Preprints of the
15th IFAC Symposium on System Identification
Saint-Malo, France, July 6-8, 2009

An LPV identification Framework Based on Orthonormal Basis Functions ⋆

R. Tóth * **P. S. C. Heuberger** * **P. M. J. Van den Hof** *

* *Delft Center for Systems and Control (DCSC), Delft University of Technology Mekelweg 2, 2628 CD Delft, The Netherlands (e-mails: r.toth@ tudelft.nl, P.S.C.Heuberger@ tudelft.nl, P.M.J.vandenHof@tudelft.nl).*

Abstract: Describing nonlinear dynamic systems by Linear Parameter-Varying (LPV) models has become an attractive tool for control of complicated systems with regime-dependent (linear) behavior. For the identification of LPV models from experimental data a number of methods has been presented in the literature but a full picture of the underlying identification problem is still missing. In this contribution a solid system theoretic basis for the description of model structures for LPV systems is presented, together with a general approach to the LPV identification problem. Use is made of a series-expansion approach, employing orthogonal basis functions.

Keywords: Identification, Linear parameter-varying systems, Models, All-pass filters

1. INTRODUCTION

Industrial processes often exhibit parameter variations due to non-stationary or nonlinear behavior or dependence on external variables. For such processes, the theory of *Linear Parameter-Varying* (LPV) systems offers an attractive modeling framework. This system class can be seen as an extension of *Linear Time-Invariant* (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. As a result of this parameter variation, the LPV system class can describe both time-varying and nonlinear phenomena and it is particularly suited to model plants that have regime dependent, e.g. position dependent, linear behavior. Practical use of this framework is stimulated by the fact that LPV control design is well worked out, extending results of optimal and robust LTI control theory to nonlinear, time-varying plants (Packard (1994); Zhou and Doyle (1998)).

In the past two decades several methods have been developed for the identification of discrete-time LPV models from measured data (a few examples are Giarré et al. (2006); Felici et al. (2006); Tóth et al. (2007b)). Most of these approaches exploit the fact that an LPV system can be viewed as a collection of "local" models connected by scheduling dependent weighting functions. The identification approaches that are presented in the literature so far all take a particular starting point of a fixed model structure and identification method, usually chosen as a direct extension of the situation of LTI systems. A general theory for identification of LPV models is still missing. To a large extent, this is due to the fact that a structured framework for the description of this model class is lacking, including well-defined notions as model transformations, equivalence classes and canonical forms. As a result the model structures, commonly used in LPV identification methods, are generally not well defined or are limiting the representation capabilities of the resulting models considerably. In this paper the behavioral framework, originally developed for LTI systems (Willems (1991)), is used and extended to the LPV system class, to overcome the indicated limitations. On the basis of a solid system-theoretic definition of LPV systems, several LPV model structures are presented and consequences for their use in identification are discussed. Particular attention will be given to a series-expansions approach in terms of *Orthonormal Basis Functions* (OBFs). The question whether the scheduling signal has a static or dynamic effect on the system coefficients is an important issue that is discussed in detail.

In this paper we will restrict attention to *single input - single output* (SISO) systems, but all results carry over to the MIMO case in a straightforward way.

2. CONCEPTS AND NOTATION

In a discrete-time setting, LPV systems are commonly described in a *state-space* (SS) form:

$$x(k+1) = A(p(k))x(k) + B(p(k))u(k), \tag{1a}$$

$$y(k) = C(p(k))x(k) + D(p(k))u(k), \tag{1b}$$

and in a *input-output* (IO) model representation:

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

where y and u are the output, respectively input of the system, x is the state variable. The real-valued system coefficients (A, B, C, D) and $\{a_i, b_j\}_{j=0,\cdots n_b}^{i=1,\cdots n_a}$, are dependent on the scheduling variable $p : \mathbb{Z} \to \mathbb{P}$, where $\mathbb{P}$ is a closed subset of $\mathbb{R}^{n_\mathbb{P}}$. Note that these representations are equivalent with their LTI counterpart for a constant trajectory of p, i.e. $p(k) \equiv \bar{p}$ for all k, where $\bar{p} \in \mathbb{P}$.

A few observations should be added to these concepts:

⋆ This work was supported by the Dutch National Science Foundation (NWO).

- The coefficients in (1a-b,2) are assumed to be a static (nonlinear) function of the instantaneous value of p.
- The McMillan degree of the LTI systems associated with (1a-b,2) for $p(k) \equiv \bar{\mathrm{p}}$ can vary for each $\bar{\mathrm{p}} \in \mathbb{P}$.
- LPV systems are closely related to *Linear Time-Varying* (LTV) systems, with the difference that the knowledge about the time-varying behavior is limited by the fact that p is generally unknown in advance but online measurable during operation.
- Virtually all control design approaches are based on LPV-SS models, often with the assumption that the dependence of the matrices on p is affine or rational.

