

LPV system identification using series expansion models

Citation for published version (APA):

Toth, R., Heuberger, P. S. C., & Hof, Van den, P. M. J. (2011). LPV system identification using series expansion models. In P. L. Santos, dos, T. P. A. Perdicoulis, C. Novara, J. A. Ramos, & D. E. Rivera (Eds.), *Linear parameter-varying system identification : new developments and trends* (pp. 259-294). (Advanced Series in Electrical and Computer Engineering; Vol. 14). World Scientific. https://doi.org/10.1142/9789814355452_0010

DOI:

[10.1142/9789814355452_0010](https://doi.org/10.1142/9789814355452_0010)

Document status and date:

Published: 01/01/2011

Document Version:

Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Chapter 10

LPV system identification using series expansion models

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

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
r.toth@tudelft.nl
p.s.c.heuberger@tudelft.nl
p.m.j.vandenhof@tudelft.nl

In this chapter an Orthonormal Basis Functions (OBFs) model structure is proposed in the LPV context with advantageous properties in terms of estimation and realization. A solid system theoretic basis for the description of LPV systems in terms of these model structures is presented together with a general approach to LPV identification. Data-driven model structure selection is also discussed in this setting and the stochastic properties of the employed identification schemes are analyzed. The introduced approaches are demonstrated on an industrially relevant example.

1. Introduction

Describing nonlinear systems by *linear parameter-varying* (LPV) models has become an attractive approach to address control of complicated systems with regime-dependent (linear) behavior. However in LPV *data-driven modeling* or *identification* of such systems, it is a delicate issue to decide what kind of model structure will be used to capture the underlying dynamic behavior. As the LPV class is more like a modeling philosophy than an actual interpretation of first-principle relations, like the *linear time-invariant* (LTI) or *affine nonlinear*, etc. model classes, the actual possibilities are numerous with a lot of sensitive details. Beside the question of deciding what an adequate choice of the *scheduling variable* p is (see [Tóth

(2010); Kwiatkowski *et al.* (2006)]), the representation form of the model structure and the parametrization itself are also of crucial importance.

In the current literature, many successful approaches have been developed using model structures that are formulated in the form of *state-space* (SS) and *linear-fractional representations* (LFR), *input-output* (IO) representations or *series-expansion* forms. For a recent overview see [Tóth (2010); Casella and Lovera (2008)]. Despite the significant advances this field have recently seen, crucial questions about consequences of using an SS or an IO model, equivalence of representations and hence model parameterizations, and the degree of usefulness of the obtained models often remain undiscussed. However, these questions become important if the methods are applied in an engineering context.

By taking a fresh look on these problems, in this chapter, an *orthonormal basis functions* (OBFs) based LPV model structure using series expansion is introduced. This model structure appears to have advantageous properties, compared to other model structures, in terms of estimation and realization. A solid system theoretic basis for the description of LPV systems by LPV-OBF models is presented, together with a general approach to LPV identification both in the *local* (identification for constant p and interpolation) and the *global* (identification with varying p) setting. Data-driven model structure selection is also discussed and the stochastic properties of the employed identification schemes are investigated. Finally, the introduced approaches are demonstrated on an industrially relevant example.

2. Perspectives of series-expansion models

In order to shed more light on the specific problems of SS or IO models in the LPV context and why a different representation can be advantageous, we will take a closer look on these LPV representation forms. In specific, we will introduce SS, LFR and IO representations, discuss some important properties, and try to seek an answer to the question: does it really matter how the model equations are formulated or it is just a matter of personal choice? As we will see, such choices might have heavy consequences and it is often attractive to avoid the involved problems, perhaps by using alternative formulations like series-expansion forms.

A general dynamic description of a discrete-time LPV system \mathcal{S} can be formalized as a convolution:

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

in terms of the input signal $u : \mathbb{Z} \rightarrow \mathbb{R}^{n_u}$ and the *scheduling variable* $p : \mathbb{Z} \rightarrow \mathbb{P}$, with range $\mathbb{P} \subseteq \mathbb{R}^{n_p}$ often called the *scheduling “space”*. Here $y : \mathbb{Z} \rightarrow \mathbb{R}^{n_y}$ denotes the output signal of \mathcal{S} and $k \in \mathbb{Z}$ is the *discrete time* (DT). The coefficients h_i of Eq. (1) are *static* or *dynamic* matrix functions of p with arbitrary complexity ranging from simple linear to rational or real *meromorphic*^a dependence bounded on \mathbb{P} . A static function (dependence) means that $h_i(p, k)$ depends only on $p(k)$. The situation where $h_i(p, k)$ depends on multiple but finite many time-shifted instances of p , like $\{p(k + \tau_1), \dots, p(k - \tau_2)\}$ with $\tau_1, \tau_2 \geq 0$, is called *dynamic dependence*. To have a convenient notation to express both static and dynamic dependencies, we introduce the operator $\diamond : (\mathcal{R}, \mathbb{P}^{\mathbb{Z}}) \rightarrow \mathbb{R}_{\infty}^{\mathbb{Z}}$, where \mathcal{R} is the set of all real meromorphic functions with finite dimensional domain, such that $(h_i \diamond p)(k) = h_i(p(k + \tau_1), \dots, p(k - \tau_2))$.

In Eq. (1), the sequence $\{h_i\}_{i=0}^{\infty}$ defines the varying linear dynamical relation between u and y . This description is 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. $q^{-i}u(k) = u(k-i)$. Thus Eq. (1) is also called the *impulse response representation* (IIR) of \mathcal{S} . It can be proven that for an asymptotically stable \mathcal{S} , the expansion (1) is convergent [Tóth (2010)].

An important property of LPV systems is that for a constant scheduling signal, i.e. $p(k) = \mathbf{p}$ for all $k \in \mathbb{Z}$, Eq. (1) is equal to a convolution describing an LTI system as each $h_i(p, k)$ is constant. Thus, LPV systems can be seen to be similar to LTI systems, but with a different signal behavior due to the variation of each h_i . Note that there are many formal definitions of LPV systems based on particular model structures and parameterizations. The convolution form (1) can be seen as their generalization.

Two important formulations are LPV *state-space* (SS) representations and LFRs, commonly used in the control literature. LPV-SS representations of a given LPV system \mathcal{S} , denoted as $\mathfrak{R}_{\text{SS}}(\mathcal{S})$, are often defined under the assumption of static dependence in the form of

$$qx = A(p)x + B(p)u, \quad (2a)$$

$$y = C(p)x + D(p)u, \quad (2b)$$

^a $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is a real meromorphic function if $h = f/g$ with f, g analytic and $g \neq 0$.

where $x : \mathbb{Z} \rightarrow \mathbb{R}^{n_x}$ is the state variable and A, B, C, D , with appropriate dimensions, are rational matrix functions of p , bounded on \mathbb{P} . The LFR of \mathcal{S} , denoted by $\mathfrak{R}_{\text{LFR}}(\mathcal{S})$, is defined as

$$\begin{bmatrix} qx \\ z \\ y \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix}, \quad (3a)$$

where $\{A, \dots, D_{22}\}$ are constant matrices with appropriate dimensions and

$$w(k) = \Delta(p(k))z(k), \quad (3b)$$

with $\Delta : \mathbb{P} \rightarrow \mathbb{R}^{n_w \times n_z}$ being the (linear) function of p . Commonly, Δ has a block diagonal structure and it is assumed to vary in a *polytope*. Additionally, x, w, z are latent (auxiliary) variables of $\mathfrak{R}_{\text{LFR}}(\mathcal{S})$.

A particular drawback of the SS and the LFR forms is that the estimation of states or latent variables together with the underlying matrix coefficients under noisy measurements of y , like $y(k) + v(k)$ where $v(k)$, is stochastic noise process (not necessary white), is difficult, commonly requiring simplifying assumptions and approximations. Due to this complexity, stochastic implications of estimation are not well understood and the curse of dimensionality casts a constant shadow over these approaches. Usually the computational and parametrization simplicity can not be exploited yet as much as in the LTI case, thus often only simple dependencies like static and linear are considered. Furthermore, estimation is effected by non-uniqueness of the parametrization in both the global and local settings, which often causes interpolation to be unpredictable in the SS case [Tóth (2010); Tóth *et al.* (2011e)]. However a major advantage of the SS and LFR representation based approaches is that the delivered models are ready for control synthesis without further processing and a state-space representation, like in the LTI case, is efficient to describe MIMO relations.

An other important class of representations are IO representations in terms of polynomial forms, denoted by $\mathfrak{R}_{\text{IO}}(\mathcal{S})$, which express the IO signal relations in their natural difference equation form:

$$y(k) = - \sum_{i=1}^{n_a} (a_i \diamond p)(k) q^{-i} y(k) + \sum_{j=0}^{n_b} (b_j \diamond p)(k) q^{-j} u(k). \quad (4)$$

The coefficients a_i and b_i are often assumed to be static or dynamic functions of p , representing from simple linear (affine) to rational or meromorphic dependence, as this setting is flexible to handle complicated dependences. Particularly attractive features of these structures are that their

identification can be addressed via the extension of the LTI *prediction-error* (PE) framework [Tóth (2010); Tóth *et al.* (2011b)] enabling the stochastic analysis of the estimates, treatment of general noise models [Tóth *et al.* (2011b); Laurain *et al.* (2010)], experiment design [Dankers *et al.* (2011); Wei and Del Re (2006); Khalate *et al.* (2009)], model structure selection and direct identification of the involved dependencies [Tóth *et al.* (2011c, 2009b); Hsu *et al.* (2008)] often in a computationally attractive manner. However, such model structures have a serious disadvantage: the delivered IO model needs to be converted to an SS or an LFR form as the main stream LPV control-synthesis approaches are formulated in terms of these representations. Due to the fact that multiplication with q in Eq. (4) is not commutative over the p -dependent coefficients a_i and b_j , the involved realization theory is more complicated than in the LTI case and often introduces complicated rational dynamic dependence on p in the resulting SS forms [Tóth *et al.* (2011e)]. Even if there exist some strategies to avoid such complications in specific situations (see [Tóth *et al.* (2011a)]), the burden of the realization is likely to rise difficulties in applications.

