kolmogorov bound and

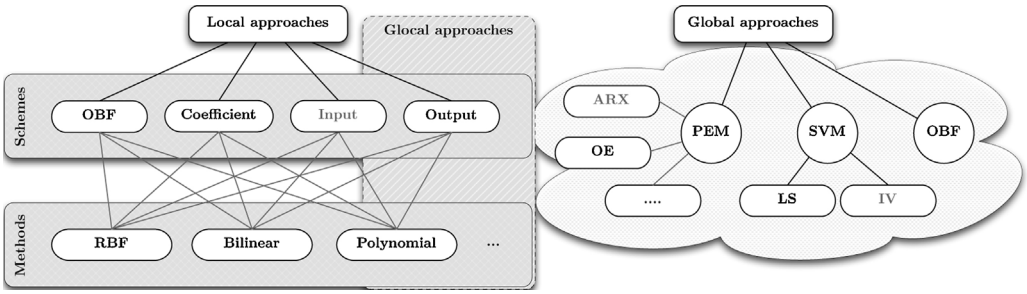

Fig. 3. Summary of the investigated IO and series expansion model structures based local and global LPV identification approaches (black) together with additionally mentioned methodologies (gray).

the n_e-width cost suggest that the OBF's are well tuned w.r.t. the estimated pole locations. The estimated poles cover a large dynamical range in the vicinity of the real axis with some poles (associated with models at the boundary of the operating regime) located close to the unit circle. Hence, the bound covers a large region in case of 5 basis poles, while the n_e-width cost describes a relatively small representation error. This implies that (i) the selected basis functions have a small representation error w.r.t. the local LTI models and (ii) further reduction of the error introduced by the series expansion would require a significantly larger number of basis poles, which only have a relatively small benefit for the local models identified at the boundary of the operating region. The resulting basis functions $\{\phi_i\}_{i=1}^{5}$ are used to form the model structure in terms of (19). Subsequently, the estimated local models are used to calculate the optimal local expansion coefficients $w_i^{(\tau)}$ via a Schur decomposition based method [44].

6.2.2. Investigated approaches

To investigate the performance of the above discussed local identification approaches for data-driven modeling of the distillation column, several combinations of the interpolation methods and schemes have been applied to the locally identified LTI models. The interpolation schemes, namely the coefficient and output approaches, are applied in combination with the RBF, polynomial and bilinear interpolation methods corresponding to 2×3 tested methods, while the OBF approach is tested only with the polynomial interpolation. In the coefficient and OBF scheme, the order of the polynomial interpolation has been chosen to be 3 by cross-validation based tuning, while in the output scheme polynomial order 4 has been found adequate. In the output and in the coefficient scheme, RBF interpolation has been accomplished with equidistantly placed basis functions centered at the interpolation points. Only in the output scheme, the RBF and polynomial interpolation methods have been implemented in the "glocal" sense by using linear regression and nonlinear optimization with $\mathcal{D}_{\text{global}}$ (using `fmincon` in MATLAB), respectively.

6.2.3. Assessment of the results

The resulting LPV models are compared in an extensive simulation study using an independent noise-free IO data set $\mathcal{D}_{\text{val}} = \{(u(k), y(k), p(k))\}_{k=1}^{N}$ with $N = 6 \times 10^3$ samples and varying operating conditions. The data set has been generated by the distillation column model using similar experimental conditions, but with a different varying p and a different realization of u than in the estimation data set(s) and without noise. Note that this data set is only used to evaluate the performance of the models and it is not required for identification. The simulation results are presented in Fig. 5 and in Table 2 where the *mean squared error* (MSE)

$$\text{MSE} = \|y(k) - \hat{y}(k)\|_2^2, \tag{32}$$

and the BFR of the simulation error (31), computed on $\mathcal{D}_{\text{val}}$, are listed.