3. LPV MODELS REVISITED

3.1 Approaches to LPV identification

For the identification of LPV models, two major different approaches can be distinguished:

(1) Local approach
  - LTI models are identified in a number of operating points corresponding to constant scheduling signals, $p(k) \equiv \bar{\mathrm{p}}_i$, $i = 1, \cdots N_{\mathrm{loc}}$.
  - The resulting N_{loc} local linear models are interpolated (possibly by using data from an additional global experiment) to an LPV model.

(2) Global approach
  - Determine a global LPV model structure and an identification criterion.
  - Use data from a global experiment, i.e. with a varying p, to estimate an LPV model.

For the estimation step in these identification approaches both prediction-error methods and subspace methods are available (Giarré et al. (2006); Felici et al. (2006)). For interpolation various techniques and approaches have been introduced, varying from interpolation on pole estimates to the technique where each local (LTI) model is converted to a SS canonical form, and subsequently the coefficients in this model are interpolated.

This sketch of possible approaches directly leads to questions about the definition and selection of appropriate model structures. While many identification-related issues are up for further exploration, as e.g. experiment design, estimation accuracy, model validation, we will focus on the questions related to the use of different model structures.

3.2 Model structure considerations

As a first indication that there are theoretical problems involved with the current practice, let's consider a LPV-SS model representation

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0 & a_2(p(k)) \\ 1 & a_1(p(k)) \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} b_2(p(k)) \\ b_1(p(k)) \end{bmatrix} u(k),$$
$$y(k) = x_2(k).$$

This system can be written in an equivalent IO form:

$$y(k) = a_1(p(k-1))y(k-1) + a_2(p(k-2))y(k-2)$$
$$+ b_1(p(k-1))u(k-1) + b_2(p(k-2))u(k-2),$$

which is clearly not in the form defined by (2). This simple example shows that LPV-SS and IO representations are inequivalent if the coefficient dependence on p is restricted to be static. In order to obtain equivalence, it is necessary to allow a dynamic mapping between p and the coefficients, i.e. $\{A, B, C, D\}$ and $\{a_i, b_j\}$ should depend on (finite many) time-shifted instances of $p(k)$ (Tóth et al. (2007a)).

Based on the observation that LPV systems are closely related to LTV systems, it follows that for the definition of state-space equivalence transformations the concepts of the LTV theory should be used (Guidorzi and Diversi (2003)). It can be shown (see Tóth et al. (2007a)) that this results in transformation matrices and consequently also in state-space matrices that depend dynamically on p.

3.3 A behavioral approach

From the previous sections it can be concluded that the classical formulation of LPV models should be adapted in order to deal with dynamic scheduling dependence. In Tóth (2008), the behavioral framework, originally developed for LTI systems (Willems (1991)), is extended to deal with LPV systems. In this framework a *discrete time* (DT) *parameter-varying* (PV) system $\mathcal{S}$ is defined as

$$\mathcal{S} = (\mathbb{T}, \mathbb{P}, \mathbb{W}, \mathfrak{B}), \qquad (3)$$

where $\mathbb{T} = \mathbb{Z}$ is the DT time axis, $\mathbb{P}$ denotes the scheduling space, $\mathbb{W}$ is the signal space (the range of the system signals) with dimension $n_{\mathbb{W}}$ and $\mathfrak{B} \subset (\mathbb{P} \times \mathbb{W})^{\mathbb{T}}$ is the *behavior* of the system ($\mathbb{X}^{\mathbb{T}}$ stands for all maps from $\mathbb{T}$ to $\mathbb{X}$). $\mathfrak{B}$ defines trajectories of $(\mathbb{P} \times \mathbb{W})^{\mathbb{T}}$ that are possible according to the system model. Note that there is no prior distinction between inputs and outputs in this setting.

We also introduce the *projected scheduling behavior*

$$\mathfrak{B}_{\mathbb{P}} = \{p \in \mathbb{P}^{\mathbb{T}} \mid \exists w \in \mathbb{W}^{\mathbb{T}} \text{ s.t. } (w, p) \in \mathfrak{B}\}, \qquad (4)$$

and for a given $p \in \mathfrak{B}_{\mathbb{P}}$, we define the *projected behavior*

$$\mathfrak{B}_p = \{w \in \mathbb{W}^{\mathbb{T}} \mid (w, p) \in \mathfrak{B}\}. \qquad (5)$$

With these concepts we can define LPV systems as follows:

Definition 1. (DT-LPV system). Let $\mathbb{T} = \mathbb{Z}$. The parameter-varying system $\mathcal{S}$ is called LPV, if the following conditions are satisfied:

- $\mathbb{W}$ is a vector-space and $\mathfrak{B}_p$ is a linear subspace of $\mathbb{W}^{\mathbb{T}}$ for all $p \in \mathfrak{B}_{\mathbb{P}}$ (linearity).
- For any $(w, p) \in \mathfrak{B}$ and any $\tau \in \mathbb{T}$, it holds that $(w(\cdot + \tau), p(\cdot + \tau)) \in \mathfrak{B}$, in other words $q^\tau \mathfrak{B} = \mathfrak{B}$ (time-invariance), with q the forward time-shift operator.