With respect to the previously mentioned representation forms, a particularly attractive model structure in the LPV case follows by the truncation of Eq. (1) to a finite number of expansion terms:

$$y(k) = \sum_{i=0}^n (h_i \diamond p)(k) u(k-i), \quad (5)$$

which is the LPV form of the well-known LTI *finite impulse response* (FIR) models. Such models have attractive properties in terms of identification in opposite to the challenging identification problem of (2a-b) or (3a-b). In particular, they benefit from the advantages of IO models as their identification can be addressed via the PE framework. An important property of Eq. (5) is *linearity-in-the-coefficients* that allows to use *linear regression* for the estimation of $\{h_i\}_{i=1}^n$ if they are linearly parameterized:

$$(h_i \diamond p)(k) = \sum_{j=0}^{n_i} \theta_{i,j} (f_{i,j} \diamond p)(k), \quad (6)$$

where $\theta_{i,j} \in \mathbb{R}^{n_y \times n_u}$ are the unknown parameters and $f_{i,j}$ are prior selected functions. Furthermore, noise or disturbances in the system can be modeled in an *output error* (OE) sense with this model structure, which allows independent parametrization of the noise model. However, a well known disadvantage of FIR models, both in the LTI and the LPV cases,

is that the expansion may have a slow convergence rate, meaning that it requires a relatively large number of parameters for an adequate approximation of the system. In order to benefit from the same properties, but achieve faster convergence rate of the expansion, it is attractive to use basis functions which, opposite to q^{-i} , have infinite impulse responses. A particular choice of such a basis follows through the use of *orthonormal basis functions* (OBF's), which are specific basis functions in \mathcal{H}_2 (Hardy space of square integrable complex functions) and have already proven their usefulness in LTI identification (see [Heuberger *et al.* (2005)]) . This is the idea of the representation we would like to use to formulate expansion-based model structures for LPV identification which are beneficial both from the estimation and the utilization point of view. As we will see, the proposed structures represent an attractive trade-off between SS and IO forms, combining the advantages of both representations based model structures.

3. Orthonormal basis function models

In this section we explore the possibilities for using series-expansion model structures for LPV systems, using the concept of OBF's. A major motivation is the linear-in-the-coefficients property of these structures, which is very beneficial in PE identification. A second merit of these structures is that they allow a relatively simple interpolation of local linear models with varying McMillan degree and they have a direct SS and LFR representation. Furthermore it was shown in [Boyd and Chua (1985)] that models composed from an OBF filter bank followed by a static nonlinearity are general approximators of nonlinear systems with fading memory.

3.1. Series-expansion representations

We will start to develop these expansions and the concepts of OBF's by following a local point of view. It is well known that an LPV system has an LTI behavior if the considered scheduling trajectory is constant, i.e. $p(k) \equiv \mathbf{p}$. Thus such a frozen aspect of the system can be represented by a transfer function $F_{\mathbf{p}}(z)$ with $z \in \mathbb{C}$ being the complex frequency. If $F_{\mathbf{p}} \in \mathcal{H}_2^{n_y \times n_u}$, then $F_{\mathbf{p}}$ can be written as

$$F_{\mathbf{p}}(z) = W_0 + \sum_{i=1}^{\infty} W_i \phi_i(z), \quad (7)$$

where $\{\phi_i\}_{i=1}^{\infty}$ is a basis for \mathcal{H}_2 and $W_i \in \mathbb{R}^{n_y \times n_u}$. In the theory of *generalized orthonormal basis functions* (GOBF's), the functions $\phi_i(z)$ are generated by applying *Gram-Schmidt orthonormalization* to the sequence of functions

$$\frac{1}{z - \xi_1}, \frac{1}{z - \xi_2}, \dots, \frac{1}{z - \xi_{n_g}}, \frac{1}{(z - \xi_1)^2}, \frac{1}{(z - \xi_2)^2}, \dots \quad (8)$$

with stable pole locations $\xi_1, \dots, \xi_{n_g} \in \mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$. The choice of these *basis poles* determines the rate of convergence of the series expansion (7). Note that it is possible to also develop such an expansion using basis functions in $\mathcal{H}_2^{n_y \times n_u}$ but for the sake of simplicity we only consider here the so called scalar case. For more on multidimensional bases see [Heuberger *et al.* (2005)]. An alternative derivation of the basis functions is based on a balanced LTI realization $\{A_g, B_g, C_g, D_g\}$ of the *inner function*

$$G_g(z) = \prod_{i=1}^{n_g} \frac{1 - z\xi_i^*}{z - \xi_i}, \quad (9)$$

where $\{\phi_i(z)\}_{i=0}^{\infty}$ are the scalar elements of the vector functions

$$(zI - A_g)^{-1} B_g G_g^j(z), \quad j \in \mathbb{N}. \quad (10)$$

By using a truncated expansion in Eq. (7), an attractive OBF model structure for LTI identification results, with a well worked-out theory in terms of variance and bias expressions [Heuberger *et al.* (2005)]. The series expansion (7) can be extended to LPV systems (see [Tóth (2010)]), via the expansion of each q^{-i} in Eq. (1) in terms of $\{\phi_i\}_{i=1}^{\infty}$. Thus, an LPV system can be written as

$$y(k) = (W_0 \diamond p)(k)u(k) + \sum_{i=1}^{\infty} (W_i \diamond p)(k)\phi_i(q)u, \quad (11)$$

where W_i are matrix functions with dynamic dependence on p . An obvious choice of model structure is to use a truncated expansion, i.e. truncating Eq. (11) to a finite sum in terms of $\{\phi_i\}_{i=1}^n$:

$$y(k) \approx (W_0 \diamond p)(k)u(k) + \sum_{i=1}^n (W_i \diamond p)(k)\phi_i(q)u. \quad (12)$$

Note that these expansions are formulated in the time domain (using the shift operator q), as there exists no frequency-domain expression for LPV systems. Similar to the FIR case, this structure is linear in the coefficients $\{W_i\}_{i=1}^n$. Furthermore, it is proven that structures like (12), i.e. an OBF

filter bank followed by a static nonlinearity, are general approximators of nonlinear systems with fading memory, i.e. nonlinear dynamic systems with convolution representation [Boyd and Chua (1985)]. An important question that arises is whether the basis functions ϕ_i can be chosen such that a fast convergence rate can be achieved for all possible trajectories of p , i.e. how $\{\phi_i(q)\}_{i=1}^n$ with minimal n should be chosen such that the approximation error is adequate for the problem at hand.

It is important to note that in case of an unstable LPV system, it is possible to factorize the system representation into a stable and unstable part by using co-prime factorization [Wood *et al.* (1996)]. A convergent impulse response representation of the unstable part can be characterized in terms of the basis $\{q^i\}_{i=1}^\infty$ and hence OBF models, like (12), of such systems can be formulated in a two-sided expansion.

3.2. Basis selection

In the previously introduced modeling concept, it has a prime importance to achieve an efficient selection of the basis $\{\phi_i\}_{i=1}^\infty$, which provides a fast convergence rate of the series-expansion (11). This makes it possible to capture the dynamics of the system with a small n in Eq. (12). By taking a closer look at Eq. (12), an important implication is that if p is constant, i.e. $p(k) \equiv \mathbf{p}$, then the error of the approximation for a given n depends on the expansion error of the frozen LTI transfer function $F_{\mathbf{p}}$ in terms of $\{\phi_i(q)\}_{i=1}^n$. This means that to achieve a fast convergence rate, i.e. small approximation error by Eq. (12), it is necessary to choose $\{\phi_i(q)\}_{i=1}^n$ such that the maximum expansion error of $F_{\mathbf{p}}$ for all $\mathbf{p} \in \mathbb{P}$ is minimal. Even if such a condition is not sufficient [Tóth *et al.* (2009a)], it gives an important tool for efficient basis selection in terms of the classical *Kolmogorov n-width* result of [Pinkus (1985)] extended to OBFs in [Oliveira e Silva (1996)]. This result states that for a given LTI *inner function* G_g with poles $\Xi_{n_g} \subset \mathbb{D}$, the OBF's generated by G_g (see Section 3.1) are optimal in the n -width sense for the set of LTI systems having poles in the region

$$\Omega(\Xi_{n_g}, \rho) = \{z \in \mathbb{D} \mid |G_g(z^{-1})| \leq \rho\}. \quad (13)$$

Here ρ is the rate of convergence in the series expansion, and n should be a multiple of the number of basis poles n_g in Ξ_{n_g} . See Fig. 1, taken from [Tóth *et al.* (2009a)], for an example of these regions.

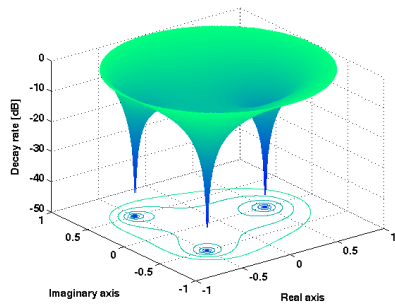


Fig. 1. Example of the function $|G_g(z^{-1})|$ and the region $\Omega(\Xi_n, \rho)$ for an inner function G_g with 3 poles and various values of ρ (decay rate). Note that if z_0 is a pole of G_g , then $G_g(z_0^{-1}) = 0$.

For the basis-selection problem we are dealing with the inverse problem, i.e. given a region of poles $\Omega_* = \{\xi \in \mathbb{C} \mid \exists p \in \mathbb{P} \text{ s.t. } \xi \text{ is a pole of } F_p(z)\}$, approximate this region as

$$\Omega_* \subseteq \Omega(\Xi_n, \rho) = \{z \in \mathbb{D} \mid |G_g(z^{-1})| \leq \rho\}. \quad (14)$$

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

$$\min_{\Xi_n \subset \mathbb{D}} \rho = \min_{\Xi_n \subset \mathbb{D}} \max_{z \in \Omega_*} |G_g(z^{-1})| = \min_{\Xi_n \subset \mathbb{D}} \max_{z \in \Omega_*} \left| \prod_{k=1}^n \frac{1 - z\xi_k^*}{z - \xi_k} \right|. \quad (15)$$

For a given $\Xi_n = \{\xi_1, \dots, \xi_n\}$ and Ω_* , the region $\Omega(\Xi_n, \rho_*)$ with ρ_* is the minimum of ρ such that (14) is satisfied, is called the *Kolmogorov bound* of Ξ_n . Smallness of this region w.r.t. Ω_* and ρ_* together indicate how well the basis functions are tuned w.r.t. Ω_* .

In order to select an efficient basis, it is obviously required to obtain knowledge about the system to be modeled. In the LPV case, this knowledge is Ω_* : the set of poles of all possible local linear models. 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 simple selection scheme, which delivers adequate results in practice, is given by the following steps:

- (1) Identify a number of local linear models in several different operating regimes p_i , i.e. using data with a constant scheduling signal $p(k) \equiv p_i$.
- (2) Plot all poles of the identified models in the complex plane.
- (3) Cluster the poles in groups and find optimal cluster centers (these centers will be used as basis poles) which minimize (15).

Alternatively, other basis optimization schemes based on recently developed sparse estimators can also be used via *orthogonal matching pursuit* or ℓ_1 -optimization approaches [Tropp and Wright (2010)]. Basis selection is recently developing.