Based on these results, the OBF scheme and the coefficient scheme with polynomial interpolation has produced the highest BFR on $\mathcal{D}_{\text{val}}$, while the output interpolation scheme has produced inferior results compared to the other schemes. It is interesting to note that if the data length was increased to contain five times more samples, the output interpolation scheme takes the lead in terms of the accuracy of the output fit. This can be explained by the "glocal" nature of this modeling approach as the interpolation parameters are optimized w.r.t. both the local and the transient behavior (using $\mathcal{D}_{\text{global}}$ data set). However, to exploit optimization of the interpolation function for modeling the transients, a rich and sufficiently long data set is needed, which certainly would be too costly to be obtained from the distillation column.

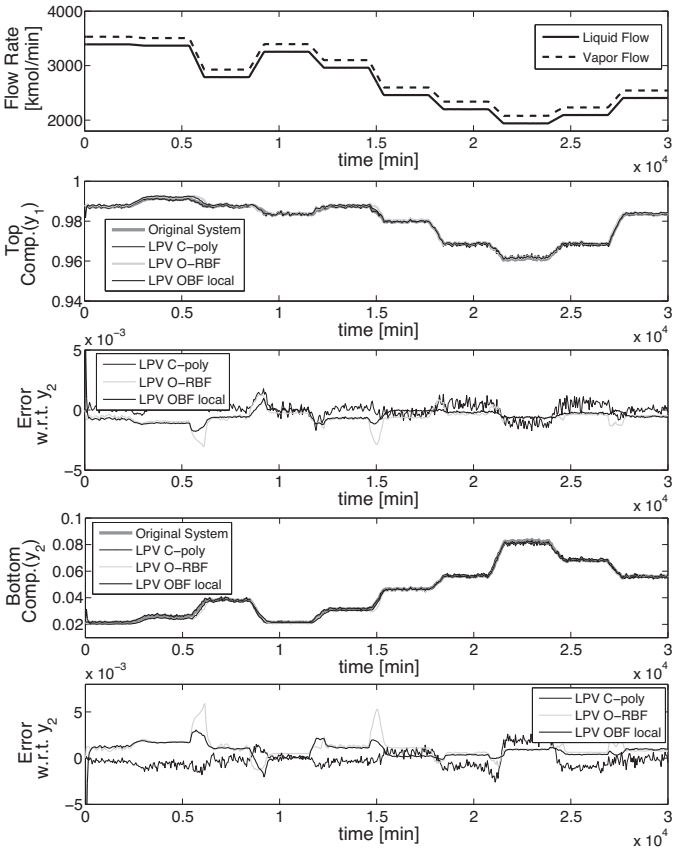

Fig. 5. The noise-free validation data $\mathcal{D}_{\text{val}}$ based comparison of the true system output and the simulated output of the LPV models estimated via the local approaches. The three best results are displayed: the output scheme with polynomial interpolation (LPV O-poly), the coefficient scheme with RBF interpolation (LPV C-RBF) and the local LPV-OBF scheme with polynomial interpolation (LPV-OBF local).

For the coefficient scheme based approaches, the polynomial interpolation method results in a better BFR compared to the other two methods. This indicates that smoother transition of the LPV model coefficients can be obtained using this approach. For the LPV OBF approach, it is observed that even without the information of global (inter-sample) behavior of the system, the LPV OBF model is able to produce a high fit ratio w.r.t. the original system. This is because the basis functions are optimally synthesized for the estimated local model poles and their possible variation due to changing operating conditions.

It is well-known in LPV identification that comparison of the simulated output response w.r.t. a particular trajectory of p may not fully reveal the quality of the identified models. A similar observation for the distillation column has been obtained in [35,27]. To systematically verify whether the models can generalize the local dynamics of the system w.r.t. other operating points, the $\mathcal{H}_2$ error

$$\varepsilon_{\bar{p}} \triangleq \|\underbrace{G_{o,\bar{p}}(z) - \hat{G}_{\bar{p}}(z)}_{\varepsilon_{\bar{p}}(z)}\|_2 = \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi} \text{tr}\left(\mathcal{E}_{\bar{p}}^{*}(e^{i\omega})\mathcal{E}_{\bar{p}}(e^{i\omega})\right) d\omega} \tag{33}$$

is computed on a dense 13×13 gridding of $\mathbb{P}$. (33) denotes the $\mathcal{H}_2$ norm of the difference of the transfer function $G_{o,\bar{p}}(z)$ of the linearized column model at operating condition $\bar{p} \in \mathbb{P}$ and the (frozen) transfer function $\hat{G}_{\bar{p}}(z)$ of the LPV model for a constant scheduling trajectory $p(k) \equiv \bar{p}$. The mean and maximum values of $\varepsilon_{\bar{p}}$ are given in Table 3.

