

Model Structures for Identification of Linear Parameter-Varying (LPV) Models

Paul M.J. Van den Hof, Roland Tóth and Peter S.C. Heuberger

Abstract Describing nonlinear dynamic systems by linear parameter-varying models has become an attractive tool for control of complex 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 models is presented, together with a general approach to the LPV identification problem. Use is made of a series expansion approach to LPV modeling, employing orthogonal basis function expansions.

1 Introduction

Many physical/chemical processes 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 [15]. This class of systems is particularly suited to deal with systems that operate in varying operating regimes. LPV systems can be seen as an extension of the class of *Linear Time-Invariant* (LTI) systems. In the LPV systems the signal relations are considered to be linear, but the model parameters are assumed to be functions of a measurable 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. Practical use of this framework is stimulated by the fact that LPV control design is well worked out, and the results of optimal and robust LTI control theory can be extended to nonlinear, time-varying plants [15, 17, 28].

In the past two decades several methods have been developed for the identification

Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands. e-mail: p.m.j.vandenhof@tudelft.nl

of discrete-time LPV models from measured data [4, 18, 16, 3, 26, 25]. Most of these approaches exploit the fact that an LPV system can be viewed as a collection of “local” models joined together by scheduling dependent weighting functions [15, 22].

The identification approaches that are presented in the literature so far all take the 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. This is also true for such 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 [13], 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 herein and the consequences of their use in identification are discussed. Particular attention will be given to a series-expansions approach in terms of orthonormal basis functions. The question whether the scheduling signal has a static or dynamic effect on the system coefficients is an important issue and is discussed in detail.

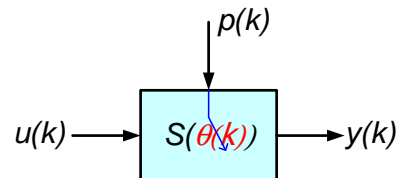
In this paper, we will restrict our attention to single input - single output (SISO) systems, but all the results can be applied to the MIMO case in a straightforward way.

2 Concepts and Notation

A conceptual view of an LPV system is depicted in Figure 1, emphasizing the fact that the system coefficients $\theta(k)$ – that are used to determine output $y(k)$ – are dependent on an external signal $p(k)$, while for a fixed $\theta_k = c$ the system S is linear time-invariant. At every time instant k , this linear dynamics is updated on the basis of the mapping $p \rightarrow \theta$.

LPV systems can be written in different representations, among which one is the LPV state-space description,

Fig. 1 LPV system representation, where for a fixed value of k , $S(\theta(k))$ describes an LTI system. The coefficient θ is a function of the scheduling variable p .



$$x(k+1) = A(p(k))x(k) + B(p(k))u(k) \quad (1a)$$

$$y(k) = C(p(k))x(k) + D(p(k))u(k) \quad (1b)$$

and the LPV IO model representation

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

Here, y and u are the output and the input of the system, respectively, x is the state vector, and the real-valued system coefficients (A, B, C, D) and $(a_i, b_j)_{i=1 \dots n_a; j=0 \dots n_b}$, are dependent on a scheduling function $p : \mathbb{Z} \rightarrow \mathbb{P}$, where $\mathbb{P} \subset \mathbb{R}^{n_p}$. It is assumed that p is either measurable or known.

A few observations should be added to these concepts:

- An important observation is that usually the map $p \rightarrow \theta$ (see Fig. 1) is assumed to be a static (nonlinear) mapping, i.e. $\theta(k)$ depends only on the value of $p(k)$. As will be shown in Section 3, this assumption is a core issue in the development of a solid theory of LPV systems.
- Note that in these representations there is no limitation or guarantee that the McMillan degree of the linear systems remains constant for every value of k .
- It is clear that LPV systems are closely related to the class of *Linear Time-Varying* (LTV) systems, with the restriction that knowledge about the time-varying behavior is limited by the fact that the scheduling signal p can generally only be measured on-line.
- With respect to control synthesis for LPV systems, it is important to note that virtually all methods are based on LPV state-space models, very often with the assumption that the dependence of the matrices on p is affine, i.e. every matrix function X in Eq. (1) can be decomposed as

$$X(p(k)) = X_0 + \sum_{i=1}^{n_p} X_i p_i(k), \quad (3)$$

where $\{X_i\}$ are real-valued matrices.

- If p is dependent on y , u or x the system is referred to as a quasi-LPV system.

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 (local) operating points corresponding to constant scheduling signals $p(k) \equiv \bar{p}_i$, $i = 1, \dots, N_l$, where N_l is the number of local models obtained in this fashion.
- The resulting local linear models are interpolated (possibly by using data from an additional global experiment) to form 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 scheduling signal, to estimate an LPV model.

For the estimation step in these identification approaches, both prediction-error methods and subspace methods are available [4, 25]. For interpolation, various techniques and approaches have been introduced, ranging from interpolation on pole estimates to the technique where each local (LTI) model is converted to a state-space model in canonical form, and subsequently the coefficients in this model are interpolated ([26]).