Note that in terms of Definition 1, for a constant scheduling trajectory, $p(k) \equiv \bar{\mathrm{p}}$, the associated system $\mathcal{G} = (\mathbb{T}, \mathbb{W}, \mathfrak{B}_{\bar{\mathrm{p}}})$ is an LTI system. In a next step, the behavior $\mathfrak{B}$ of LPV systems has to be specified in terms of mathematical representations. The coefficients in these representations will become (nonlinear) functions of p. In order to describe this functional dependence of a single real-valued coefficient, we employ functions $r : \mathbb{R}^n \to \mathbb{R}$ that are considered to be in the set $\mathcal{R} = \cup_{n \in \mathbb{N}} \mathcal{R}_n$, where $\mathcal{R}_n$ is the set of essentially [1] n-dimensional real-meromorphic functions (being a quotient of analytical functions). This function specifies how the resulting coefficient is dependent on n variables, that are selected -in a unique ordering- from the set $\{q^i p_j\}_{j=1,\cdots,n_{\mathbb{P}}}^{i \in \mathbb{Z}}$. In order to specify the (time-varying) coefficient we introduce the operator

$$\diamond : (\mathcal{R}, \mathfrak{B}_{\mathbb{P}}) \to \mathbb{R}^{\mathbb{Z}} \quad \text{defined by} \quad r \diamond p = r\left(p, qp, q^{-1}p, \ldots\right).$$

[1] In the sense that $r(\mathtt{x}_1, \cdots, \mathtt{x}_n)$ does depend on $\mathtt{x}_n$.

Thus the value of a (p-dependent) coefficient r in an LPV system representation at time k is given by $(r \diamond p)(k)$.

In the sequel the (time-varying) coefficient sequence $(r \diamond p)$ will be used to operate on a signal w (like $a_i(p)$ in (2)), giving the varying coefficient sequence of the representations. In this respect an important property of the $\diamond$ operation is that multiplication with the shift operator q is not commutative, in other words

$$q(r \diamond p)w \neq (r \diamond p)qw.$$

To handle this multiplication, for $r \in \mathcal{R}$ we define the shift operations $\overrightarrow{r}$, $\overleftarrow{r}$ as

$$\begin{aligned} \overrightarrow{r} &= r' \in \mathcal{R} \quad \text{s.t.} \quad r' \diamond p = r \diamond (qp), \\ \overleftarrow{r} &= r'' \in \mathcal{R} \quad \text{s.t.} \quad r'' \diamond p = r \diamond (q^{-1}p), \end{aligned}$$

for $p \in (\mathbb{R}^{n_{\mathbb{P}}})^{\mathbb{Z}}$. With these notions we can write

$$q(r \diamond p)w = (\overrightarrow{r} \diamond p)qw \quad \text{and} \quad q^{-1}(r \diamond p)w = (\overleftarrow{r} \diamond p)q^{-1}w.$$

Next we introduce difference equations with varying coefficients as the representation of the behavior. Let $\mathcal{R}[\xi]^{n_r \times n_w}$ be the ring of matrix polynomials in the indeterminant ξ and with coefficients in $\mathcal{R}^{n_r \times n_{\mathbb{W}}}$, then a PV difference equation is defined as follows:

$$(R(q) \diamond p)w := \sum_{i=0}^{n_\xi} (r_i \diamond p) q^i w = 0, \tag{6}$$

where $R(q) = \sum_{i=0}^{n_\xi} r_i q^i$, $n_\xi = \deg(R)$, and $r_i \in \mathcal{R}^{n_r \times n_{\mathbb{W}}}$. In this notation q operates on the signal w, while the operation $\diamond$ takes care of the time/scheduling-dependent coefficient sequence. Note that as the indeterminant ξ is associated with q, multiplication with ξ is non commutative on $\mathcal{R}[\xi]^{n_r \times n_{\mathbb{W}}}$, i.e. $\xi r = \overrightarrow{r}\xi$ and $r\xi = \xi\overleftarrow{r}$.

3.4 LPV kernel representation

Using the previously introduced concepts, we can define the *kernel representation* (KR) of an LPV system in the form of (6). More precisely, we call (6) a representation of the LPV system $\mathcal{S} = (\mathbb{Z}, \mathbb{R}^{n_{\mathbb{P}}}, \mathbb{R}^{n_{\mathbb{W}}}, \mathfrak{B})$ with scheduling signal p and signals w if

$$\mathfrak{B} = \{(w,p) \in (\mathbb{R}^{n_{\mathbb{W}}} \times \mathbb{R}^{n_{\mathbb{P}}})^{\mathbb{Z}} \mid (R(q) \diamond p)\, w = 0\}. \tag{7}$$

In the sequel we only consider LPV systems, whose behavior can be described by (7). An important property of these systems is that they have a kernel representation where R has full row rank (Tóth (2008)).