Table 2
$\mathcal{D}_{val}$ based comparison of the simulated output error of the LPV models estimated via the local approaches.

		Output interpolation		Coefficient interpolation		Local LPV OBF	
		MSE	BFR	MSE	BFR	MSE	BFR
Polynomial interpolation	y_1	9.12×10^{-7}	89.88%	3.56×10^{-7}	93.68%	3.27×10^{-7}	93.94%
	y_2	2.13×10^{-6}	92.16%	9.10×10^{-7}	94.88%	9.78×10^{-7}	94.69%
RBF interpolation	y_1	7.10×10^{-7}	91.07%	4.61×10^{-7}	92.80%	–	–
	y_2	2.13×10^{-6}	92.17%	1.29×10^{-6}	93.90%	–	–
Bilinear interpolation	y_1	7.12×10^{-7}	91.05%	5.08×10^{-7}	92.45%	–	–
	y_2	2.13×10^{-6}	92.16%	1.44×10^{-6}	93.54%	–	–

Table 3 indicates that the coefficient interpolation and LPV-OBF based approaches result in a significantly smaller local $\mathcal{H}_2$ error in comparison with the output interpolation based approaches. In the polynomial output scheme, the error is even infinity which indicates a serious error in terms of an unstable $\mathcal{E}_{\overline{p}}$. This implies that the local properties of the model (e.g., stability) are better preserved by the OBF and coefficient schemes approaches. This explains that it is reasonable that the "glocal" approaches can achieve better BFR with enough data points, 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 in controller synthesis. On the other hand, the "glocal" method using trapezoidal approach enables preserving the local properties of the model fairly 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 [27,35] using the polynomial interpolation method. However, it is important to highlight that the problem is not with the polynomial method, but with the "glocal" nature of the applied scheme.

As the quality of the LPV models heavily depends on the quality of the local LTI models, one way to improve the identification results is to increase the model order in the local OE identification. However, raising the model order leads to a significant increase in the parameter variance of the identified models. This is because when operating the distillation column in the low purity region, the y_1 channel has significant higher-order (local) modes that disappear (become unidentifiable) in the high-purity region. The same holds for the significant higher-order (local) modes of y_2 in the high-purity region vanishing in the low purity region. These phenomena are highly undesirable, as the interpolation methods will be heavily influenced by the noise and, therefore, will result in random interpolated behavior of the identified 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. It also explains why the approach reported in [35] is applied successfully to other applications without this property. In the LPV-OBF identification approach, the flexible series expansion model structure is resilient to the changing local model order behavior, leading to promising results in both model simulation fitness and low local $\mathcal{H}_2$ error ratio. This makes the OBF scheme to be an attractive methodology for identifying the dynamics of the high-purity column model using a local approach.

6.3. Results of global LPV identification

Compared to the local approaches, the global approaches provide direct estimation of LPV models from data; hence it becomes possible to assess their stochastic performance. For this reason, the previously discussed estimators have been tested in a Monte-Carlo study using 100 global data sets generated with the same input and noise conditions as $\mathcal{D}_{global}$ described in Section 6.2.1 (same deterministic input trajectory, but new realizations of the random input component and the noise). In this study, the global approaches are applied to each of these data sets separately to identify MIMO LPV models.

6.3.1. Choices of model structures

Using cross-validation based order selection, a 4th-order model structure has been used in the LPV-OE, LS-SVM estimators (i.e., $n_a = 4$, $n_b = 4$) while the same basis functions found in the local case have been used in the global OBF approach (i.e., with $n_g = 5$). In the OBF scheme, the same basis functions $\{\phi_i\}_{i=1}^{5}$ as in the local approach are applied in the global estimator (25). In the parametric OE and OBF approaches, the parametrization of the coefficients of (21) is done using a 2-variable polynomial basis function of order n (i.e., $\Psi_s(p(k)) = p_1^i(k)p_2^j(k)$ such that $i, j \in \mathbb{I}_0^n$, $i \cdot j > 0$ and $s = i \cdot (n+1) + j$). This choice is motivated by the fact that it is known from the first-principles model that the system behavior can be well approximated by polynomial nonlinearities. Based on cross-validation based order selection, the polynomial order n has been chosen to be 4 in the OE case and 3 in the OBF case. It has been observed that a larger polynomial order quickly leads to an over-parametrization problem (significant increase of the estimation variance) and numerical problems in the OE case due to the nonlinear optimization.