In the next section we present an efficient approach to obtain a simultaneous solution for the problems of reconstructing Ω_* from experimental data and the minimization of the Kolmogorov measure.

3.3. A fuzzy clustering approach

Objective-function-based fuzzy clustering algorithms, such as *fuzzy c-max* (FcM) clustering, have been used in a wide collection of applications [Bezdek (1981); Kaymak and Setnes (2002)]. Generally, FcM partitions the data into overlapping groups to capture the underlying structure [Jain and Dubes (1988)]. In this section we describe the extension of the classical FcM approach to the so-called *Fuzzy-Kolmogorov c-Max* (FKcM) algorithm, originally developed in [Tóth *et al.* (2009a)], which enables the determination of the region Ω_* on the basis of observed frozen poles with membership based, overlapping areas. We assume that we are given a set of observed/identified poles $Z = \{z_1, \dots, z_N\} \subset \Omega_*$.

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. Denote $\mathbb{I}_{\tau_1}^c = \{n \in \mathbb{N} \mid \tau_1 \leq n \leq \tau_2\}$. 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* $0 < \varepsilon \leq 1$, we obtain a set

$$\Omega_\varepsilon = \{z \in \mathbb{D} \mid \exists i \in \mathbb{I}_1^c, \mu_i(z) \geq \varepsilon\}. \quad (16)$$

To measure dissimilarity of Z with respect to each cluster, we introduce distances $d_{i,k} = \kappa(v_i, z_k)$ between v_i and z_k , where κ , defined by

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

is a metric on \mathbb{D} , referred to as the *Kolmogorov metric*.

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

$$\sum_{i=1}^c \mu_{i,k} = 1 \quad \text{for } \forall k \in \mathbb{I}_1^N \quad \text{and} \quad 0 < \sum_{k=1}^N \mu_{i,k} \quad \text{for } \forall i \in \mathbb{I}_1^c.$$

With these preliminaries we can now formulate the problem we will consider.

Problem 10.1. For a set of pole locations Z and for a given number of clusters c , find a set of cluster centers $\{v_i\}_{i=1}^c$, a set of membership functions $\{\mu_i\}_{i=1}^c$, and the maximum of ε , such that

- Ω_ε contains Z and it describes the underlying distribution of Z in terms of a chosen dissimilarity measure κ .
- With respect to Ω_ε , the OBF's, with poles Ξ_c in the cluster centers $\{v_i\}_{i=1}^c$, are optimal in the KnW sense, where $n = c$ and with the corresponding decay rate ρ as small as possible.

Let $V = [v_1 \dots v_c]^\top$ and $U_k = [\mu_{1,k} \dots \mu_{c,k}]^\top$ and denote by U the matrix with U_k 's as columns. Fuzzy clustering can be viewed as the minimization of the FcM-functional [Bezdek (1981)], \mathcal{J}_m , which in the FKcM case can be formulated as

$$\mathcal{J}_m(V, U) = \max_{k \in \mathbb{I}_1^N} \sum_{i=1}^c \mu_{i,k}^m d_{i,k}. \quad (18)$$

Here the design parameter $m \in (1, \infty)$ defines the *fuzziness* of the resulting partition in the sense that m determines how sharp the separation is between the clusters by the membership functions. For $1 \ll m$, each μ_i flattens till all poles in Z belong to all clusters with equal membership, while for $m = 1$, each $z \in Z$ belongs to only one cluster with a nonzero membership. It can be observed that Eq. (18) corresponds to a *worst-case (max) sum-of-error* criterion, contrary to the *mean-squared-error* criterion of the original FcM, see [Bezdek (1981)].

The crucial property of this functional is that it can be shown [Tóth *et al.* (2009a)] that for large values of m minimization of J_m is equivalent to the Kolmogorov measure minimization problem (15):

Theorem 1. Given a data set $Z \subset \mathbb{D}$ with N elements, and a vector of cluster centers $V \in \mathbb{D}^c$, such that $d_{i,k} = \kappa(v_i, z_k) \neq 0$ for all $(i, k) \in \mathbb{I}_1^c \times \mathbb{I}_1^N$. Define U_m as a membership matrix of V minimizing Eq. (18) for $m > 1$. Then $\mathcal{J}_m(U_m, V) = c^{1-m} \max_{k \in \mathbb{I}_1^c} (\prod_{i=1}^c d_{i,k})^{1/c} + \mathcal{O}(e^{-m})$. Furthermore, $\mathcal{J}_m(U_m, V)$ decreases monotonically with m , and $\lim_{m \rightarrow \infty} \mathcal{J}_m(U_m, V) = 0$.

This theorem gives that, by solving Problem 10.1 via the solution of Eq. (18), which can be obtained in terms of an alternating optimization (*Picard iteration*) [Tóth *et al.* (2009a)], one can simultaneously cluster the observed poles in such a way that the resulting cluster centers will approximate arbitrary well (governed by m) the n -width optimal OBF poles for the reconstructed frozen pole regions. See Fig. 2 for an example of the basis selection mechanism. For further details as well as a description of the optimization algorithm see [Tóth (2010)], where also the robust extension of the basis selection is discussed. This robust extension allows to solve the optimal basis selection problem when the local pole estimates are given up to an uncertainty region due to the effects of the measurement noise.

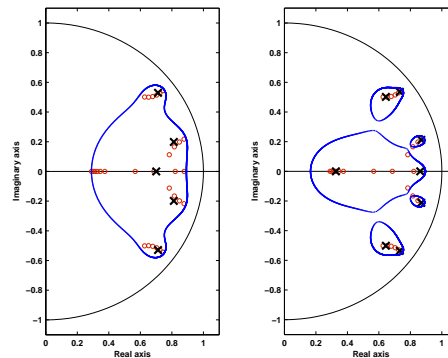


Fig. 2. Example of the basis selection procedure, using fuzzy clustering with fuzziness parameter $m = 8$. The 30 observed poles (i.e. the set Z) are given with circles. The resulting cluster centers are depicted with a black x. The lines represent the Kolmogorov bound $\Omega(\Xi_c, \rho_*)$ w.r.t. Z . On the left hand side $c = 5$ clusters are determined, on the right hand side $c = 8$.

For the determination of the actual number of clusters in these algorithms, *adaptive cluster merging* can be applied (see [Kaymak and Setnes (2002); Tóth *et al.* (2009a)]). Starting from an 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.

3.4. OBF's-based model structures

Now we can define the OBF model structures we intend to use for addressing the identification of LPV systems. Assume that the basis selection step has been completed and we are given a set of n_f basis functions $\{\phi_i(z)\}_{i=1}^{n_f}$ and the data-generating LPV system \mathcal{S}_o is affected by a stochastic disturbance $v(k)$ in an output additive sense. Note that v can represent a wide variety of noise processes from pure (white) measurement noise to process noise correlated with y , p and/or u . The input-output dynamics of a truncated

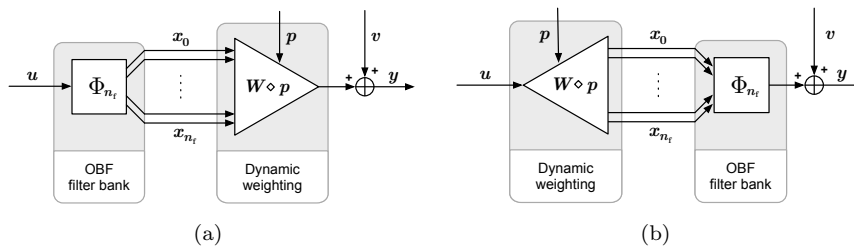


Fig. 3. IO signal flow graph of (a) the W-LPV OBF model described by (19) and (b) the H-LPV OBF model described by (20).

LPV series-expansion model in terms of $\{\phi_i(z)\}_{i=1}^{n_f}$ can now be written as

$$y(k) = \sum_{i=0}^{n_f} (W_i \diamond p)(k) \phi_i(q) u(k) + v(k), \tag{19}$$

with $\phi_0(\bullet) = 1$. Equation (19) corresponds to a so-called *output error* (OE) model. Introduce $\Phi_{n_f} = [\phi_1 \dots \phi_{n_f}]^\top$ and $W = [W_1 \dots W_{n_f}]^\top$. Then the model structure (19) can be visualized as in Fig. 3a, where $x_i(k) = \phi_i(q)u(k)$. 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. 3b, is the so-called *Hammerstein LPV OBF* model:

$$y(k) = \sum_{i=0}^{n_f} \phi_i(q) (W_i \diamond p)(k) u(k) + v(k). \tag{20}$$

The truncated expansion (20) can be obtained by deriving the series expansion (11) such that $\{W_i\}_{i=0}^\infty$ appear after the basis $\{\phi_i\}_{i=0}^\infty$. Such a reordering has no effect in the LTI case, but for LPV systems, due to the non-commutativity of multiplication of any p -dependent coefficients with q^{-1} , it results in a set of different expansion coefficients. Even if such expansions are equal in the asymptotic sense, in case of a finite truncation they have different approximation capabilities (see [Tóth (2010)]).

In the sequel we will restrict attention to the Wiener model structure. A particularly interesting feature of Eq. (19) is that it can be written in a state-space form

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

$$y = (W \diamond p)x + (W_0 \diamond p)u + v, \tag{21b}$$

where the constant matrices A and B are completely determined by the basis functions $\{\phi_i\}_{i=1}^{n_f}$. This illustrates that the dependency on p is only present in the output equation (21b). The direct SS realization (21a-b) avoids complications that are present in the general IO case and at the same time allows to directly control the resulting dependency of the SS model by the chosen parametrization of each W_i . Similarly, direct LFR realization schemes of these models also exist in case the dependencies of each expansion coefficients are polynomial. This is a clear benefit over IO models. Moreover, as we will see, we can preserve all the advantages of the IO setting for the identification of (19) which is attractive compared to SS identification.

4. Identification via OBF models

In this section, the identification of the previously introduced OBF model structure (19) is discussed in a PE setting. With respect to the actual estimation of (19) we distinguish between two methods: a local and a global approach, which not only correspond to different estimation concepts but, as we will see, can also result in rather different model estimates. However, to derive these estimation schemes, first the question of parametrization of the coefficients is investigated.

4.1. Parametrization of the coefficient dependence

Beside the selection of basis functions in the considered LPV-OBF setting, model structure selection also contains an equally important part: the parametrization of the dependence of the coefficients W_i on p :

$$W_i \diamond p = \psi_i(\theta) \diamond p, \quad (22)$$

where ψ_i is function defined via some constant parameters $\theta \in \mathbb{R}^n$. The aim of the identification is to estimate θ based on a measured data record.