3.5 IO representation

For practical applications, a partitioning of the signals w into input signals $u \in (\mathbb{R}^{n_{\mathbb{U}}})^{\mathbb{Z}}$ and output signals $y \in (\mathbb{R}^{n_{\mathbb{Y}}})^{\mathbb{Z}}$, i.e. $w = \mathrm{col}(u, y)$, is often convenient. Note that this partitioning is not trivial and can neither be chosen freely. For details see (Willems (1991); Tóth (2008)). Using an IO partition, the *IO representation* of $\mathcal{S}$ is defined as

$$(R_{\mathrm{y}}(q) \diamond p)\, y = (R_{\mathrm{u}}(q) \diamond p)\, u, \tag{8}$$

where R_{u} and R_{y} are matrix polynomials with meromorphic coefficients, and where R_{y} is full row rank and $\deg(R_{\mathrm{y}}) \geq \deg(R_{\mathrm{u}})$. Using the same type of decomposition as in (6), we derive the following form of an IO representation

$$\sum_{i=0}^{n_a} (a_i \diamond p)\, q^i y = \sum_{j=0}^{n_b} (b_j \diamond p)\, q^j u. \tag{9}$$

It is apparent that (9) is the 'dynamic-dependent' counterpart of (2).

3.6 State-space representation

Without going into details about the definition of so called latent variables, we formulate the discrete-time SS representation, based on an IO partition (u, y), as a first-order PV difference equation system in the latent variable $x : \mathbb{Z} \to \mathbb{X}$:

$$qx = (A \diamond p)x + (B \diamond p)u, \tag{10a}$$

$$y = (C \diamond p)x + (D \diamond p)u, \tag{10b}$$

where $\mathbb{X} \subseteq \mathbb{R}^{n_{\mathbb{X}}}$ is the state space and (A, B, C, D) are matrices of appropriate dimensions with their entries being meromorphic functions in $\mathcal{R}$. It is apparent that (10a-b) are the 'dynamic-dependent' counterparts of (1a-b).

3.7 Properties

Using the behavioral framework, it is now possible to consider equivalence of behaviors, and related equivalent transformations between the different LPV system representations. For details see (Tóth (2008)). In this framework, transformations between different representations as well as state transformations into a different coordinate system generally involve dynamically dependent relations.

4. AN ORTHONORMAL BASIS FUNCTIONS APPROACH

4.1 Series-expansion representations

In this section we introduce a series-expansion type of model structure for LPV systems, via the concept of OBFs (Heuberger et al. (2005)). A major motivation is the linear-in-the-parameters property of these structures, which is beneficial in prediction-error identification. A second merit of these structures is that they allow a relatively simple interpolation of local LTI models with varying McMillan degree. Furthermore it was shown in Boyd and Chua (1985) for nonlinear Wiener models (LTI system followed by a static nonlinearity) that, if the LTI part is an OBF filter bank, then such models are general approximators of nonlinear systems with fading memory.

The transfer function $F \in \mathcal{H}_2$ of a (local) linear model can be written as

$$F(z) = D + \sum_{i=1}^{\infty} w_i \phi_i(z), \tag{11}$$

where $\{\phi_i\}_{i=1}^{\infty}$ is a basis for $\mathcal{H}_2$ and $w_i \in \mathbb{R}$. In the theory of *Generalized Orthonormal Basis Functions*, the functions $\phi_i(z)$ are generated by applying a *Gram-Schmidt orthonormalization* to the sequence of functions

$$\frac{1}{z - \lambda_1}, \cdots \frac{1}{z - \lambda_{n_{\mathrm{b}}}}, \frac{1}{(z - \lambda_1)^2}, \cdots$$

with stable pole locations $\lambda_1, \cdots \lambda_{n_{\mathrm{b}}}$. The choice of these *basis poles* determines the rate of convergence of the series expansion (11).