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

3.2 Model structure considerations

To get an indication that there are theoretical problems involved with the current practice, let us consider the LPV model representations in state-space and IO form in Eq. (1-2) and evaluate whether these two representations are equivalent, as is the case for LTI systems, or not. A simple example shows that they are not equivalent for LPV systems, if the mapping $p \rightarrow \theta$ is restricted to be static. Consider the following second-order state-space model in the form of Eq. (1):

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

With simple manipulations this system can be transcribed into IO form:

$$\begin{aligned} y(k) &= a_2(p(k-1))y(k-1) + a_1(p(k-2))y(k-2) \\ &\quad + b_2(p(k-1))u(k-1) + b_1(p(k-2))u(k-2), \end{aligned}$$

which is clearly not in the form defined by Eq. (2). In order to obtain equivalence between the state-space and IO representations, it is necessary to allow for a dynamic mapping $p \rightarrow \theta$, i.e. allow $\theta(k)$ to depend on $\{\dots, p(k-1), p(k), p(k+1), \dots\}$ [20].

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 [6]. It can be shown (see [20]) that this results in transformation matrices – and consequently also in state-space matrices – that depend dynamically on the scheduling parameter p . Note that by using only LTI based state-space transformations, there is no guarantee that the resulting state vectors have a common basis.

It can be concluded that transforming estimated IO models to state-space descriptions, while retaining a static dependence on the scheduling function, as well as, using LTI state-space transformations on local models (before interpolation) may result in errors. This especially holds for situations with rapidly varying scheduling signals, as further illustrated in [20].

3.3 A behavioral approach

From the previous sections it can be concluded that the classical formulation of LPV models should be adapted to deal with dynamic scheduling dependence. In [19] the behavioral framework, originally developed¹ for LTI systems ([13]), is extended to deal with LPV systems. In this framework, a parameter-varying system S is defined as a quadruple

$$S = (\mathbb{T}, \mathbb{P}, \mathbb{W}, \mathfrak{B}), \quad (4)$$

where \mathbb{T} is called the time axis, \mathbb{P} denotes the scheduling space (i.e. $p(k) \in \mathbb{P}$), \mathbb{W} is the signal space with dimension n_w and $\mathfrak{B} \subset (\mathbb{P} \times \mathbb{W})^{\mathbb{T}}$ is the *behavior* of the system. The set \mathbb{T} defines the time-axis of the system, describing continuous, $\mathbb{T} = \mathbb{R}$, and discrete, $\mathbb{T} = \mathbb{Z}$, systems alike, while \mathbb{W} gives the range of the system signals. \mathfrak{B} defines the physical laws, i.e. the rules for selecting which trajectories of $(\mathbb{P} \times \mathbb{W})^{\mathbb{T}}$ are possible. In the sequel, we restrict attention to the discrete-time case. Note that there is no prior distinction between inputs and outputs in this setting.

We also introduce the so-called *projected scheduling behavior*

$$\mathfrak{B}_{\mathbb{P}} = \{p \in \mathbb{P}^{\mathbb{T}} \mid \exists w \in \mathbb{W}^{\mathbb{T}} \text{ s.t. } (w, p) \in \mathfrak{B}\}, \quad (5)$$

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

$$\mathfrak{B}_p = \{w \in \mathbb{W}^{\mathbb{T}} \mid (w, p) \in \mathfrak{B}\}. \quad (6)$$

With these concepts we can define LPV systems as follows:

Definition 1. (LPV system) The parameter-varying system S is called LPV, if the following conditions are satisfied:

¹ In the past decades this framework has been extended to LTV ([27, 8]), and even *nonlinear* (NL) systems ([13, 14]).

- \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).
- \mathbb{T} is closed under addition.
- For any $(w, p) \in \mathfrak{B}$ (a signal trajectory associated with a scheduling trajectory) 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)².

In the next step, the behavior of LPV systems has to be specified in terms of mathematical representations. The coefficients in these representations will become (nonlinear) functions of the scheduling signal p . In order to describe this functional dependence of a single real-valued coefficient in one of the representations to be introduced in the next section, we employ functions

$$r: \mathbb{R}^n \rightarrow \mathbb{R},$$

that are considered to be in the set $\mathfrak{R} = \bigcup_n \mathfrak{R}_n$, where \mathfrak{R}_n is the set of essentially³ 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 elements of the set $\{q^i p_j\}_{i \in \mathbb{Z}; j=1, \dots, n_p}$. In order to specify the (time-varying) coefficient, we introduce new notation through the operator

$$\diamond: (\mathfrak{R}, \mathfrak{B}_{\mathbb{P}}) \rightarrow \mathbb{R}^{\mathbb{Z}}, \text{ defined by } (r \diamond p)(k) = r(x(k)) \quad (7)$$

where x is a vector of n signals, being constructed by taking the first n components of the signal vector

$$[p^T \quad q^{-1}p^T \quad qp^T \quad q^{-2}p^T \quad q^2p^T \quad \dots]^T. \quad (8)$$

A (scheduling-dependent) coefficient in an LPV system representation is evaluated by an operation $(r \diamond p)(k)$.