To simplify the estimation problem, in the LPV literature often a linear parametrization of the structural dependence is used. In fact, coefficients like W_i are considered to be a linear combination of fixed matrix functions $\psi_{i,j} : \mathbb{P} \rightarrow \mathbb{R}^{n_y \times n_u}$:

$$W_i = \sum_{j=0}^{n_\psi} \theta_{i,j} \odot \psi_{i,j}, \quad (23)$$

where $\psi_{i,0} = 1$, $\theta_{i,j} \in \mathbb{R}^{n_y \times n_u}$, and \odot denotes the *Hadamard*, i.e. element-by-element, matrix product. A linear parametrization not only reduces the complexity of the associated estimation problem but also makes the problem of adequate selection of the underlying structural dependence well-posed [Tóth (2010)]. In terms of Eq. (23), the selection problem of an adequate parametrization translates to a search for a set of functions $\{\psi_i\}_{i=0, j=0}^{n_f, n_\psi}$ such that the true expansion coefficients W_i^o of the system with respect to the used basis functions $\{\phi(q)\}_{i=1}^{n_\psi}$ satisfy $W_i^o \in \text{Span}(\{\psi_{i,j}\}_{j=0}^{n_\psi})$. In case of a black-box scenario, the choice of $\{\psi_{i,j}\}$ can be arbitrary. One can consider all $\{\psi_{i,j}\}$ to be rational functions or polynomials with a fixed degree and a fixed order of dynamic dependence. However the number of possible choices is enormous. Including a too large set of functions $\{\psi_{i,j}\}$ can easily lead to over-parametrization, while restriction of $\{\psi_{i,j}\}$ to only a few basic functions can lead to serious structural bias. In order to assist the selection of an efficient set of functional dependencies in the parametrization of linear regression models, recently practically applicable approaches have been proposed in [Tóth *et al.* (2011c, 2009b)] and [Hsu *et al.* (2008)]. In [Hsu *et al.* (2008)] a *dispersion functions* based method while in [Tóth *et al.* (2011c)] a *support vector machine* approach, both originating from the *machine learning* field, has been developed to basically learn the underlying static or dynamic nonlinear dependence of the coefficients with great efficiency. In [Tóth *et al.* (2009b)] a coefficient shrinkage method, the so-called *non-negative garrote* (NNG) approach originating from statistics, has been introduced for this purpose. The NNG uses regularization in terms of weights to penalize individual elements of the parameter vector θ . In this way, the approach starts with a relatively large set of possible functional dependencies from which those functions that do not contribute significantly to the validity of the estimated model are eliminated by decreasing their weights.

4.2. Prediction error concept

Based on the considered parametrization of $\{W_i\}_{i=0}^{n_f}$ given by Eq. (23), the deterministic part of (19) can be written in the operator form

$$G(q, \theta) = \sum_{i=0}^{n_f} \sum_{j=0}^{n_\psi} (\theta_{i,j} \odot \psi_{i,j}) \phi_i(q), \quad (24)$$

with an overall parameter vector $\theta \in \mathbb{R}^{n_\theta}$ containing the elements of $\theta_{i,j}$.

Similarly, we can introduce a parametrized noise model w.r.t. v such as

$$v(k) = (H(q, \theta) \diamond p) e(k), \quad (25)$$

where e is a zero mean white noise process,

$$H(q, \theta) = \sum_{i=0}^{\infty} (h_i(\theta) \diamond p) q^{-i}, \quad (26)$$

is the IRR of a parametrized asymptotically stable LPV filter and θ also contains the parameters associated with this noise model. Then, the overall model $\mathcal{M}_\theta = (G(q, \theta), H(q, \theta))$ is represented as

$$y_\theta(k) = (G(q, \theta) \diamond p)(k) u(k) + (H(q, \theta) \diamond p)(k) e(k). \quad (27)$$

If there is a θ_o such that the data-generating system \mathcal{S}_o satisfy that $\mathcal{S}_o = \mathcal{M}_{\theta_o}$, then the aim of identification is to estimate θ_o based on measured data \mathcal{D}_N . In any other case, an \mathcal{M}_θ is searched for that describes the behavior of \mathcal{S}_o , the best in terms of a given criterion function. To simplify the following discussion, in the sequel we treat only the case of $H(q, \theta) = I$ and e being a vector of independent zero white noise processes. However, we will briefly return to the concept of general noise models later.

It is possible to show (see [Tóth (2010); Tóth *et al.* (2011b)]) that w.r.t. (19) with $H(q, \theta) = I$, the conditional expectation of $y_\theta(k)$ in the ℓ_2 sense under the information set of $\{u(\tau)\}_{\tau=1}^k$ and $\{p(\tau)\}_{\tau=1}^k$ is equal to

$$\hat{y}_\theta(k|k-1) := (G(q, \theta) \diamond p)(k) u(k). \quad (28)$$

Then the basic philosophy of PE based identification is that w.r.t. a given model set $\mathcal{M} = \{\mathcal{M}_\theta \mid \theta \in \mathbb{R}^{n_\theta}\}$ with parameter space $\Theta \subseteq \mathbb{R}^{n_\theta}$ and a data set \mathcal{D}_N , to find θ such that the *one-step-ahead predictor* (28) associated with \mathcal{M}_θ provides a *prediction error*

$$e_\theta(k) = y(k) - \hat{y}_\theta(k|k-1), \quad (29)$$

which resembles a zero mean white noise “as much as possible”.

Based on the predictor form (28), many different classical identification criteria can be applied to meet the above goal. A particularly interesting choice is the *least-squared* (LS) prediction-error criterion

$$\mathcal{V}(\theta, \mathcal{D}_N) = \frac{1}{N} \sum_{k=1}^N e_\theta^\top(k) e_\theta(k) = \frac{1}{N} \|e_\theta\|_{\ell_2}^2 \quad (30)$$

where the residual $e_\theta(k)$ is given by (29). As we will see, linear parametrization of the expansion coefficients (see Eq. (23)) yields that w.r.t. (30), the estimation of θ reduces to a linear regression problem. In other cases, when the parametrization of the coefficients is nonlinear, the estimation corresponds to a nonlinear optimization problem.

To guarantee a unique solution of (30), one condition is that the set of functions $\{\psi_{i,j}\}$ are chosen such that \mathcal{M}_θ is globally identifiable:

Definition 10.1. The model structure $(G(q, \theta), H(q, \theta))$, defined by Eq. (27), with a parameter domain $\Theta \subseteq \mathbb{R}^{n_\theta}$ and with $H(q, \theta) = I$ is globally identifiable, if for any $\theta_1, \theta_2 \in \Theta$, the corresponding one-step-ahead predictors (see Eq. (28)) are distinguishable:

$$G(q, \theta_1) = G(q, \theta_2) \Rightarrow \theta_1 = \theta_2.$$

In terms of the 1-step ahead predictor (28) of the considered OBF model, a necessary and sufficient condition to guarantee identifiability is that in the parameterization, each function set $\{\psi_{i,j}\}_{j=0}^{n_\psi}$ is a set of linearly independent functions. Another condition for the unique solution of (30), is the *informativity* of the data set \mathcal{D}_N :

Definition 10.2. For a model structure $(G(q, \theta), H(q, \theta))$, defined by Eq. (27), with a parameter domain $\Theta \subseteq \mathbb{R}^{n_\theta}$ and with $H(q, \theta) = I$, a data set $\mathcal{D}_N = \{u(k), y(k), p(k)\}_{k=1}^N$ is called informative, if the following holds for the one-step-ahead predictors (see Eq. (28)):

$$(G(q, \theta_1) \diamond p)u = (G(q, \theta_2) \diamond p)u \Rightarrow G(q, \theta_1) = G(q, \theta_2).$$

This means that if \mathcal{D}_N is informative w.r.t. a globally identifiable \mathcal{M}_θ , then the global optimum of (30) is unique. Now, having the model structure and the required concepts of identification established, we can introduce identification schemes of W-LPV-OBF models.

4.3. Local estimation approach

As a first approach, we aim at the identification of (19) based on the frozen aspects of the underlying data-generating system. This so-called local approach uses a number N_{loc} of “local” experiments, i.e. data collection with a constant scheduling signal $p(k) \equiv p_\tau \in \mathbb{P}$, resulting in data sequences

$\mathcal{D}_{N,p_\tau} = \{u(k), y(k), p_\tau\}_{k=1}^N$ for $\tau \in \mathbb{I}_1^{N_{\text{loc}}}$. Based on these data, N_{loc} LTI-OBF models defined as

$$G_\tau(q, \eta_\tau) = \sum_{i=0}^{n_f} W_{i,\tau} \phi_i(q), \quad H(q, \eta_\tau) = I, \quad (31)$$

with $\eta_\tau = [W_{0,\tau} \dots W_{n_f,\tau}] \in \mathbb{R}^{n_y \times (n_u \cdot (n_f+1))}$, are estimated using the LS criterion. Note that – under the condition that the data sets are informative – there exist unique analytic solutions to these estimation problems. The estimated coefficients can now be considered as “samples” of the function $W_i \diamond p$, in the sense that $W_i \diamond p$ with $p(k) \equiv p_\tau$ is equal to $W_{i,\tau}$. As a second step, we use interpolation to obtain estimates of the function $W_i \diamond p$, for instance by assuming a polynomial dependency of $(W_i \diamond p)(k)$ on $p(k)$, or by making use of splines etc. Note that as the only information available about the system is in the form of \mathcal{D}_{N,p_τ} where p is constant, the estimation assumes that each W_i has only static dependence on p .

Let $\text{Vec}(\cdot)$ denote *vectorization* of a matrix in the sense that

$$M = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \Rightarrow \text{Vec}(M) = [\alpha_{1,1} \ \alpha_{1,2} \ \alpha_{2,1} \ \alpha_{2,2}]^\top.$$

Furthermore, introduce the diagonal matrix row construction as

$$\text{Diag}_{\text{row}}(M) = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & 0 & 0 \\ 0 & 0 & \alpha_{2,1} & \alpha_{2,2} \end{bmatrix}.$$

Then the local concept of estimation is formalized in terms of Algorithm 10.1.

4.4. Global estimation approach

Opposite to the local approach, in the global case we aim at the estimation of (19) in terms of the parametrized model structure (27) with $H(q, \theta) = I$ using a data set \mathcal{D}_N where p is varying. This data set is assumed to be informative w.r.t. (27). Then, in terms of the LS criterion, a unique analytic solution – under the condition that Eq. (27) is globally identifiable – can be obtained for θ via Algorithm 10.2.