An alternative derivation is based on a balanced realization $\{A_{\mathrm{b}}, B_{\mathrm{b}}, C_{\mathrm{b}}, D_{\mathrm{b}}\}$ of the inner function

$$G_{\mathrm{b}}(z) = \prod_{i=1}^{n_{\mathrm{b}}} \frac{1 - z\lambda_i^*}{z - \lambda_i}, \tag{12}$$

where the functions $\{\phi_i(z)\}$ are the scalar elements of the vector functions

$$(zI - A_\mathrm{b})^{-1} B_\mathrm{b} G_\mathrm{b}^j(z), \; j = 1, 2, \cdots. \quad (13)$$

By using a truncated expansion in (11) an attractive model structure for LTI identification results, with a well worked-out theory in terms of variance and bias expressions. The series expansion (11) can be extended to LPV systems, such that for a given basis $\{\phi_i\}_{i=1}^{\infty}$ and a specific IO-partitioning (u, y), an LPV system can be written as

$$y = (D \diamond p)u + \sum_{i=1}^{\infty} (w_i \diamond p)\, \phi_i(q) u, \quad (14)$$

with $w_i \in \mathcal{R}$. An obvious approach is to use a truncated expansion, i.e. with $\{\phi_i\}_{i=1}^{n}$, as a model-structure candidate for LPV identification. Note that these expansions are formulated in the time domain (using the shift operator q), as there exist no frequency-domain expressions for LPV systems. Similar to the LTI case, this structure is linear in the coefficients $\{w_i\}_{i=1}^{n}$. An important question that arises is wether the basis functions ϕ_i can be chosen such that a fast rate of convergence can be accomplished for all possible scheduling trajectories p. Note that the representation (14) is equivalent with a state-space description (10a-b), where the matrices A and B are independent of p.

4.2 Basis selection

In order to select a basis, it is obviously required to obtain knowledge about the system to be modeled. For the LTI case it is well-known that an optimal basis can be chosen using knowledge about the system poles. It can be shown that the same property holds for LPV systems, where knowledge of the poles of all possible local LTI models is required. In practice this knowledge is generally not available and one has to resort to limited prior-information resources, such as expert knowledge or preliminary identification experiments.

A possible scheme for the basis selection is given by the following steps:

(1) Identify a number of local linear models in several different operating regimes $\bar{\mathrm{p}}_i$, i.e. using data with a constant scheduling signal $p(k) \equiv \bar{p}_i$.
(2) Cluster the poles in groups of the complex plane and find optimal cluster centers (these centers will be used as basis poles)

In this procedure use is made of minimization of a distance measure, which is relevant for the worst-case approximation error of the representation (14). This scheme is motivated by the extension of the classical *Kolmogorov n-width* result of Pinkus (1985) to OBFs, as obtained by Oliveira e Silva (1996). These results states that for a given LTI inner function G_b, the first n OBF's generated by G_b (see Section 4.1) are optimal in the n-width sense for the set of LTI systems having poles in the region

$$\{z \in \mathbb{D} \mid |G_\mathrm{b}(z^{-1})| \le \rho\}.$$

Here ρ is the rate of convergence in the series expansion, $\mathbb{D}$ is the unit disc, and n is a multiple of the number of basis poles n_b. For the basis-selection problem we are dealing with the inverse problem, i.e. given a region of poles Ω, approximate this region as

$$\Omega \approx \Omega(\Xi, \rho) = \{z \in \mathbb{D} \mid G_\mathrm{b}(z^{-1}) \le \rho\}. \quad (15)$$

The n optimal OBF poles $\Xi = \{\lambda_1, \cdots, \lambda_n\}$ are therefore obtained by solving the following *Kolmogorov measure minimization* problem,

$$\min_{\Xi \subset \mathbb{D}} \rho = \min_{\Xi \subset \mathbb{D}} \max_{z \in \Omega} \left| G_\mathrm{b}(z^{-1}) \right| = \min_{\Xi \subset \mathbb{D}} \max_{z \in \Omega} \left| \prod_{k=1}^{n} \frac{1 - z\lambda_k^*}{z - \lambda_k} \right|$$

As stated above, in a practical situation the knowledge about the pole region Ω is limited. In the next section we present an approach to obtain a simultaneous solution for the problems of reconstructing Ω from experimental data and the Kolmogorov measure minimization problem.

4.3 A fuzzy clustering approach

Objective-function-based fuzzy clustering algorithms, such as *Fuzzy c-Mean clustering* (FcM), have been used in a wide collection of applications (Bezdek (1981)). In this section we give the extension of FcM to the so-called *Fuzzy-Kolmogorov c-Max* (FKcM) algorithm, which enables the determination of the region Ω on the basis of observed poles with membership based, overlapping areas. We assume that we are given a set of poles $Z = \{z_1, \cdots, z_N\}$.

Let c be the number of clusters, that we wish to discern and let $v_i \in \mathbb{D}$ denote the *cluster center* of the i-th cluster. We define membership functions $\mu_i : \mathbb{D} \to [0,\, 1]$, that determine for each $z \in \mathbb{D}$ the 'degree of membership' to cluster i. By using a *threshold value* ε, we obtain a set

$$\Omega = \{z \in \mathbb{D} \mid \exists i \in \{1, \ldots, c\},\ \mu_i(z) \ge \varepsilon\}. \quad (16)$$

With these preliminaries we can now formulate the problem we consider:

Problem 2. For a given c, find a region Ω, as described above, such that Ω contains Z, and such that the OBFs, with poles in the cluster centers $\{v_i\}_{i=1}^{c}$, are optimal in the Kolmogorov n-width sense, $n = c$, with respect to Ω and with a minimal corresponding decay rate ρ.