Example 2. (Coefficient function) Let $\mathbb{P} = \mathbb{R}^{n_p}$ with $n_p = 2$. Consider the coefficient

$$\frac{1 + p_1(k-1)}{1 - p_2(k)}.$$

In order to describe this coefficient with a real-meromorphic function r , we need a function with dimension 3, i.e. $x = [p_1, p_2, q^{-1}p_1]$ specified by

$$r(x_1, x_2, x_3) = \frac{1 + x_3}{1 - x_2}.$$

With this specification of r , $(r \diamond p)(k) = \frac{1 + p_1(k-1)}{1 - p_2(k)}$.

² Here q denotes the standard shift operator, $q^i x(t) = x(t+i)$, $i \in \mathbb{Z}$.

³ In the sense that $r(x_1, \dots, x_n)$ does depend on x_n .

In the sequel the (time-varying) coefficient sequence $(r \diamond p)$ will be used to operate on a signal w . In this respect it is important to note that multiplication with the shift operator q is not commutative, in other words $q(r \diamond p)w \neq (r \diamond p)qw$.

Shift operations $\vec{r}, \overleftarrow{r}$ can be defined by the equations $\vec{r} \diamond p = r \diamond (qp)$, and $\overleftarrow{r} \diamond p = r \diamond (q^{-1}p)$, respectively. With this notion, it follows that $q(r \diamond p)w = (\vec{r} \diamond p)qw$ and $q^{-1}(r \diamond p)w = (\overleftarrow{r} \diamond p)q^{-1}w$.

The considered operator \diamond can straightforwardly be extended to matrix functions $r \in \mathfrak{R}^{n_r \times n_w}$ where \diamond is applied to each scalar entry of the matrix, as well as to polynomial matrices in q . Let $\mathfrak{R}[q]^{n_r \times n_w}$ denote the set of polynomial matrices in q with coefficients in \mathfrak{R} , then

$$(R(q) \diamond p)w := \sum_{i=0}^{n_q} (r_i \diamond p) q^i w \quad (9)$$

where $R(q) = \sum_{i=0}^{n_q} r_i q^i$, n_q is the order of R , and r_i is a $n_r \times n_w$ -dimensional matrix with elements in \mathfrak{R} . In this notation, the shift operation q operates on the signal w , while operation \diamond takes care of the time/schedule-dependent coefficient sequence.

3.4 LPV system representations

Kernel representation

Using the behavioral framework, we can introduce the so-called kernel representation of an LPV system. By employing the notation presented in the previous section, a kernel representation of an LPV system is written as

$$(R(q) \diamond p)w = 0. \quad (10)$$

We call difference equation (10) a discrete-time *kernel representation* of LPV system $S = (\mathbb{T}, \mathbb{P}, \mathbb{W}, \mathfrak{B})$ with scheduling signal p and signals w if

$$\mathfrak{B} = \{(w, p) \in (\mathbb{R}^{n_w} \times \mathbb{R}^{n_p})^{\mathbb{Z}} \mid (R(q) \diamond p)w = 0\}. \quad (11)$$

In the sequel we only consider LPV systems, whose behavior can be described by Eq. (11). An important property of these systems is that they have a kernel representation where R has full row rank ([19]).

IO representation

For practical applications one will often need a partitioning of the signals w in input signals $u \in (\mathbb{R}^{n_u})^{\mathbb{Z}}$ and output signals $y \in (\mathbb{R}^{n_y})^{\mathbb{Z}}$. Note that this partitioning is not trivial and can neither be chosen freely. For details see [13, 19].

Using an IO partitioning, we can define the *IO representation* as

$$(R_y(q) \diamond p)y = (R_u(q) \diamond p)u, \quad (12)$$

where R_u and R_y are again matrix polynomials with meromorphic coefficients, and where R_y is full row rank with $\text{order}(R_y) \geq \text{order}(R_u)$.

Using the same notation and decomposition as in (9), it follows that

$$\sum_{i=0}^{n_a} (a_i \diamond p) q^i y = \sum_{j=0}^{n_b} (b_j \diamond p) q^j u, \quad (13)$$

where $n_a \geq n_b \geq 0$, and $n_a \geq 1$.

State-space representation

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

$$qx = (A \diamond p)x + (B \diamond p)u \quad (14a)$$

$$y = (C \diamond p)x + (D \diamond p)u, \quad (14b)$$

where $\mathbb{X} \subset \mathbb{R}^{n_x}$ is called the state space and the (parameter-varying) state space matrices (A, B, C, D) are matrices of appropriate dimensions with their entries being meromorphic functions in \mathfrak{R} . Note that the latent variable x in (14) qualifies as a state variable.

It is apparent that (13) and (14) are the ‘dynamic-dependency’ counterparts of (2) and (1), respectively.

3.5 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 [19].