A similar algorithm can be introduced for the identification of the Hammerstein LPV-OBF model structure (20), including the estimation of initial conditions, see [Tóth (2010)] for a detailed discussion. In case the noise model is parameterized in terms of an LPV filter, i.e. $H(q, \theta) \neq I$, the

Algorithm 10.1 OBF based LPV identification, local method

Input: an OBF set $\Phi_{n_f} = \{\phi_i\}_{i=0}^{n_f}$ with $\phi_0(\cdot) = 1$, scheduling points $\mathcal{P} = \{\mathbf{p}_\tau\}_{\tau=1}^{N_{\text{loc}}} \subset \mathbb{P}$, data records $\mathcal{D}_{N, \mathbf{p}_\tau} = \{u(k), y(k), \mathbf{p}_\tau\}_{k=1}^N$ of \mathcal{S} , an identification criterion \mathcal{V} , and the local OBF model structure (31) with $\eta_\tau = [W_{0, \tau} \dots W_{n_f, \tau}]$. Assume that each $\mathcal{D}_{N, \mathbf{p}_\tau}$ is informative w.r.t. (31).

- 1: for each $\tau \in \mathbb{I}_1^{N_{\text{loc}}}$, calculate $\hat{\eta}_\tau = \arg \min_{\eta_\tau} \mathcal{V}(\eta_\tau, \mathcal{D}_{N, \mathbf{p}_\tau})$.
- 2: choose a set of matrix functions $\{\psi_{i,j} : \mathbb{P} \rightarrow \mathbb{R}^{n_y \times n_u}\}_{i=0, j=0}^{n_f, n_\psi}$ where the scalar functions $\{[\psi_{i,j}]_{k,l}\}_{j=0}^{n_\psi}$ are linearly independent for each (i, k, l) and all $\hat{\theta}_{i,j} = \arg \min_{\theta_{i,j} \in \mathbb{R}^{n_y \times n_u}} \|\hat{\Gamma}_i - \sum_{j=0}^{n_\psi} \Psi_{i,j} \text{Vec}(\theta_{i,j})\|$ achieves the least interpolation error in terms of the considered norm $\|\cdot\|$, where $\hat{\Gamma}_i = [\hat{W}_{i,1}^\top \dots \hat{W}_{i, N_{\text{loc}}}^\top]^\top$ and $\Psi_{i,j} = [\text{Diag}_{\text{row}}(\psi_{i,j}(\mathbf{p}_1))^\top \dots \text{Diag}_{\text{row}}(\psi_{i,j}(\mathbf{p}_{N_{\text{loc}}}))^\top]^\top$.
- 3: **return** estimated model (27).

estimation problem becomes non-linear in θ . If $v(k)$ is not white or correlated with u , then Algorithm 10.2 results in a biased estimate. To handle such cases it is possible to use a recently developed *instrumental variable* approach [Laurain *et al.* (2010)], which can provide statistically efficient estimation for linear regression models such as (32).

Algorithm 10.2 OBF based LPV identification, global method

Input: an OBF set $\Phi_{n_f} = \{\phi_i\}_{i=0}^{n_f}$ with $\phi_0(\cdot) = 1$, matrix functions $\{\psi_{i,j} : \mathbb{P} \rightarrow \mathbb{R}^{n_y \times n_u}\}_{i=0, j=0}^{n_f, n_\psi}$ where the scalar functions $\{[\psi_{i,j}]_{k,l}\}_{j=0}^{n_\psi}$ are linearly independent for each (i, k, l) , a data record $\mathcal{D}_N = \{u(k), y(k), p(k)\}_{k=1}^N$ of \mathcal{S} , an identification criterion \mathcal{V} , and the OBF model structure:

$$G(q, \theta) = \sum_{i=0}^{n_f} \sum_{j=1}^{n_\psi} (\theta_{i,j} \odot \psi_{i,j}) \phi_i(q), \quad H(q, \theta_\tau) = I, \quad (32)$$

with $\theta = [\theta_{0,0} \dots \theta_{n_f, n_\psi}]$. Assume that \mathcal{D}_N is informative.

- 1: calculate the signals $x_{i,j} = \psi_{i,j} \odot (1_{n_y} \otimes \phi_i(q) u^\top)$ and let $\Gamma = \text{Diag}_{\text{row}}([x_{0,0} \dots x_{n_g, n_\psi}])$ giving that $y(k) = \Gamma(k) \text{Vec}(\theta) + \varepsilon_\theta(k)$.
- 2: estimate θ in terms of $\hat{\theta} = \arg \min_{\theta} \mathcal{V}(\theta, \mathcal{D}_N)$. In case of (30), $\hat{\theta} = (\frac{1}{N} \Gamma_d^\top \Gamma_d)^{-1} (\frac{1}{N} \Gamma_d^\top Y)$ with $Y = [y^\top(1) \dots y^\top(N)]^\top$ and $\Gamma_d = [\Gamma^\top(1) \dots \Gamma^\top(N)]^\top$.
- 3: **return** estimated model (27).

4.5. *Properties of estimation*

Similar to the classical LTI identification framework, it is possible to show that under minor conditions, the parameter estimates of local and global W-LPV approaches are convergent and consistent. Convergence means that for $N \rightarrow \infty$ the parameter estimate $\hat{\theta}$ converges, i.e. $\hat{\theta} \rightarrow \theta_*$, with probability 1, while consistency means that the convergence point θ_* is equal to θ_o , the parameters of the data generating system \mathcal{S}_o . Obviously, the latter property requires that the data generating system is in the model class \mathcal{M} . Furthermore, asymptotic bias and variance expressions can be derived, see [Tóth (2010)] for details.

It is important to mention that all the above presented results and algorithms are implicitly based on the assumption that the sequence of p is measurable/observable in the system without any measurement noise or disturbance. In the LPV literature, such an assumption is generally taken as a technical necessity and the resulting methods based on it are almost exclusively applied in practical situations where measurements of p are polluted by noise with various stochastic properties. The reason for this theoretical gap lies in the difficulty to establish a conditional expectation of $y(k)$ in the situation when instead of $p(k)$ only its noisy observations are available. Recently it has been proved that using estimated moments, a one-step-ahead predictor of $y(k)$ can be formulated if p is observed up to an additive white noise independent from v and the resulting formulation still allows linear-regression based estimation under an LS criterion [Tóth *et al.* (2011b)]. However, consistency and convergence properties of estimation in that case are currently not well-understood.

4.6. *Global versus local approach*

As demonstrated, both the global and the local approaches provide attractive ways of identifying an LPV system. An obvious question is when to use which approach. In most situations the global approach is considered to be more attractive as it provides estimation of the system with a varying trajectory of p , giving a better possibility to approximate the global dynamic behavior of the system instead of just the frozen aspects. As shown, estimation in the global case can be formulated in a simple least-squares setting and cumbersome problems of interpolation are avoided due to the fixed functional dependencies of the parametrization. The better understood behavior of the stochastic nature of estimation in the global setting

also suggests that it is a theoretically more sound approach than the local method if informativity of \mathcal{D}_N can be guaranteed.

However, practical use of LPV identification on industrial problems often turns out to favor different properties. In most practical applications, identification must be accomplished in a closed-loop setting due to instability of the plant or because the current production can not be disturbed in the favor of identification. In terms of the local approach, the well worked out methods of the LTI framework can be fully used to solve the identification problem in a divide-and-conquer manner. The use of frequency-based identification is also supported in the local setting. The latter is important in mechatronic applications where the often tight modeling specifications with respect to the frozen behaviors are only available in the frequency domain. Usually, such specifications can not be addressed in the global setting. On the other hand, interpolation can result in unexpected global behavior as the local identification approach only focuses on the frozen aspects of the system. However, such a drawback can be avoided by using data with varying p to assist the interpolation. As a general recipe, the use of the global approach is advised whenever there is enough possibility to perturb the system for an informative data record and if the model specifications are not given in the frequency domain. In other situations, the use of a local approach is advised due to its higher capability to meet the target performance under the given information content of available data sets.

There is an important aspect of the proposed identification methods if in the data-generating system p is not an external (free) signal but depends on internal signals like inputs, outputs or state-variables. Such systems are called *quasi*-LPV. For many quasi LPV systems, p can generally be not held constant. In such cases, only the global method is applicable, as the local approach needs identification of the system w.r.t. constant scheduling trajectories. Violation of the freedom of p and how this affects the previous results are generally not well understood and these questions are subject of current research.

5. Identification of a high-performance positioning device

In the sequel, the benefits of the proposed LPV-OBF identification approach are demonstrated on the data-driven modeling of an industrially relevant application: an xy -positioning table.

The conventional design of high-performance positioning devices used com-

monly in the production of *integrated circuits* (ICs) – with usual servo error requirements in the order of $[1, 50]$ [nm] – involves a *long stroke*, called the *xy*-positioning table, moved by two linear motors on parallel rails, see Fig. 4. This table is controlled in the *x*, *y*-translational and the *z*-rotational motion *degrees of freedom* (DOF's). The linearized (or so called local) dynamics of this device varies with its actual position, often manifesting in terms of position-dependent resonant dynamics. Based on [Tóth *et al.* (2011d)], we present in the following an LPV identification study of an *xy*-table aiming at a highly accurate fit of the resulting model w.r.t. the frozen frequency responses of the original plant. To generate-data, we will use a first-principle model which makes it possible to compare the results of the modeling analytically with the original system.

5.1. First-principle modeling

The first-principle modeling concept of a conventional *xy*-positioning table is described in Fig. 4. In this modeling concept, it is assumed that the long stroke has no displacement in the *x*-direction, i.e. $x_2 = 0$. Under this assumption, the dynamical behavior of this multiple mass-damper-spring system \mathcal{S} can be described via the following differential equation:

$$r_M \ddot{w} + r_B(x_1) \dot{w} + r_K(x_1) w = r_F u \tag{33}$$

where $w = [x_1 \ y_1 \ R_{z1} \ y_2 \ R_{z2}]^T$, $u = [F_x \ F_{left} \ F_{right}]^T$ and r_M and r_F are full-rank block diagonal matrices with appropriate dimensions and r_B and r_K are linear matrix functions of x_1 . By taking $p = x_1$ as

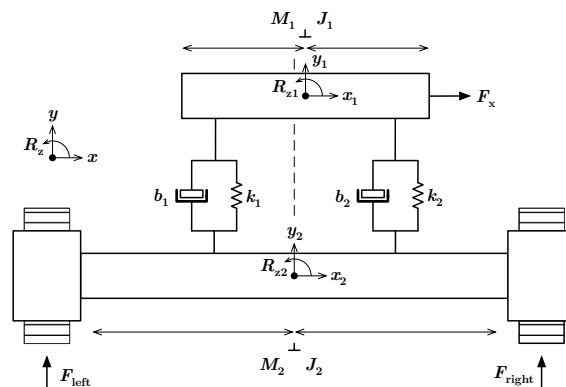


Fig. 4. First-principle modeling concept of *xy*-positioning tables.

a scheduling variable with $\mathbb{P} = [x_{\min}, x_{\max}] \subset \mathbb{R}$, the differential equation Eq. (33) corresponds to an LPV-IO representation $\mathfrak{R}_{\text{IO}}(\mathcal{S})$, with input-output partition $(u, [x_1 \ y_1 \ R_{z1}])$.