To measure dissimilarity of Z with respect to each cluster, we introduce distances $d_{ik} = \kappa(v_i, z_k)$ between v_i and z_k, where κ is the *Kolmogorov metric* on $\mathbb{D}$, defined by

$$\kappa(\mathtt{x}, \mathtt{y}) = \left| \frac{\mathtt{x} - \mathtt{y}}{1 - \mathtt{x}^* \mathtt{y}} \right|. \quad (17)$$

Analogously we define $\mu_{ik} = \mu_i(z_k)$ and we regulate the membership functions by the so-called *fuzzy constraints*:

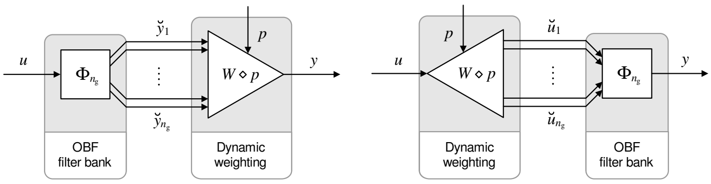
$$\sum_{i=1}^{c} \mu_{ik} = 1 \quad \text{and} \quad 0 < \sum_{k=1}^{N} \mu_{ik} < N.$$

Fuzzy clustering can be viewed as the minimization of the FcM-*functional* (Bezdek (1981)), J_m, which in the FKcM case can be formulated as

$$J_m = \max_{1 \le k \le N} \sum_{i=1}^{c} \mu_{ik}^m d_{ik}. \quad (18)$$

Here the design parameter $m \in (1, \infty)$ determines the fuzziness of the resulting partition. Note that J_m is a function of the membership data μ_{ik} and the cluster centers v_i. It can be observed, that (18) corresponds to a *worst-case (max) sum-of-error* criterion, contrary to the *mean-squared-error* (MSE) criterion of the original FcM.

The crucial property of (18) is that it can be shown (Tóth et al. (2009)) that for large values of m minimization of J_m is equivalent to the Kolmogorov measure minimization problem. For details as well as a detailed description of



(a) W-LPV OBF model (b) H-LPV OBF

Fig. 1. IO signal flow graph of (a) the W-LPV OBF model (see (19)) and (b) the H-LPV OBF model (see (20)).

the optimization algorithm see Tóth et al. (2009). In Tóth (2008), also the robust extension of the basis selection is discussed, i.e. not only pole estimates are considered, but also their corresponding uncertainty regions.

For the determination of the actual number of clusters in these algorithms, so-called adaptive cluster merging is applied. Starting from a relatively large initial number of clusters (typically around $N/2$), the adaptive merging steers the algorithm towards the natural number of groups that can be observed in the data, assisting the selection of the required number of OBFs in (14).

4.4 OBF-based model structures

We assume that the basis-selection step has been completed and we are given a set of n_g basis functions $\{\phi_i(z)\}_{i=1}^{n_\mathrm{g}}$ with good approximation properties for the set of local LTI behaviors corresponding to constant scheduling signals. In the next step we construct model structures for the identification of an LPV system $\mathcal{S}$. To keep the notation simple, we restrict attention to strictly proper models ($D = 0$ in (14)). The IO dynamics of the LPV model can now be written as

$$y(k) = \sum_{i=1}^{n_\mathrm{g}} \left(w_i \diamond p\right)(k)\phi_i(q)u(k). \tag{19}$$

Introduce $\Phi_{n_\mathrm{g}} = [\, \phi_1 \, \dots \, \phi_{n_\mathrm{g}} \,]^\top$ and $W = [\, w_1 \, \dots \, w_{n_\mathrm{g}} \,]^\top$ as shorthand notations. Then (19) can be visualized as in Fig. 1a, where $\breve{y}_i(k) = \phi_i(q)u(k)$. Because of the close resemblance of this structure to classical Wiener models this model structure is referred to as a Wiener LPV OBF (W-LPV OBF) model. A closely related model structure, depicted in Fig. 1b, is the Hammerstein LPV OBF model (H-LPV OBF), that results from the description

$$y(k) = \sum_{i=1}^{n_\mathrm{g}} \phi_i(q)\left(w_i \diamond p\right)(k)u(k). \tag{20}$$