6.3.2. Implementation of the global approaches

While implementation of the LPV-OBF and LS-SVM schemes follows from the solution of a simple least squares problem (besides the choice of some hyper parameters), the LPV-OE scheme requires the solution of a nonlinear minimization problem (see (22) and (23)). The latter minimization problem is implemented using a gradient descent based optimization with initial parameter estimate provided by an LPV-ARX estimator. To tune the hyper parameters $\{\gamma_i\}_{i=1}^{2}$ and $\{\sigma_{i,j}\}_{i=1,j=1}^{n_g,2}$ in the LPV LS-SVM approach, a data set $\mathcal{D}_{tune} = \{(u(k), y(k) + v(k), p_k)\}_{k=1}^{N}$, with $N = 2500$, independently generated from $\mathcal{D}_{global}$ and $\mathcal{D}_{val}$, has been

Table 3
The local $\mathcal{H}_2$ error of the estimated LPV models w.r.t. the transfer functions of the linearized and discretized system (computed on a dense 13×13 gridding of $\mathbb{P}$).

	Output interpolation		Coefficient interpolation		Local LPV OBF	
	Average	Max	Average	Max	Average	Max
Poly. int.	∞	∞	2.89×10^{-4}	1.04×10^{-3}	6.42×10^{-5}	1.39×10^{-4}
RBF int.	2.84×10^{-4}	1.04×10^{-3}	2.83×10^{-4}	1.04×10^{-3}	–	–
Bilin. int.	4.25×10^{-3}	1.58×10^{-2}	2.91×10^{-4}	1.05×10^{-3}	–	–

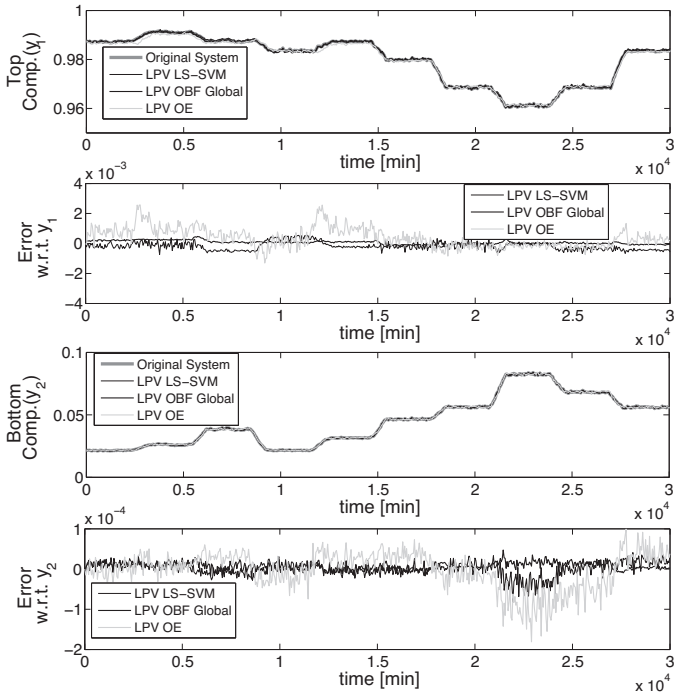

Fig. 6. Comparison of the simulated output of the LPV models estimated using the global approaches and the true system output based on the noise-free validation data $\mathcal{D}_{\text{val}}$.

recorded from the high-purity distillation column model using the same input and noise settings. To simplify the tuning problem in the LS-SVM approach, it is assumed that $\sigma_j = \sigma_{i,j}$ for all $i \in \mathbb{I}_1^{n_g}$, suggesting equal width for all kernels. Tuning of the regularization and hyper parameters is conducted by gridding the search space and, subsequently, using cross-validation with $\mathcal{D}_{\text{tune}}$. The resulting optimized values are $\gamma^{(1)} = 5.74 \times 10^4$, $\gamma^{(2)} = 1.27 \times 10^3$ and $\sigma_i^{(1)} = 0.58$, $\sigma_i^{(2)} = 0.21$ using equal width for all kernels.