Transformations between different representations, as well as, state transformations into a different coordinate system, generally involve dynamically dependent relations. For instance, the transformation of an LPV state-space model (14) into an observable canonical form requires a transformation matrix $T \in \mathfrak{R}^{n_x \times n_x}$ to obtain a new state

$$x' = (T \diamond p)x$$

and state-space matrices $A' = \overrightarrow{T}AT^{-1}$, $B' = \overrightarrow{T}B$, $C' = CT^{-1}$, $D' = D$.

Here, the matrix T is constructed from the LPV observability matrix, which is in the SISO case built up from

$$C, \overrightarrow{C}A, \overrightarrow{\overrightarrow{C}A}A, \dots.$$

Suppose that the original state-space model has static dependency on the scheduling function, so at time instant k the matrix functions depend on values of $p(k)$ only, then the construction of the transformation matrix T , as well as, the calculation of the new state-space matrices immediately imply that the new matrices depend – at time k – on *future* values $p(k + \tau)$ ($\tau > 0$), as well.

This problem can be circumvented by using a reachability canonical form, in which case the transformation only involves backward shift operations ([19]).

4 An orthonormal basis functions approach

4.1 Series-expansion representations

In this section we explore the possibilities for using a series-expansion type of model structure for LPV systems, using the concept of orthonormal basis functions (OBF) [7]. A major motivation is the linear-in-the-parameters property of these structures, which is very beneficial in prediction-error identification. A second merit of these structures is that they allow a relatively simple interpolation of local linear models with varying McMillan degree. Furthermore, it was shown in [2] for nonlinear Wiener-models (that is a 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.

For a (local) linear model $G \in H_2$ it holds that G can be written as

$$G(z) = D + \sum_{k=1}^{\infty} c_k F_k(z), \quad (15)$$

where $\{F_k\}$ is a basis for H_2 . In the theory of generalized orthonormal basis functions (GOBF's), the functions $F_k(z)$ are generated by applying a Gram-Schmidt orthonormalization to the sequence of functions

$$\frac{1}{z - \xi_1}, \dots, \frac{1}{z - \xi_{n_b}}, \frac{1}{(z - \xi_1)^2}, \dots$$

with stable pole locations ξ_1, \dots, ξ_{n_b} . The choice of these *basis poles* determines the rate of convergence of the series expansion (15).

An alternative derivation of the basis functions is based on a balanced realization $\{A_b, B_b, C_b, D_b\}$ of the inner function

$$G_b(z) = \prod_{k=1}^{n_b} \frac{1 - z\xi_k^*}{z - \xi_k}, \quad (16)$$

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

$$(zI - A_b)^{-1} B_b G_b^i(z), \quad i = 1, 2, \dots$$

By using a truncated expansion in (15), an attractive model structure for LTI identification results, with a well worked-out theory in terms of variance and bias expressions. The series expansion (15) can be extended to LPV systems, such that for a given basis $\{F_k\}$ and a specific IO-partitioning (u, y) , an LPV system can be written as

$$y(t) = \left((D \diamond p) + \sum_{k=1}^n (w_k \diamond p) F_k(q) \right) u(t), \quad (17)$$

where $n = \infty$. It is an obvious step to use a truncated expansion, i.e. (17) with finite 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. Similarly to the LTI case, this structure is linear in the parameters. An important question that arises is whether the basis functions F_k can be chosen in such a manner that a fast rate of convergence can be accomplished for all possible scheduling trajectories p . Note that the representation (17) is equivalent with a state-space description (14), where the matrices A and B are independent of the scheduling function.

4.2 Basis selection

In order to select a basis, it is important to obtain some knowledge about the system to be modeled. For the LTI case, it is well-known that – if the underlying system can be well approximated by an LTI model – an optimal basis can be chosen relying on the knowledge about the system poles. It can be shown that the same property holds for LPV systems, where the knowledge of the poles of all possible local linear 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{p}_i , i.e., by using data with a constant scheduling signal $p(k) \equiv \bar{p}_i$.

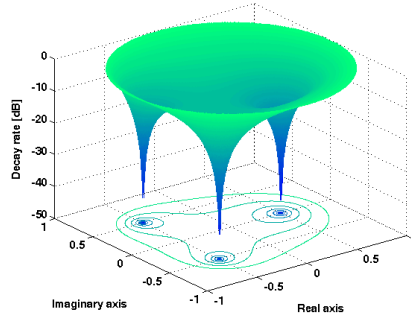
2. Plot all poles of the identified models in the complex plane.
3. Cluster the poles into groups and find optimal cluster centers (these centers will be used as basis poles)

In this procedure, use is made of minimization of a distance measure, which is relevant for the worst-case approximation error of the representation (17). This scheme is motivated by the extension of the classical *Kolmogorov n-width* result of [12] to OBFs, as obtained by [11]. This result states that for a given LTI inner function G_b , the 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

$$\{z \in \mathbb{D} \mid |G_b(z^{-1})| \leq \rho\}.$$

Here ρ is the rate of convergence in the series expansion, and n should be a multiple of the number of basis poles n_b . See Fig. 2, taken from [23], for an example of such regions.