In Eq. (33), forces in one direction have influence on the movements in other directions (see [Tóth *et al.* (2011d)]). Thus, to enable the design of SISO controllers, the plant dynamics are commonly decoupled in practice by using pre- and post-transformation matrices T_u and T_y implemented directly into the hardware. As in the LPV case full decoupling of the IO channels is currently not a well-understood concept, the decoupling of the plant is developed by using a rigid-body formulation of Eq. (33), providing approximately decoupled dynamics (i.e. approximately diagonal) in the low-frequency region.

Based on the above given considerations, the rigid-body decoupled plant can be written as:

$$\mathfrak{R}'_{\text{IO}}(\mathcal{S}) = T_y(p) * \mathfrak{R}_{\text{IO}}(\mathcal{S}) * T_u(p). \quad (34)$$

The matrix T_y is defined by the variables to be controlled: $y' = [y_1 - R_{z1}x_1 \ R_{z1}]^T$ which are the actual measurements available from xy -positioning tables (besides the measurement of x_1). T_u is developed by assuming arbitrary slow variation of x_1 and aiming for $T_u P_0 T_y = I$ where P_0 is the static gain of the system, giving a set of new input variables u' satisfying $[F_x \ F_{\text{left}} \ F_{\text{right}}]^T = T_u(p)u'$.

The frozen FRF's of the first-principle model of a real-life xy -positioning table w.r.t. (u', y') , are depicted at different x_1 -positions in terms of Bode magnitude plots in Fig. 5. To protect the interest of the manufacturer, frequency and time have been scaled throughout this example. The following observations are crucial:

- The system dynamics can clearly be separated into an unstable rigid body part dominant in the low frequency band (below 1) and a x_1 -position dependent stable flexible part dominant in the frequency band $[1, 3]$ which is symmetric in magnitude to the $x_1 = 0$ position (phase has a 180° drop at $x_1 = 0$ due to sign change).
- In the diagonal channels, rigid body dynamics correspond to a 2nd order integrator, while in the off-diagonal channels, due to the decoupling, only a small proportional term can be observed.
- At $x_1 = 0$, the off-diagonal transfer functions become approximately

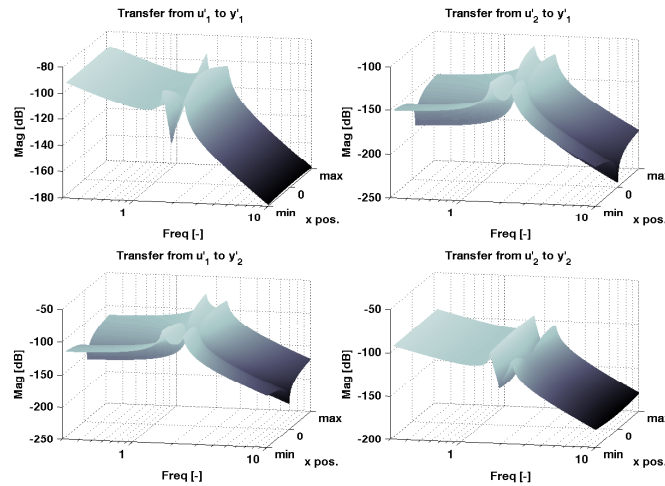


Fig. 5. Bode magnitude plot of the 2×2 MIMO xy -table model at different x -positions.

zero, indicating perfect decoupling. The order of the transfer functions of the diagonal channels (each with order 6) also drops by 2 at $p = 0$.

5.2. Simulation conditions

As the underlying system is unstable, meaningful simulations or measurements can only be obtained under closed-loop conditions. For this purpose robust continuous-time LTI single-loop controllers $K_{y_1}(s)$ and $K_{Rz_1}(s)$ have been designed for the model satisfying moderate specs. in terms of performance. The complete closed control loop of the system is given in Fig. 6, which corresponds to a simplified control architecture used in practice. Additionally, to record DT data for identification purposes, the inputs and outputs of the xy -positioning table in Fig. 6 are sampled with a sampling frequency of 20 (i.e. $10 \times$ the highest interesting frequency point: 2).

To give a realistic setting for identification, noise affecting both the closed loop control and the data acquisition is also considered in the form:

$$\hat{y}_1(k) = y'_1(k) + v_1(k), \quad \hat{y}_2(k) = y'_2(k) + v_2(k), \quad (35)$$

with v_1 and v_2 independent white noise processes with normal distributions: $v_1(k) \in \mathcal{N}(0, \frac{1}{3} \cdot 10^{-7})$ and $v_2(k) \in \mathcal{N}(0, \frac{5}{3} \cdot 10^{-6})$. Such noise levels are typical for the considered laser-interferometers based high-accuracy position measurements. Note that these noise conditions seem to be not so

significant, but due to the relatively small range of movement and the tight error specifications they are challenging enough.

5.3. OBFs based LPV identification

In modeling of *xy*-tables it is important to achieve accuracy of the model w.r.t. both constant and varying trajectories of $p(t) = x_1(t)$. For constant values of p , typical specifications require that the magnitude of the error in terms of the frozen *frequency response function* (FRF) of the system, i.e. $F_p(i\omega)$ needs to be less than -40dB . It is generally true that it is very difficult to include identification constraints into global methods which would guarantee a specified upper bound of the model accuracy for constant p (local fit). As LPV systems do not have a transfer function representation, it is only possible to include frequency domain constraints for local type of approaches. Thus in the following we will study local identification of the *xy*-table using the introduced OBF approach.

5.3.1. *Choice of model structure*

Based on the observations in Section 5.1, it is attractive to separate the system dynamics into an additive “rigid-body part,” which is not dependent on p , and a remaining “flexible part” contains the varying-poles related dynamical aspects of the system. Identifying the flexible part with a fixed DT rigid body filter, provides the means to enforce the well-known fact that the low-frequency behavior of the system is governed by decoupled 2nd-order integrators with an additional zero at -1 for each diagonal IO channel:

$$\phi_R(z) = \frac{z + 1}{(z - 1)^2} \tag{36}$$

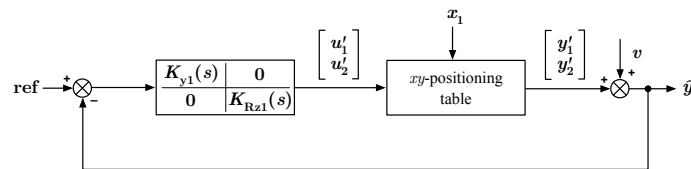


Fig. 6. Simplified closed-loop control scheme of the *xy*-table mechanism with measurement error v .

It can also be observed in Fig. 5 that the “moving” pole locations of the underlying IO channels of the system are the same. This implies that the optimal set of OBFs, which provide the fastest convergence rate, is the same for each channel. Furthermore, local approaches can only identify coefficients, like W_i in Eq. (12), with static dependence. Thus the overall model structure can be chosen as

$$\hat{y}_\theta = \begin{bmatrix} c_1 \phi_R(q) & 0 \\ 0 & c_2 \phi_R(q) \end{bmatrix} u + \sum_{i=1}^{n_f} W_i(p) \phi_i(q) u \quad (37)$$

where $\{\phi_i(q)\}_1^{n_f}$ is a set of SISO OBF's and $c_1, c_2 \in \mathbb{R}$ with $W_i : \mathbb{P} \rightarrow \mathbb{R}^{2 \times 2}$ are the unknown coefficients to be estimated. As a next step, it is important to design our experiments which will give the information upon which adequate selection of the basis functions and the estimation of the expansion coefficients will be accomplished.

5.3.2. Experiment design and data generation

The first step of experiment design for local identification is the gridding of \mathbb{P} . This refers to designing the points on the x_1 -axis around which local LTI identification of the setup will be performed. It is important that the gridding must be dense enough to capture important dynamic changes of the plant for different x_1 -positions. By analyzing the rate of change of the frozen poles and zeros of the system w.r.t. $\mathbb{P} = [x_{\min}, x_{\max}]$, a grid of 21 equidistant points is chosen.

In order to generate informative data for frequency-domain identification at the designated x_1 -positions, orthogonal multisines with normalized amplitude are generated based on 2^{14} equidistant frequency points $\mathcal{W} = \{\omega_k\}_{k=1}^{2^{14}}$ in the range $[10^{-4}, 10]$. This frequency range has been chosen to contain the relevant dynamical aspect of the plant in terms of rigid body and flexible modes. The orthogonality of the generated multisine signals $r_{11}, r_{12}, r_{21}, r_{22}$ can be understood in the following manner: the *discrete-time Fourier transforms* $R_{11}(\omega), \dots, R_{22}(\omega)$ of these signals, satisfy that

$$R(\omega_k) R^H(\omega_k) = \begin{bmatrix} \lambda_1(\omega_k) & 0 \\ 0 & \lambda_2(\omega_k) \end{bmatrix} \prec I, \quad \omega_k \in \mathcal{W},$$

where $R(\omega_k) = \begin{bmatrix} R_{11}(\omega_k) & R_{12}(\omega_k) \\ R_{21}(\omega_k) & R_{22}(\omega_k) \end{bmatrix}$,

where \cdot^H denotes the *Hermitian conjugate*. This property ensures high

accuracy frequency domain estimates in closed loop even under heavy measurement noise. In the experiments, first signals r_{11} and r_{21} are used as references for y'_1 and y'_2 (see Fig. 6) and with constant x_1 equal to a grid point. Then the whole experiment is repeated by using r_{12} and r_{22} . The two set of responses for y'_1 and y'_2 are required to uniquely estimate the 2x2 MIMO FRF of the plant at the considered x_1 position. Note that the normalized reference signals are multiplied with 10^{-4} to remain in the operating range of the setup.

With the designed multisine sequence, data is generated based on the closed loop model starting from zero initial conditions. To generate an appropriately long data record for the attenuation of both the transient and noise effects, the designed multisines are repeated 25 times. For validation purposes, noise free data records are also generated.