6.3.3. Results of the Monte-Carlo study

To investigate the performance of the above discussed approaches, the simulation error of a randomly chosen identified LPV models in the 100 runs is computed w.r.t. $\mathcal{D}_{\text{val}}$, as depicted in Fig. 6. The mean and std of the BFR and the MSE measures of the simulation error w.r.t. $\mathcal{D}_{\text{val}}$ computed over the entire Monte-Carlo samples are presented in Table 4.

Table 4 indicates that the global approaches result in LPV models that exhibit high average BFR and low MSE if compared to the results of the local approaches. In addition, the std of the errors are reasonably low over the entire 100 Monte-Carlo runs, suggesting that the identified models have low variance. In the LPV LS-SVM approach, the std of the errors (for both BFR and MSE) is somewhat larger than that of the other two approaches, but the average is significantly better than in the OE case. This is explained by the ability of the non-parametric estimator to better represent the underlaying coefficient dependencies of the modeled process rather than using fixed polynomial basis. Furthermore, the increased std in this case results from the compensation of the bias due to the regularization term. This is important to note because under the OE noise setting, the basic LPV LS-SVM approach (which is based on the LPV-ARX model structure) always results in biased estimations, even though the 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 [54] or possible extension of results reported in [49,50]. However, due to the already adequate fit of the model on independently generated validation data, further refinement w.r.t. possible noise bias is not necessary (in case of insignificant noise bias, higher numerical complexity of the refined approaches diminishes the benefits of a more sophisticated handling of the noise).

Table 4
$\mathcal{D}_{\text{val}}$ based comparison of the simulated output error of the LPV models estimated via the global approaches.

		LPV LS-SVM	LPV OE	Global LPV-OBF
y_1	Avg. BFR	97.02%	92.49%	97.68%
	STD BFR	0.34%	0.23%	0.29%
	Avg. MSE	8.27×10^{-8}	5.03×10^{-7}	5.14×10^{-8}
	STD MSE	1.97×10^{-8}	5.79×10^{-8}	1.30×10^{-8}
y_2	Avg. BFR	99.89%	99.77%	99.90%
	STD BFR	0.02%	0.02%	0.03%
	Avg. MSE	5.14×10^{-10}	2.02×10^{-9}	3.79×10^{-10}
	STD MSE	2.02×10^{-10}	4.79×10^{-10}	2.43×10^{-10}

In the LPV-OE approach, it is observed that the initial model estimate provided by the ARX estimator needed only a slight refinement by the nonlinear optimization scheme to minimize the OE residual error in almost all the Monte-Carlo runs. Analysis of the estimated models revealed that the surface of the cost function is ill-conditioned (i.e., flat) and even the initial ARX estimation needs some regularization to better condition the estimation problem. This also explains the relatively small but erratic oscillations of the estimated model (see Fig. 6). Note that the regularization of the LS-SVM scheme inherently compensates for this ill-conditioning.

In the LPV-OBF approach, the basis functions are optimally selected w.r.t. the local dynamics of the process. Hence, the use of a data set that contains information about the transient behavior among the operating points results in a model with improved accuracy in comparison with the local approach. Among the global approaches investigated here, the OBF approach provides the best results in terms of both the average and std of the errors. In addition, the latter approach ensures that the identified model is stable, whereas the stability properties remain unconstrained in the OE and the LS-SVM estimators. However, this comes at the price of increased model order (5 compared to 4 for the other approaches). For the sake of fair comparison, it is important to note that the obtained SVM model can be used for prediction of the behavior of the system like in forecasting or even in model predictive control, while it is not directly useful to design controllers based on synthesis procedures like in optimal $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control. In that case, a closed form of the expanded functions in (30) needs to be obtained via interpolation, which might affect (in an unpredictable way) the overall model accuracy.