This latter structure can be motivated from the LTI series expansion (11), by changing the order of the arguments. This change has no effect in the LTI case, but certainly results in a different LPV structure. In the sequel we will restrict attention to the Wiener model structure. Furthermore we assume that the coefficient functions w_i have only a static dependence on p, so we can write $\left(w_i \diamond p\right)(k) = w_i(p(k))$ in (19). As stated before, the W-LPV OBF structure has a direct SS realization:

$$qx = Ax + Bu \tag{21a}$$

$$y = \left(W \diamond p\right)x, \tag{21b}$$

where the constant matrices A and B are completely determined by $\{\phi_i\}_{i=1}^{n_\mathrm{g}}$. This illustrates that the dependence on p is only present in the output equation, with the result that the assumption of static dependence is much less restrictive than in (10a-b). With respect to the actual estimation with these model structures we again distinguish a local and a global approach (Tóth et al. (2007b)):

Local estimation approach: This approach is based on a number N_loc of "local" experiments, i.e. data collection with a constant scheduling signal $p(k) \equiv \bar{\mathsf{p}}_j \in \mathbb{P}$, resulting in data sequences $\{u_j(k), y_j(k)\}$ for $j = 1, \dots, N_\mathrm{loc}$. Based on these data, N_loc LTI-OBF models are estimated using a standard least-squares criterion in a one-step-ahead prediction error setting with *Output Error* (OE) type of noise model:

$$\varepsilon(k) = y_j(k) - \sum_{i=1}^{n_\mathrm{g}} \theta_{ij}\phi_i(q)u_j(k), \tag{22}$$

where $\{\theta_{ij}\}$ are real-valued coefficients. Note that – under the condition that the data are informative – there exists a unique analytic solution to this estimation problem. These estimated coefficients can now be considered as "samples" of the function $w_i(p)$, in the sense that $w_i(\bar{\mathsf{p}}_j) = \theta_{ij}$. As a second step we use interpolation to obtain estimates of the function $w_i(p)$, for instance by assuming a polynomial dependence of w_i on p, or by making use of splines etc.

Global estimation approach: For this approach we need to assume a specific functional dependence of the functions w_i on p and we propose to use a linear parametrization for this purpose, such as a polynomial dependence

$$w_i(p(k)) = \theta_{i0} + \theta_{i1}p(k) + \ldots + \theta_{in_\mathrm{r}}p^{n_\mathrm{r}}(k).$$

Here we assumed for simplicity that p is a one-dimensional signal. A global data set $\{u(k), y(k), p(k)\}_{k=1}^{N}$ is collected, which is assumed to be informative with respect to (19). It is straightforward that – using a least-squares criterion– a unique analytic solution can be obtained for the parameters θ_{ij}. Note that the restriction to static dependence can be relaxed for the global approach by allowing w_i to depend on time-shifts of $p(k)$ as well. Because of the postulated OBF structure, both approaches will always result in asymptotically stable models.

4.5 Approximation of dynamic dependence

To alleviate the restrictions caused by the assumption of static dependence in the suggested model structures, extensions for these structures were proposed in Tóth et al. (2008). Here we only consider the Wiener case. The idea is still to use weighting functions with static dependence, but with the introduction of an additional feedback loop around each basis component with a gain incorporating also static dependence. For this new model structure what we call Wiener Feedback (WF-LPV OBF) model, it is apparent that by setting the feedback gains to zero, the previous structure result. This immediately indicates an increase in the representation capability. The W-LPV OBF can be represented in SS form by

$$qx = \left(A - BV(p)C\right)x + Bu \tag{23a}$$

$$y = W(p)Cx, \tag{23b}$$

where the constant matrices A, B and C are again completely determined by $\{\phi_i\}_{i=1}^{n}$ and $V(p)$ is a diagonal matrix. For the estimation of the functions W and V again a linear parametrization using polynomials or spline functions is suggested. To overcome the nonlinear optimization problem associated with the parallel estimation

Table 1. Validation results of 100 identification experiments with the Wiener (W), Hammerstein (H), and the Wiener Feedback (WF) model structures.

	MSE (dB)			VAF (%)		
SNR	W	H	WF	W	H	WF
no noise	-34.96	-30.73	-39.75	99.00	97.47	99.42
35 dB	-34.77	-30.42	-39.17	98.99	97.21	99.39
20 dB	-32.75	-28.96	-35.01	98.71	96.94	99.00
10 dB	-31.81	-27.72	-32.38	98.19	96.01	98.59

of the parameters of W and V, the approach uses a separable least squares optimization scheme. For algorithmic details see Tóth et al. (2008). It should be noted that the better representation capability comes at a price. First of all, there is no longer an analytic solution available. Secondly, there is no guarantee that the resulting models are asymptotically stable.

5. EXAMPLE

To illustrate the applicability of the introduced model structures, we consider the following asymptotically stable LPV system $\mathcal{S}$, given in LPV-IO form:

$$a_0(p)\, y = b_1(p)\, q^{-1} u - \sum_{i=1}^{5} a_i(p)\, q^{-i} y,$$

where $p : \mathbb{Z} \to [0.6, 0.8]$ and $a_0(p) = 0.58 - 0.1p$, $a_1(p) = -\frac{511}{860} - \frac{48}{215}p^2 + 0.3(\cos(p) - \sin(p))$, $a_2(p) = \frac{61}{110} - 0.2\sin(p)$, $a_3(p) = -\frac{23}{85} + 0.2\sin(p)$, $a_4(p) = \frac{12}{125} - 0.1\sin(p)$, $a_5(p) = -0.003$, $b_1(p) = \cos(p)$.