5.3.3. FRF estimate of the local behaviors

The data records that are collected in the previous step can now be used to deliver estimates of the FRF of the system at the chosen x_1 -positions. Consider the data sets $\mathcal{D}_{p,1} = \{y'_{\text{ref1}}(k), u'_{\text{ref1}}(k), r_{\text{ref1}}(k)\}_{k=1}^{N_d}$, $\mathcal{D}_{p,2} = \{y'_{\text{ref2}}(k), u'_{\text{ref2}}(k), r_{\text{ref2}}(k)\}_{k=1}^{N_d}$ collected from the model with $x_1 = \mathbf{p}$ and reference signals $r_{\text{ref}i} = [r_{1i} \ r_{2i}]^\top$. Denote the *fast Fourier transform* (FFT) of these signals taken on one period of the time-domain data as $R_{\text{ref1}}(\omega)$, $U'_{\text{ref1}}(\omega)$, $Y'_{\text{ref1}}(\omega)$ and $R_{\text{ref2}}(\omega)$, $U'_{\text{ref2}}(\omega)$, $Y'_{\text{ref2}}(\omega)$ respectively. Due to the periodic nature of the excitation, it is true that after the transients have died out, the FFTs of each period of the measured data records only differ from each other in terms of the additive noise. Therefore, by chopping off the transient part of the data records (first 5-10 periods) and averaging the results of the FFT on the remaining periods, the effect of the noise can be averaged out. Thus in the sequel consider these spectra as the averaged FFT of the non-transient periods. Let

$$\begin{aligned}\bar{U}(\omega) &= [U'_{\text{ref1}}(\omega) \ U'_{\text{ref2}}(\omega)], \\ \bar{Y}(\omega) &= [Y'_{\text{ref1}}(\omega) \ Y'_{\text{ref2}}(\omega)], \\ \bar{R}(\omega) &= [R_{\text{ref1}}(\omega) \ R_{\text{ref2}}(\omega)].\end{aligned}$$

The classical way to estimate the FRF of the plant for a given frequency point $\omega_k \in \mathcal{W}$ is $\hat{F}(\omega_k) = \bar{Y}(\omega_k) \cdot \bar{U}^{-1}(\omega_k)$. However, it is well known that such an *empirical transfer function estimate* is biased in case of closed-loop

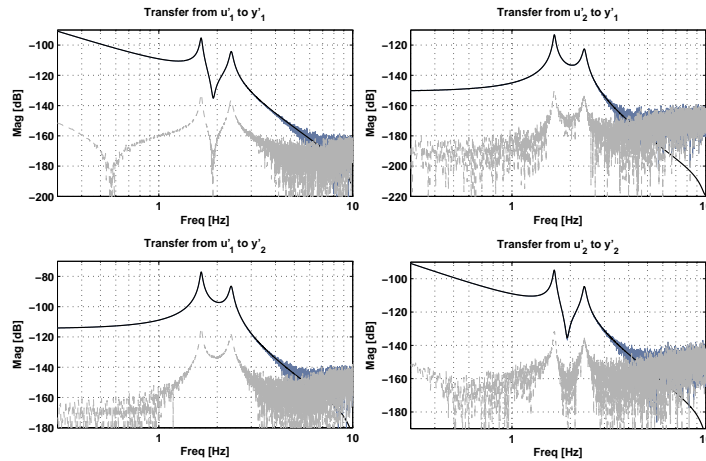


Fig. 7. Bode magnitude plot of the estimated FRF of the plant at position x_{\min} . Original plant (black), estimated FRF (grey), error (light grey).

data. To have an unbiased estimate it is better to consider

$$\hat{F}(\omega_k) = (\bar{Y}(\omega_k)\bar{R}^H(\omega_k)) \cdot (\bar{U}(\omega_k)\bar{R}^H(\omega_k))^{-1}. \quad (38)$$

Among many choices of unbiased closed loop estimators, Eq. (38) has also been observed in the literature to deliver good results under heavy noise settings [Wernholt and Gunnarsson (2007)].

By using the data records and the estimation approach Eq. (38), FRF estimates of the plant at the considered scheduling points have been calculated. During the calculation the first 10 periods in the records have been removed to attenuate the effect of initial conditions. The results at position x_{\min} are depicted in Fig. 7. From this figure it is obvious that the method delivers almost perfect estimates of the frozen FRFs on each IO channel. Furthermore the considered noise only significantly affects the high-frequency band beyond the flexible modes, which shows that accurate frequency-domain information is available to recover the most important dynamical aspects of the plant from measured data.

5.3.4. Selection of the OBF filter banks

To arrive at an adequate selection of the OBF functions in Eq. (12) the FKcM approach introduced in Section 3.3 is used. To obtain an estimate of the frozen pole locations of the xy -positioning table model at the considered

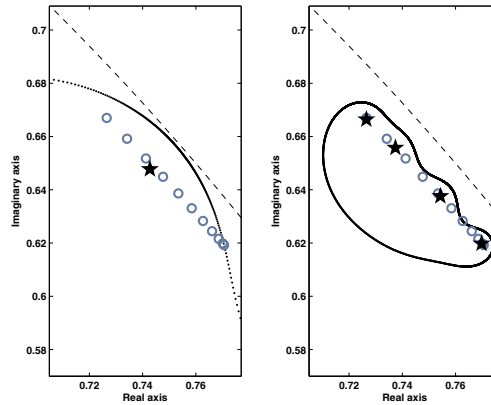
Table 1. Achieved Kolmogorov n -width cost by the FKcM provided OBF's for fuzziness $m = 35$ and different number of OBFs n_g .

n_g	estimated frozen poles	true frozen poles
4	-61.061 dB	-1.651 dB
8	-125.884 dB	-7.084 dB
12	-197.604 dB	-12.031 dB
16	-257.905 dB	-16.845 dB

x_1 -positions, a general *curve fitting* method can be applied on the previously obtained FRF estimates. Here the approach of the FREQID toolbox has been used [de Callafon and Van den Hof (1996)]. To arrive at the correct number of poles a MIMO common denominator model with 8th order has been estimated with curve fitting. The worst-case absolute error of the resulting pole estimates w.r.t. the true pole locations of the system at the given x -positions has been 0.07%.

Next the FKcM approach is applied on the obtained highly-accurate pole estimates. By analyzing the results of the algorithm based on the estimated pole locations, it has been observed that nearly optimal basis selection can be achieved if the fuzziness m is set to 35. Using this fuzziness value the algorithm has been executed on the estimated pole locations. The algorithm has been used with different number of optimized basis functions n_g and the results are summarized in Table 1 and given in Fig. 8. In this table, as a performance measure, the Kolmogorov n -width cost (see Eq. (15)) with $n = n_g$ has been computed in dB both for the obtained pole estimates and also for all true frozen pole locations Ω_* of the xy -positioning table in the considered x_1 -region. From Fig. 8 and Table 1 it follows that the obtained OBF poles achieve very small representation error w.r.t the estimated poles (small Kolmogorov cost). However, w.r.t. the true pole locations a dramatic difference can be observed between the set of 4 or 16 basis functions. Further analysis of the results shows that at least 12 basis functions are needed to adequately represent the varying system dynamics on \mathbb{P} . After posteriori assessment of the estimation with 12 and 16 basis functions it is concluded that 16 basis functions are required to meet with the specifications. It can also be shown that beyond 16 basis functions the improvement in model accuracy is not significant. Thus in the sequel, we will consider the OBF set with 16 basis functions selected by the FKcM approach to formulate Eq. (37).

Fig. 8. Zoomed in figures depicting the results of the basis selection w.r.t. the estimated frozen poles (i.e the set Z) given with circles. The cluster centers (basis poles) are given with black \star . The black lines represent the Kolmogorov bound $\Omega(\Xi_c, \rho_*)$ w.r.t. the frozen pole set Ω_* of the system, while the dashed lines are the perimeter of the unit disc. On the left hand side $c = 4$ clusters are determined, on the right hand side $c = 16$.



5.3.5. Estimation of the expansion coefficient dependence

By having the OBF filter banks chosen, the last remaining step of identification is to estimate the constants c_1 , c_2 and the expansion coefficients $W_i(p)$ in Eq. (37). For this purpose the already calculated FRF estimates of the system are used. Note that the frequency response of the OBF filters and the first-principles suggested rigid body filters can be computed w.r.t. the frequency points of the FRF estimate and in terms of the model structure these frequency responses should approximate the estimated FRFs by linear combination. Thus, estimation of the samples of the expansion coefficients W_i at each considered grid point $\mathbf{p} \in \mathcal{P}$ reduces to a simple linear regression. After solving the linear regression, the resulting samples of each W_i can be interpolated using any approach like polynomial, spline, Chebyshev, etc. After investigation of the obtained results with each method, it has been concluded that a polynomial interpolation provides the most efficient solution in terms of the complexity/accuracy trade off. Regarding polynomial interpolation it has been concluded that for the case of 16 OBF functions a polynomial order of 15 is minimally required to achieve a good approximation of the frozen dynamics. By using the FRF estimates, the samples of the expansion coefficients of the OBF filter banks obtained in the previous section with $n_g = 16$ have been estimated and these samples have been interpolated with 17th order polynomials. The results are depicted in Fig. 9. These figures show that using only a few estimated samples of the coefficient functions, a close approximation of the polynomial dependencies can be obtained. This concludes the identification as the delivered model now can be explicitly realized in an LFR form.

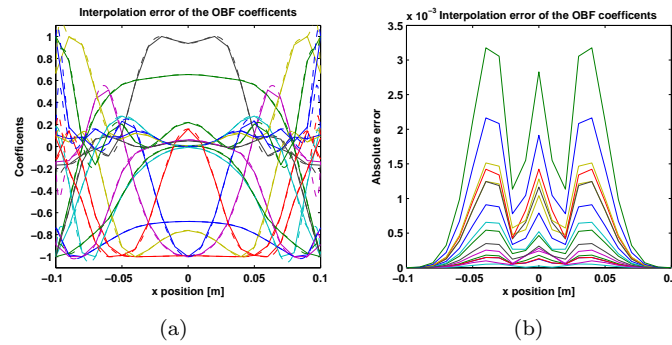


Fig. 9. Optimal coefficient functions W_i of the OBF's $n_g = 16$ (solid lines) w.r.t. the frozen transfer functions between y'_1 and u'_1 at the grid positions x_1 together with their polynomial approximation (dashed lines).

5.4. Validation of the model

As a final step it remains to validate the obtained model in both the frequency and the time domain.

5.4.1. Frequency-domain validation

The obtained LPV-OBF model can be compared in terms of its frozen frequency responses to the behavior of the first-principle model. For the OBF model with 16 basis functions and 17th-order polynomial dependence, the frozen frequency responses for each IO channel have been computed on a fine grid $\mathcal{P} \subset \mathbb{P}$ (1000 points) together with the response of the true system and visualized in terms of Bode plots. The results at x_{\min} are given in Fig. 10 which corresponds to the worst-case model fit. By analyzing these results the following observations can be made:

- The overall difference between the magnitude error and magnitude of the transfer functions is approx. 40 dB.
- However, the error increases with an approx. 20 dB around the anti-resonance mode. This results as a side effect of linear regression. Using better tuned weights this error can be decreased if necessary.