7. Conclusion

Local and global methodologies of *linear parameter-varying* identification have been reviewed and compared in the data-driven modeling of a high-purity distillation column. It has been shown that among the local approaches, the coefficient and series expansion (OBF) interpolation schemes demonstrate a superior performance in terms of the generalization capability of the identified models (describing the system behavior for all possible signal trajectories). The simulation results have revealed that, due to the varying order of the local models, the local LPV identification approaches may encounter significant difficulties when input–output or state-space model structures are used. On the other hand, it has been shown that the investigated global approaches, namely the least-squares support vector machine, the orthonormal basis functions based estimator and the prediction error minimization approach, have high-performance in describing

the input–output dynamics. Among the global approaches, the OBF method has demonstrated the most promising performance, although the difference in accuracy w.r.t. the other global schemes has not been significant.

Therefore, in terms of recommendations, the use of global approaches is suggested if the transient behavior of the system is needed to be adequately modeled (frequent variations of the operating conditions) while in case the operating condition is expected to vary slowly with respect to a large operating regime, then the use of local approaches is suggested which can exploit data sets gathered around steady state conditions. Note that, in the global case, the experimental campaign might be shorter, but the process operation is needed to be perturbed to capture the transient behavior, while in the local case, small perturbation based data can be gathered in batches when the process operates at particular given operating points. The properties of the OBF method in terms of dealing with locally changing model order, guaranteeing stability of the resulting description and well-conditionedness of the estimation problem make it a primary candidate for both local and global identification. However, in the global case, it also requires local experiments to synthesize the basis functions. The LS-SVM approach in the global case can be successfully applied to explore the required coefficient dependencies, which information can be used to improve parametrization of the OE and the OBF methods. In the local case, the coefficient scheme is also suggested due to its easy applicability and good performance. Based on the given study, use of "glocal" approaches for process systems is not suggested due their high experimentation costs and sensitive performance.

Acknowledgements

The authors thank Dr. Dario Piga for the fruitful discussions as well as his advice on software implementation of the LPV identification approaches.