Fig. 2 Example of the function $|G_b(z^{-1})|$ and the region $\{z \in \mathbb{D} \mid |G_b(z^{-1})| \leq \rho\}$ for an inner function G_b with 3 poles and various values of ρ . Note that if z_0 is a pole of G_b , then $G_b(z_0^{-1}) = 0$.



For the basis-selection problem, we are dealing with its inverse problem, i.e. given a region of poles Ω , approximate this region as

$$\Omega \approx \Omega(\Xi, \rho) = \{z \in \mathbb{D} \mid |G_b(z^{-1})| \leq \rho\}. \quad (18)$$

The n optimal OBF poles $\Xi = \{\xi_1, \dots, \xi_n\}$ are therefore obtained by solving the following *Kolmogorov measure minimization* problem,

$$\min_{\Xi \subset \mathbb{D}} \rho = \min_{\Xi \subset \mathbb{D}} \max_{z \in \Omega} |G_b(z^{-1})| = \min_{\Xi \subset \mathbb{D}} \max_{z \in \Omega} \left| \prod_{k=1}^n \frac{1 - z\xi_k^*}{z - \xi_k} \right| \quad (19)$$

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 -max clustering (FcM), have been used in a wide collection of applications [1, 10]. Generally, FcM partitions the data into overlapping groups, that describe an underlying structure within the data [9]. In this section, we describe 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 $Z = \{z_1, \dots, z_N\}$.

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. Furthermore, we define membership functions $\mu_i : \mathbb{D} \rightarrow [0, 1]$, 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

$$\Omega = \{z \in \mathbb{D} \mid \exists i, \mu_i(z) \geq \varepsilon\}. \quad (20)$$

With these preliminaries, we can now formulate the problem.

Problem 3. For a given c , find a region Ω , as described above, such that Ω contains all pole locations in 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 the corresponding decay rate ρ as small as possible.

To measure the 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 a metric on \mathbb{D} , defined by

$$\kappa(x, y) = \left| \frac{x - y}{1 - x^*y} \right|, \quad (21)$$

referred to as the *Kolmogorov metric*.

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* [1], J_m , which – in the FKcM case – can be formulated as

$$J_m = \max_{1 \leq k \leq N} \sum_{i=1}^c \mu_{ik}^m d_{ik}. \quad (22)$$

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 (22) corresponds to a *worst-case (therefore max) sum-of-error* criterion, contrary to the *mean-squared-error* criterion of the original FcM,

see [1].

The crucial property of this functional is that it can be shown (see [21]) that for large values of m the minimization of J_m is equivalent to the Kolmogorov measure minimization problem (19). For details, as well as, a detailed description of the optimization algorithm, see [19], where also the robust extension of the basis selection is discussed. In that case, not only pole estimates are considered, but also the corresponding uncertainty regions of these estimates. See Fig. 3 for an example of the basis selection mechanism.

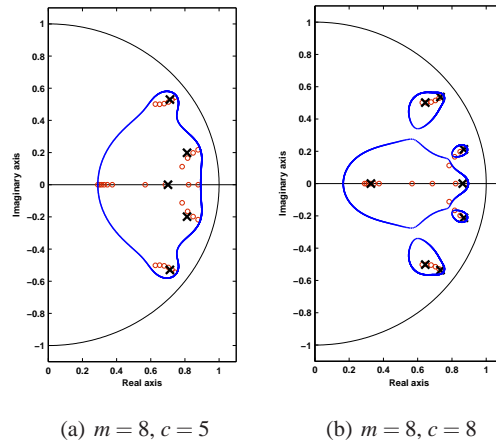
For the determination of the actual number of clusters in these algorithms, so-called adaptive cluster-merging is applied. Starting from a relatively large initial number of clusters (typically around $N/2$), the adaptive merging steers the algorithm towards the natural number of groups that can be observed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

4.4 OBF-based model structures

We assume that the basis selection step has been completed and we are given a set of n_f basis functions $\{F_k(z)\}_{k=1,\dots,n_f}$ with good approximation properties for the set of local LTI behaviors corresponding to constant scheduling signals. In the next step, we can construct model structures for the identification of an LPV system S . To keep the notation simple, we restrict attention to strictly proper models (i.e., $D = 0$ in (17)). The input-output dynamics of the LPV model can be written as

$$y(t) = \sum_{k=1}^{n_f} (w_k \diamond p)(t) F_k(q) u(t). \quad (23)$$

Fig. 3 Example of the basis selection procedure, using fuzzy clustering with fuzziness parameter $m = 8$. The 30 observed poles (i.e the set Z) are given with red circles. The resulting cluster centers are depicted with a black x. The blue lines represent the region Ω as in Eq. (18), obtained by using the cluster centers as basis poles. On the left hand side $c = 5$ clusters were determined, on the right hand side $c = 8$.