Using 8 basis functions, obtained through the FK*c*M algorithm (see Tóth et al. (2009, 2007b) for details) and a 2^{nd}-order polynomial-based parametrization of the coefficients W and a 3^{rd}-order polynomial-based parametrization of V, identification of $\mathcal{S}$ with the global approach has been carried out, with the W-LPV, H-LPV, and WF-LPV OBF model structures. Each experiment has been repeated 100 times with different realizations of input, scheduling signals in 4 different white output noise setting. See Table 1 for the results in terms of average MSE and VAF (Variance Accounted For). The resulting *signal-to-noise ratio* (SNR) of the noise settings is also indicated. As expected, all approaches identified the system with adequate MSE and VAF even in case of extremely heavy output noise, which underlines the effectiveness of the proposed identification philosophy. As expected, the W-LPV and H-LPV structures based on coefficients with static dependence could not fully cope with the variations in the parameters $\{a_i\}_{i=0}^{5}$. However, the W-LPV identification provided better results than the H-LPV structure due to the different approximation capabilities of these models (see Tóth et al. (2007b)). For all cases, the WF-LPV model provided better estimates than the pure static-dependence based model estimates. This clearly shows the improvement in the approximation capability due to the approximation of dynamic dependence with feedback-based weighting.

6. CONCLUSION

Using a solid system theoretic definition of LPV systems in terms of behaviors, LPV model representations are presented and brought into a unifying framework. Real-valued meromorphic functions are used to specify dynamic dependence of the system coefficients on the scheduling signal. A series expansion approach is presented for modeling LPV systems, including an optimization procedure for selecting optimal basis functions. The series-expansion models can be used in both local and global identification methods, and their use is illustrated in an example.

REFERENCES

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press.

Boyd, S. and Chua, L.O. (1985). Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. on Circuits and Systems*, 32(11), 1150–1161.

Felici, F., van Wingerden, J.W., and Verhaegen, M. (2006). Subspace identification of MIMO LPV systems using a periodic scheduling sequence. *Automatica*, 43(10), 1684–1697.

Giarré, L., Bauso, D., Falugi, P., and Bamieh, B. (2006). LPV model identification for gain scheduling control: An application to rotating stall and surge control problem. *Control Engineering Practice*, 14, 351–361.

Guidorzi, R. and Diversi, R. (2003). Minimal representations of MIMO time-varying systems and realization of cyclostationary models. *Automatica*, 39(11), 1903–1914.

Heuberger, P.S.C., Van den Hof, P.M.J., and Bo Wahlberg (2005). *Modeling and Identification with Rational Orthonormal Basis Functions*. Springer-Verlag.

Oliveira e Silva, T. (1996). A *n*-width result for the generalized orthonormal basis function model. In *Proc. of the 13th IFAC World Congress*, 375–380. Sydney, Australia.

Packard, A. (1994). Gain scheduling via linear fractional transformations. *System Control Letters*, 22(2), 79–92.

Pinkus, A. (1985). *n-Widths in Approximation Theory*. Springer-Verlag.

Tóth, R. (2008). *Modeling and Identification of Linear Parameter-Varying Systems, an Orthonormal Basis Function Approach*. Ph.D. thesis, Delft University of Technology.

Tóth, R., Felici, F., Heuberger, P.S.C., and Van den Hof, P.M.J. (2007a). Discrete time LPV I/O and state space representations, differences of behavior and pitfalls of interpolation. In *Proc. of the European Control Conf.*, 5418–5425. Kos, Greece.

Tóth, R., Heuberger, P.S.C., and Van den Hof, P.M.J. (2007b). LPV system identification with globally fixed orthonormal basis functions. In *Proc. of the 46th IEEE Conf. on Decision and Control*, 3646–3653. New Orleans, Louisiana, USA.

Tóth, R., Heuberger, P.S.C., and Van den Hof, P.M.J. (2008). Flexible model structures for LPV identification with static scheduling dependency. In *Proc. of the 45th IEEE Conf. on Decision and Control*, 4522–4527. Cancun, Mexico.

Tóth, R., Heuberger, P.S.C., and Van den Hof, P.M.J. (2009). Asymptotically optimal orthonormal bases functions for LPV system identification. *To appear in Automatica.*

Willems, J.C. (1991). Paradigms and puzzels in the theory of dynamical systems. *IEEE Tran. on Automatic Control*, 36(3), 259–294.

Zhou, K. and Doyle, J.C. (1998). *Essentials of Robust Control*. Prentice-Hall.