References

[1] S. Skogestad, M. Morari, LV-control of a high-purity distillation column, Chemical Engineering Science 43 (1988) 33–48.
[2] M.V. Finco, W.L. Luyben, R.E. Polleck, Control of distillation columns with low relative volatilities, Industrial & Engineering Chemistry Research 28 (1989) 75–83.
[3] V. Gokhale, S. Hurowitz, J.B. Riggs, A comparison of advanced distillation control techniques for a propylene-propane splitter, Industrial & Engineering Chemistry Research 34 (1996) 4413–4419.
[4] H.G. Pandit, R.R. Rhinehart, J.B. Riggs, Experimental demonstration of nonlinear model-based control of a nonideal binary distillation column, in: Proc. of the American Control Conf., 2013, pp. 625–629.
[5] M. Barolo, G.B. Guarise, S.A. Rienzi, A. Trotta, Nonlinear model-based startup and operation control of a distillation column: an experimental study, Industrial & Engineering Chemistry Research 33 (1994) 3160–3167.
[6] G.R. Sriniwas, Y. Arkun, I.L. Chien, B.A. Ogunnaike, Nonlinear identification and control of a high-purity distillation column: a case study, Journal of Process Control 5 (1995) 149–162.
[7] M. Diehl, I. Uslu, R. Findeisen, S. Schwarzkopf, F. Allgower, H.G. Bock, T. Burner, E.D. Gilles, A. Kienle, J.P. Schloder, E. Stein, Real-time optimization for large scale processes: nonlinear model predictive control of a high purity distillation column, in: M. Grotschel, S.O. Krumke, J. Rambau (Eds.), Online Optimization of Large Scale Systems: State of the Art, Springer-Verlag, Berlin, 2001, pp. 363–382.
[8] A. Kumar, P. Daoutidis, Nonlinear model reduction and control for high-purity distillation columns, Industrial & Engineering Chemistry Research 42 (2003) 4495–4505.
[9] W. Rugh, J.S. Shamma, Research on gain scheduling, Automatica 36 (2000) 1401–1425.
[10] R. Tóth, Modeling and Identification of Linear Parameter-Varying Systems, Lecture Notes in Control and Information Sciences, vol. 403, Springer, Heidelberg, 2010.
[11] A. Kwiatkowski, H. Werner, J.P. Blath, A. Ali, M. Schultalbers, Linear parameter varying PID controller design for charge control of a spark-ignited engine, Control Engineering Practice 17 (2009) 1307–1317.
[12] T. Besselmann, J. Löfberg, M. Morari, Explicit model predictive control for linear parameter-varying systems, in: Proc. of the 47th IEEE Conf. on Decision and Control, Cancun, Mexico, 2013, pp. 848–3853.
[13] A. Packard, Gain scheduling via linear fractional transformations, Systems & Control Letters 22 (1994) 79–92.
[14] K. Zhou, J.C. Doyle, Essentials of Robust Control, Prentice-Hall, New Jersey, 1998.
[15] L. Giarré, D. Bauso, P. Falugi, B. Bamieh, LPV model identification for gain scheduling control: an application to rotating stall and surge control problem, Control Engineering Practice 14 (2006) 351–361.
[16] J.W. van Wingerden, M. Verhaegen, Subspace identification of bilinear and LPV systems for open- and closed-loop data, Automatica 45 (2009) 372–381.
[17] R. Tóth, P.S.C. Heuberger, P.M.J. Van den Hof, Asymptotically optimal orthonormal basis functions for LPV system identification, Automatica 45 (2009) 1359–1370.
[18] R. Tóth, P.S.C. Heuberger, P.M.J. Van den Hof, An LPV identification framework based on orthonormal basis functions, in: Proc. of the 15th IFAC Symposium on System Identification, Saint-Malo, France, 2013, pp. 1328–1333.
[19] Y. Zhu, 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, 2013, pp. 1674–1679.
[20] V. Laurain, M. Gilson, R. Tóth, H. Garnier, Refined instrumental variable methods for identification of LPV Box-Jenkins models, Automatica 46 (2010) 959–967.
[21] Y. Zhao, B. Huang, H. Su, J. Chu, Prediction error method for identification of LPV models, Journal of Process Control 22 (2012) 180–193.
[22] F. Casella, M. Lovera, LPV/LFT modelling and identification: overview, synergies and a case study, in: IEEE International Symposium on Computer-Aided Control System Design, San Antonio, TX, USA, 2013, pp. 852–857.
[23] K. Hsu, T.L. Vincent, 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, TX, USA, 2013, pp. 846–851.
[24] C.P.S. El-Dine, S.M. Hashemi, H. Werner, Black-box versus grey-box LPV identification to control a mechanical system, in: Proc. of the 51st IEEE Conf. on Decision and Control, Maui, HI, USA, 2013, pp. 5152–5157.
[25] A. Wills, B. Ninness, System identification of linear parameter varying state-space models, in: P.L. dos Santos, C. Novara, D. Rivera, J. Ramos, T. Perdicoúlis (Eds.), Linear Parameter-Varying System Identification: New Developments and Trends, World Scientific Publishing, Singapore, 2011, pp. 295–313.
[26] Y. Zhu, X. Xu, A method of LPV model identification for control, in: Proc. of the 17th IFAC World Congress, Seoul, Korea, 2013, pp. 5018–5024.
[27] J. Huang, G. Ji, Y. Zhu, P. van den Bosch, Identification of multi-model LPV models with two scheduling variables, Journal of Process Control 22 (2012) 1198–1208.
[28] R. Tóth, P.M.J. Van den Hof, J.H.A. Ludlage, P.S.C. Heuberger, Identification of nonlinear process models in an LPV framework, in: Proc. of the 9th International Symposium on Dynamics and Control of Process Systems, Leuven, Belgium, 2013.
[29] J. Caigny, R. Pintelon, J.F. Camino, J. Swevers, Interpolated modeling of LPV systems based on observability and controllability, in: Proc. 16th IFAC Symposium on System Identification, Brussels, Belgium, 2013, pp. 1773–1778.
[30] M. Steinbuch, R. van de Molengraft, A. van der Voort, Experimental modeling and LPV control of a motion system, in: Proc. of the American Control Conf., Denver, CO, USA, 2013, pp. 1374–1379.
[31] B. Bamieh, L. Giarré, Identification of linear parameter varying models, International Journal of Robust and Nonlinear Control 12 (2002) 841–853.
[32] R. Tóth, V. Laurain, W. Zheng, K. Poolla, Model structure learning: a support vector machine approach for LPV linear-regression models, in: Proc. of the 50th IEEE Conf. on Decision and Control, Orlando, FL, USA, 2013, pp. 3192–3197.
[33] J. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific, New Jersey, 2002.
[34] D. Kauven, D. Piga, R. Tóth, Adaptive Model Refinement – Advanced Autonomous Model-Based Operation of Industrial Process Systems, Technical Report FP7-257059-D2.2, RWTH Aachen, 2013.
[35] J. Huang, G. Ji, Y. Zhu, Some study on the identification of multi-model LPV models with two scheduling variables, in: Proc. of the the 16th IFAC Symposium on System Identification, Brussels, Belgium, 2012.
[36] X. Jin, B. Huang, D.S. Shook, Multiple model LPV approach to nonlinear process identification with EM algorithm, Journal of Process Control 21 (2011) 182–193.
[37] A.A. Khalate, X. Bombois, R. Tóth, 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, 2013, pp. 162–167.
[38] L. Ljung, System Identification, Theory for the User, 2nd ed., Prentice-Hall, New Jersey, 1999.
[39] X. Bombois, G. Scorletti, Design of least costly identification examples – the main philosophy accompanied by illustrative examples, Journal Européen des Systèmes Automatisés 46 (2012) 587–610.
[40] C.A. Larsson, M. Annergren, H. Hjalmarsson, C.R. Rojas, X. Bombois, A. Mesbah, P.E. Modén, Model predictive control with integrated experiment design for output error systems, in: Proc. of the European Control Conf., Zurich, Switzerland, 2013.
[41] K.E. Atkinson, An Introduction to Numerical Analysis, John Wiley and Sons, New York, 1989.
[42] A.A. Bachnas, Linear parameter-varying modelling of a high-purity distillation collumn, Delft University of Technology, 2012, Master's thesis.
[43] G. Mercère, M. Lovera, 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, FL, USA, 2013, pp. 783–790.