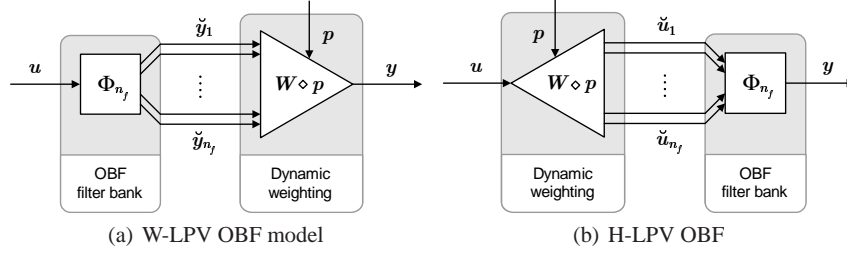


Fig. 4 IO signal flow graph of (a) the W-LPV OBF model described by (23) and (b) the H-LPV OBF model described by (26)

Introduce Φ_{n_f} and W as shorthand notation for the vectors with functions F_k and coefficients w_i , respectively.

$$\Phi_{n_f} = [F_1 \cdots F_{n_f}]^T \quad (24)$$

$$W = [w_1 \cdots w_{n_f}]^T. \quad (25)$$

Then the model structure (23) can be visualized as shown in Fig. 4a, where $\check{y}_i(t) = F_i(q)u(t)$. 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. 4b, is the so-called Hammerstein LPV OBF model (H-LPV OBF), that is derived from the description

$$y(t) = \sum_{k=1}^{n_f} F_k(q) (w_k \diamond p)(t) u(t). \quad (26)$$

This latter structure can be generated from the LTI series expansion (15), by changing the order of the arguments. This change has no effect in the LTI case, but 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_k have only a static dependency on the scheduling function p , so we can write $(w_k \diamond p)(t) = w_k(p(t))$ in (23). As stated before, we can write the W-LPV OBF structure also in a state-space form,

$$qx = Ax + Bu \quad (27a)$$

$$y = (W \diamond p)x, \quad (27b)$$

where the constant matrices A and B are completely determined by the basis functions $\{F_k\}$. This illustrates that the dependency on the scheduling signal is only present in the output equation, with the result that the assumption of static dependency is much less restrictive than in the general case (14).

With respect to the actual estimation with these model structures, we again distinguish between a local and a global approach.

Local estimation approach

This approach is based on a number N_l of ‘local’ experiments, i.e. data collection with a constant scheduling signal $p(k) \equiv p_i \in \mathbb{P}$, resulting in data sequences $\{u_i(t), y_i(t)\}$ for $i = 1, \dots, N_l$. Based on these data, N_l LTI models are estimated using a standard least-squares criterion, on the basis of the one-step-ahead prediction error in Output Error (OE) form:

$$\varepsilon(t) = y_i(t) - \sum_{k=1}^{n_f} w_{ik} F_k(q) u_i(t), \quad (28)$$

where $\{w_{ik}\}$ are real-valued coefficients. Note that – under the condition that the data are persistently exciting – there exists a unique analytic solution to this estimation problem. These estimated coefficients can be considered as ‘samples’ of the function $w_k(p(t))$, in the sense that $w_k(p_i) = w_{ik}$. As a second step, we use interpolation to obtain estimates of the function $w_k(p(t))$, for instance by assuming a polynomial dependency of w_k on p , or by making use of splines etc.

Global estimation approach

For this approach we need to assume a specific functional dependency of the functions w_k on $p(t)$ and we propose to use a linear parametrization for this purpose, such as a polynomial dependency

$$w_k(p(t)) = w_{k0} + w_{k1}p(t) + \dots + w_{kr}p^r(t).$$

Here we assumed for simplicity that p is a one-dimensional signal. Now we collect a global data set $\{u(t), y(t), p(t), t = 1, \dots, T\}$, which is assumed to be persistently exciting the system at hand⁴. It is straightforward that – using a least-squares criterion – a unique analytic solution can be obtained for the parameters w_{ki} .

Note that the restriction to static dependency can be relaxed for the global approach by allowing a dependency of w_k on time-shifts of $p(t)$, as well.

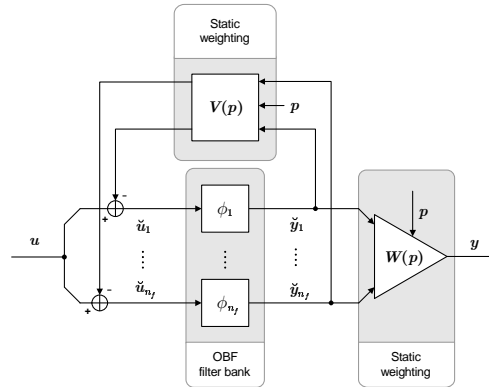
Because of the postulated OBF structure, both approaches will always result in asymptotically stable models.

4.5 Approximation of dynamic dependency

In order to alleviate the restrictions caused by the assumption of static dependency in the suggested model structures, extensions for these structures were proposed in [24]. Here, we only consider the extension of the W-LPV OBF model structure. The

⁴ Persistency of excitation for LPV systems is not yet completely understood.