This means that specifications in terms of frequency-domain accuracy could be achieved with the investigated identification approach.

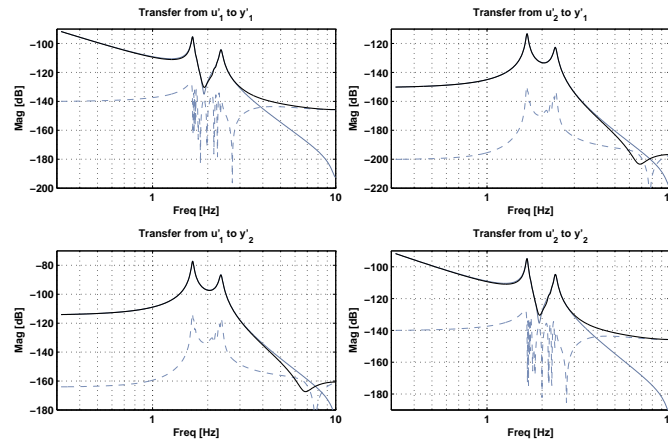


Fig. 10. Bode magnitude plot of the frozen frequency response at position x_{\min} of the estimated LPV-OBF model obtained with 16 basis functions and 17th-order polynomial coefficient dependence. Original plant (black), estimated OBF model (grey), approximation error (dashed grey).

5.4.2. Time-domain validation

As a next step we investigate the time-domain behavior of the identified LPV-OBF model. First the open-loop response of the model is computed by using recorded u and p signals from a closed-loop simulation of the original xy -positioning table model for a monotone increasing p which corresponds to a fast sweep over \mathbb{P} . The used reference signals here are a zero set-point for R_{z1} and a typical step-like pattern for y_1 , designed in terms of optimal speed, acceleration and jerk profile. The resulting responses of the LPV-OBF model (after re-transformation with $T_y(p)$) are given in gray in Fig. 11 while the response of the original plant is given in black. The error is dominated by a small difference that looks like the step response of an integrator. This yields that the identified LPV-OBF model is capable to reproduce the response of the system with high accuracy and the main source of the error is related to small differences between the unstable part of system and the LPV-OBF model. This hypothesis is also validated by the closed-loop response of the LPV-OBF model given with dashed gray in Fig. 11. It is important to note that we considered validation of the model with varying p , while the LPV-OBF model was obtained purely on the basis of the frozen behavior of the system. By achieving an acceptable error which meets the aimed specs. we conclude that the proposed identification approach can deliver high-quality model estimates.

5.4.3. Economical size

As we could see, a high number of OBFs and a high-order polynomial coefficient dependence were needed to capture the dynamics of the xy -table with the desired accuracy. This means that the final LFR form of the identified model (37) is relatively large with $\dim(x) = 4 + 2 \cdot 16 = 36$ and $\dim(z) = 2 \cdot 17 = 34$. However, by applying recent methods in LPV model reduction, like the approach of [Pettersson and Löfberg (2009)], this LFR form can be reduced to state dimension 8 and with $\dim(z) = 5$, without a significant loss of accuracy. The explanation lays in the fact that in the considered model structure Eq. (37) all dependencies on p are at the output-side. Therefore in terms of realization, there is a certain freedom to consider states and input contributions which also depend on p and hence the total dimension of the model can be reduced.

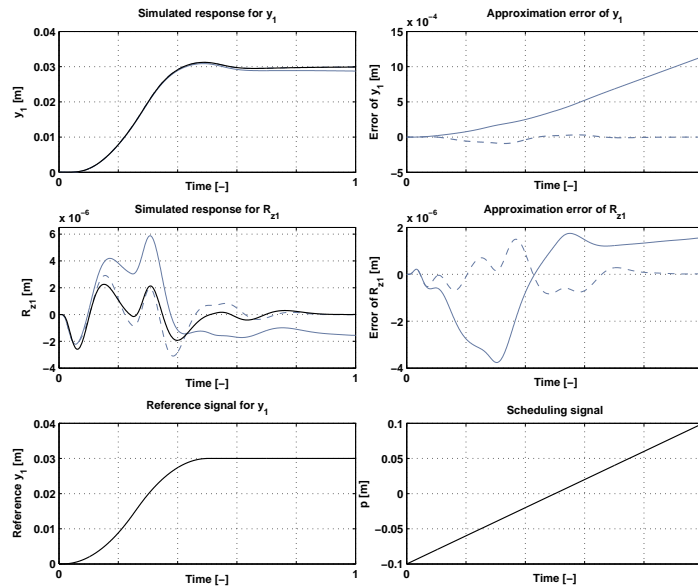


Fig. 11. Time-domain validation of the estimated LPV-OBF model using 16-basis functions with 17th-order polynomial coefficient dependence. Simulated closed-loop response of the original plant is given with black for the reference and scheduling signals depicted in the lower two figures. The simulated response of the OBF model based on the closed loop input signals of the plant is given with grey (open-loop validation) while the closed-loop response of the model for the given reference signal is depicted with dashed grey (closed-loop validation).

6. Conclusion

Based on a series-expansion representation of LPV systems, model structures using orthonormal basis functions have been developed which allow system identification in an attractive manner. These models represent an interesting trade-off between state-space and input-output models by having a direct state-space realization and allowing the efficient use of linear regression for the estimation of the underlying parameters with well-understood stochastic properties. Identification schemes of these models have been proposed both in the local and global setting and attractive properties of these approaches have been demonstrated on the identification of an industrially relevant application.

References

- Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms* (Plenum Press, New York, NY, USA).
- Boyd, S. and Chua, L. O. (1985). Fading memory and the problem of approximating nonlinear operators with Volterra series, *IEEE Transactions on Circuits and Systems* **32**, 11, pp. 1150–1161.
- Casella, F. and Lovera, M. (2008). LPV/LFT modelling and identification: overview, synergies and a case study, in *Proceedings of the IEEE Multi-conference on Systems and Control* (San Antonio, TX, USA).
- Dankers, A. G., Tóth, R., Heuberger, P. S. C., Bombois, X. and den Hof, P. M. J. V. (2011). Identifiability and the informativity of data sets for LPV-ARX identification, in *Proceedings of the IEEE Conference on Decision and Control* (Orlando, FL, USA).
- de Callafon, R. A. and Van den Hof, P. M. J. (1996). FREQID - frequency domain identification toolbox for use with Matlab, *Selected Topics in Identification, Modeling and Control* **9**, pp. 129–134.
- Heuberger, P. S. C., Van den Hof, P. M. J. and Bo Wahlberg (2005). *Modeling and identification with rational orthonormal basis functions* (Springer-Verlag, London, UK).
- Hsu, K., Vincent, T. L. and Poolla, K. (2008). Nonparametric methods for the identification of linear parameter varying systems, in *Proceedings of the International Symposium on Computer-Aided Control System Design* (San Antonio, TX, USA).
- Jain, A. K. and Dubes, R. C. (1988). *Algorithms for clustering data* (Prentice Hall, Englewood Cliffs, NJ, USA).
- Kaymak, U. and Setnes, M. (2002). Fuzzy clustering with volume prototypes and adaptive cluster merging, *IEEE Transactions on Fuzzy Systems* **10**, 6, pp. 705–711.
- Khalate, A. A., Bombois, X., Tóth, R. and Babuška, R. (2009). Optimal experi-

- mental design for LPV identification using a local approach, in *Proceedings of the IFAC Symposium on System Identification* (Saint-Malo, France).
- Kwiatkowski, A., Werner, H. and Boll, M. T. (2006). Automated generation and assessment of affine LPV models, in *Proceedings of the IEEE Conference on Decision and Control* (San Diego, CA, USA).
- Laurain, V., Gilson, M., Tóth, R. and Garnier, H. (2010). Refined instrumental variable methods for identification of LPV Box-Jenkins models, *Automatica* **46**, 6, pp. 959–967.
- Oliveira e Silva, T. (1996). A n -width result for the generalized orthonormal basis function model, in *Proceedings of the IFAC World Congress* (Sydney, Australia).
- Petersson, D. and Löfberg, J. (2009). Optimization based LPV-approximation of multiple model systems, in *Proceedings of the European Control Conference* (Budapest, Hungary).
- Pinkus, A. (1985). *n -widths in approximation theory* (Springer-Verlag, New York, NY, USA).
- Tóth, R. (2010). *Modeling and identification of linear parameter-varying systems*, Lecture Notes in Control and Information Sciences (Springer-Verlag).
- Tóth, R., Abbas, H. and Werner, W. (2011a). On the state-space realization of LPV input-output models: practical approaches, *in print, IEEE Transactions on Control Systems Technology*.
- Tóth, R., Bitar, E., Heuberger, P. S. C., Van den Hof, P. M. J. and Poolla, K. (2011b). A prediction-error identification framework for linear parameter-varying systems, *Submitted to Automatica*.
- Tóth, R., Heuberger, P. S. C. and Van den Hof, P. M. J. (2009a). Asymptotically optimal orthonormal basis functions for LPV system identification, *Automatica* **45**, 6, pp. 1359–1370.
- Tóth, R., Laurain, V., Zheng, W., and Poolla, K. (2011c). A support vector machine approach for LPV linear-regression models, in *Proceedings of the IEEE Conference on Decision and Control* (Orlando, FL, USA).
- Tóth, R., Lyzell, C., Enqvist, M., Heuberger, P. S. C. and Van den Hof, P. M. J. (2009b). Order and structural dependence selection of LPV-ARX models using a nonnegative garrote approach, in *Proceedings of the IEEE Conference on Decision and Control* (Shanghai, China).
- Tóth, R., van de Wal, M., Heuberger, P. S. C. and Van den Hof, P. M. J. (2011d). LPV identification of high performance positioning devices, in *Proceedings of the American Control Conference* (San Francisco, CA, USA).
- Tóth, R., Willems, J. C., Heuberger, P. S. C. and Van den Hof, P. M. J. (2011e). The behavioral approach to linear parameter-varying systems, *in print, IEEE Transactions on Automatic Control*.
- Tropp, J. A. and Wright, S. J. (2010). Computational methods for sparse solution of linear inverse problems, *Proceedings of the IEEE* **98**, 6, pp. 948–958.
- Wei, X. and Del Re, L. (2006). On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque, in *Proceedings of the IFAC Symposium on System Identification* (Newcastle, Australia).

- Wernholt, E. and Gunnarsson, S. (2007). Analysis of methods for multivariable frequency response function estimation in closed loop, in *Proceedings of the Conference on Decision and Control* (New Orleans, LA, USA).
- Wood, G. D., Goddard, P. J. and Glover, I. (1996). Approximation of linear parameter-varying systems, in *Proceedings of the IEEE Conference on Decision and Control* (Kobe, Japan).