[44] P.S.C. Heuberger, P.M.J. Van den Hof, Bo Wahlberg, Modeling and Identification with Rational Orthonormal Basis Functions, Springer-Verlag, London, 2005.
[45] B.M. Ninness, F. Gustafsson, A unifying construction of orthonormal bases for system identification, IEEE Transactions on Automatic Control 42 (1997) 515–521.
[46] T. Oliveira e Silva, A n-width result for the generalized orthonormal basis function model, in: Proc. of the 13th IFAC World Congress, Sydney, Australia, 2013, pp. 375–380.
[47] V. Laurain, R. Tóth, W. Zheng, 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, 2013, pp. 1761–1766.
[48] R. Tóth, P.S.C. Heuberger, P.M.J. Van den Hof, Pediction error identification of LPV systems: present and beyond, in: J. Mohammadpour, C.W. Scherer (Eds.), Control of Linear Parameter Varying Systems with Applications, Springer, Heidelberg, 2012, pp. 27–60.
[49] J. Suykens, J. Vandewalle, Recurrent least squares support vector machines, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 47 (2000) 1109–1114.
[50] T. Falck, J. Suykens, 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, 2013, pp. 6389–6394.
[51] E. Jacobsen, P. Lundström, 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, 2013, pp. 242–248.
[52] J.B. Waller, Directionality and Nonlinearity-Challenges in Process Control, Åbo Akademi University, 2003, Ph.D. thesis.
[53] L. Ljung, System Identification Toolbox, for use with Matlab, The Mathworks Inc., Natick, MA, 2006.
[54] V. Laurain, W. Zheng, R. Tóth, Introducing instrumental variables in the LS-SVM based identification framework, in: Proc. of the 50th IEEE Conf. on Decision and Control, Orlando, FL, USA, 2013, pp. 3198–3203.