Fig. 5 IO signal flow graph of the WF-LPV OBF model with feedback-based static weighting functions V and W .



idea is still to use weighting functions with static dependency, but with the introduction of an additional feedback loop around each basis component with a gain incorporating also static dependency. In this way, the filter bank of OBFs – as a dynamical LTI system – is reused to provide dynamic weighting functions that can approximate the required class of dependency for W-LPV OBF models. The introduction of feedback-based weighting leads to a new model structure given in Fig. 5, which we call Wiener Feedback LPV (WF-LPV OBF) models. See [19] for an analogous extension of the H-LPV OBF model structure. For these new model structures, it is apparent that by setting the feedback gains to zero, the previous structures result. This immediately indicates an increase in the representation capability of the extended structures. The W-LPV OBF can be represented in state-space form by

$$qx = (A - BV(p)C)x + Bu \quad (29a)$$

$$y = W(p)Cx, \quad (29b)$$

where the constant matrices A , B and C are again completely determined by the basis functions. These equations illustrate how the addition of the feedback loops introduces dependency on the scheduling signal in the state equations. 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 of the whole parameter set (i.e. the coefficients of W and V), the approach utilizes a separable least squares optimization scheme [5]. In each iteration cycle of this scheme, one set of the parameters is fixed to enable a linear-regression-based estimation of the other set. This results in a steepest descend algorithm which is guaranteed to converge to a saddle point or a local minimum, depending on the initial values of the parameters. For algorithmic details see [24, 19]. 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.

Table 1 Validation results of 100 identification experiments by the global and local methods using the W-LPV OBF and H-LPV OBF model structures. The results are given in terms of the average MSE and VAF of the simulated output signals of the model estimates.

Model	Case	MSE (dB)	VAF (%)
W-LPV	loc.	-18.01	94.12%
	glob.	-31.03	98.23%
H-LPV	loc.	-10.16	85.69%
	glob.	-26.41	96.18%

5 Example

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

$$\sum_{i=0}^5 (a_i \diamond p) q^i y = (b_4 \diamond p) q^4 u, \quad (30)$$

with $\mathbb{P} = [0.6, 0.8]$ and coefficients

$$\begin{aligned} a_0 \diamond p &= -0.003, & a_3 \diamond p &= \frac{61}{110} - 0.2 \sin(q^5 p), \\ a_1 \diamond p &= \frac{12}{125} - 0.1 \sin(q^5 p), & a_4 \diamond p &= -\frac{511 + 192q^5 p^2 - 258(\cos(q^5 p) - \sin(q^5 p))}{860}, \\ a_2 \diamond p &= -\frac{23}{85} + 0.2 \sin(q^5 p), & a_5 \diamond p &= 0.58 - 0.1q^5 p, \\ b_4 \diamond p &= \cos(q^5 p). \end{aligned}$$

Using 8 basis functions, obtained through the FKcM algorithm (see [24, 23] for details) and a 2nd order polynomial-based parametrization of the coefficients, identification of S with the local, as well as, the global approach has been carried out, with the W-LPV OBF and the H-LPV OBF model structures. Each experiment has been repeated 100 times with different realizations of input, scheduling and noise signals. The signal-to-noise ration was 20 dB in the resulting data records with a relative signal-to-noise amplitude of 25 %. See Table 1 for the results in terms of average MSE and VAF (Variance Accounted For).

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 $\{a_l\}_{l=0}^5$. However, the global W-LPV identification provided quite acceptable results for such a heavily nonlinear system. The explanation why the H-LPV structure gave a worse result lies in the different approximation capabilities of these models.

To illustrate the effect of incorporating feedback, we used the same example system and identified it with the WF-LPV OBF model structure, as well as, with the W-LPV OBF structure. For both structures, the coefficients in W are parametrized as

Table 2 Validation results of 100 identification experiments with the Wiener (W) and the Wiener Feedback (WF) model structures. The results are given in terms of the average MSE and VAF of the simulated output signals of the model estimates.

SNR	MSE (dB)		VAF (%)	
	W	WF	W	WF
no noise	-34.96	-39.75	90.04	99.42
35 dB	-34.77	-39.17	98.99	99.39
20 dB	-32.75	-35.01	98.71	99.00
10 dB	-31.81	-32.38	98.19	98.59

a 2nd order polynomial and for V a 3rd order polynomial was used. Identification of S with the global approach was accomplished 100 times in 4 different noise settings with both the Wiener and the Wiener-feedback model structures. See Table 2 for the results. As expected, both 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. For all cases, the WF-LPV model provided better estimates than the pure static-dependence based W-LPV model estimate. This clearly shows the improvement in the approximation capability due to the approximation of dynamic dependence with feedback-based weighting.

6 Conclusions

On the basis of a solid system theoretic definition of LPV systems in terms of system behaviors, several LPV model representations are presented and brought into a unifying framework. Real-valued meromorphic functions are used to specify dynamic dependency of the system coefficients on the scheduling signal. A series expansion approach is presented for modelling 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 are illustrated in an example.

References

1. Bezdek, J.C.: Pattern recognition with fuzzy objective function algorithms. Plenum Press (1981)
2. Boyd, S., Chua, L.O.: Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. on Circuits and Systems* **32**(11), 1150–1161 (1985)
3. Felici, F., van Wingerden, J.W., Verhaegen, M.: Subspace identification of MIMO LPV systems using a periodic scheduling sequence. *Automatica* **43**(10), 1684–1697 (2006)



Fig. 6 Professor Paul M.J. Van den Hof delivering his lecture

4. Giarré, L., Bauso, D., Falugi, P., Bamieh, B.: LPV model identification for gain scheduling control: an application to rotating stall and surge control problem. *Control Engineering Practice* **14**, 351–361 (2006)
5. Golub, G.H., Pereyra, V.: The differentiation of pseudo-inverses and nonlinear least squares problems whose variables are separate. *SIAM Journal on Numerical Analysis* **10**(2), 413–432 (1973)
6. Guidorzi, R., Diversi, R.: Minimal representations of MIMO time-varying systems and realization of cyclostationary models. *Automatica* **39**(11), 1903–1914 (2003)
7. Heuberger, P.S.C., Van den Hof, P.M.J., (Eds.), B.: Modeling and identification with rational orthonormal basis functions. Springer-Verlag (2005)
8. Ilchmann, A., Mehrmann, V.: A behavioral approach to time-varying linear systems. Part 1: General theory. *SIAM Journal of Control Optimization* **44**(5), 1725–1747 (2005)
9. Jain, A.K., Dubes, R.C.: Algorithms for clustering data. Prentice Hall (1988)
10. Kaymak, U., Setnes, M.: Fuzzy clustering with volume prototypes and adaptive cluster merging. *IEEE Trans. on Fuzzy Systems* **10**(6), 705–711 (2002)
11. Oliveira e Silva, T.: A n -width result for the generalized orthonormal basis function model. In: Proc. of the 13th IFAC World Congress, Sydney, Australia (1996), pp. 375–380.

12. Pinkus, A.: *n*-widths in approximation theory. Springer-Verlag (1985)
13. Polderman, J.W., Willems, J.C.: Introduction to mathematical systems theory, a behavioral Approach. Springer (1991)
14. Pommaret, J.F.: Partial differential control theory, vol. I, Mathematical tools. Kluwer Academic Publishers (2001)
15. Rugh, W., Shamma, J.: Research on gain scheduling. *Automatica* **36**(10), 1401–1425 (2000)
16. dos Santos, P.L., Ramos, J.A., de Carvalho, J.L.M.: Identification of linear parameter varying systems using an iterative deterministic-stochastic subspace approach. In: Proc. of the European Control Conf., pp. 4867–4873. Kos, Greece (2007)
17. Scherer, C.W.: Mixed H_2/H_∞ control for time-varying and linear parametrically-varying systems. *Int. Journal of Robust and Nonlinear Control* **6**(9-10), 929–952 (1996)
18. Sznaier, M., Mazzaro, C., Inanc, T.: An LMI approach to control oriented identification of LPV systems. In: Proc. of the American Control Conf., Chicago, Illinois, USA (2000), pp. 3682–3686.
19. Tóth, R.: Modeling and identification of linear parameter-varying systems. An orthonormal basis function approach. Ph.D. thesis, Delft University of Technology, The Netherlands (2008)
20. Tóth, R., Felici, F., Heuberger, P.S.C., Van den Hof, P.M.J.: Discrete time LPV I/O and state space representations, differences of behavior and pitfalls of interpolation. In: Proc. of the European Control Conf., pp. 5418–5425. Kos, Greece (2007)
21. Tóth, R., Heuberger, P.S.C., Van den Hof, P.M.J.: Orthonormal basis selection for LPV system identification, the Fuzzy-Kolmogorov *c*-Max approach. In: Proc. of the 45th IEEE Conf. on Decision and Control, pp. 2529–2534. San Diego, California, USA (2006)
22. Tóth, R., Heuberger, P.S.C., Van den Hof, P.M.J.: LPV system identification with globally fixed orthonormal basis functions. In: Proc. of the 46th IEEE Conf. on Decision and Control, pp. 3646–3653. New Orleans, Louisiana, USA (2007)
23. Tóth, R., Heuberger, P.S.C., Van den Hof, P.M.J.: Asymptotically optimal orthonormal bases functions for LPV system identification. Provisionally accepted to *Automatica* (2008)
24. Tóth, R., Heuberger, P.S.C., Van den Hof, P.M.J.: Flexible model structures for LPV identification with static scheduling dependency. To appear in Proc. 47th IEEE Conf. on Decision and Control (2008)
25. Verdult, V., Verhaegen, M.: Subspace identification of multivariable linear parameter-varying systems. *Automatica* **38**(5), 805–814 (2002)
26. Wassink, M.G., van de Wal, M., Scherer, C.W., Bosgra, O.: LPV control for a wafer stage: beyond the theoretical solution. *Control Engineering Practice* **13**, 231–245 (2004)
27. Zerz, E.: An algebraic analysis approach to linear time-varying systems. *IMA Journal of Mathematical Control and Information* **23**, 113–126 (2006)
28. Zhou, K., Doyle, J.C.: *Essentials of robust control*. Prentice-Hall (1998)