



Asymptotically optimal orthonormal basis functions for LPV system identification[☆]

Roland Tóth^{*}, Peter S.C. Heuberger, Paul M.J. Van den Hof

Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands

ARTICLE INFO

Article history:

Received 27 February 2008

Received in revised form

29 June 2008

Accepted 22 January 2009

Available online 10 March 2009

Keywords:

Linear parameter-varying systems

System identification

Orthonormal basis functions

Pole clustering

ABSTRACT

A global model structure is developed for parametrization and identification of a general class of Linear Parameter-Varying (LPV) systems. By using a fixed orthonormal basis function (OBF) structure, a linearly parametrized model structure follows for which the coefficients are dependent on a scheduling signal. An optimal set of OBFs for this model structure is selected on the basis of local linear dynamic properties of the LPV system (system poles) that occur for different constant scheduling signals. The selected OBF set guarantees in an asymptotic sense the least worst-case modeling error for any local model of the LPV system. Through the fusion of the Kolmogorov n -width theory and Fuzzy c -Means clustering, an approach is developed to solve the OBF-selection problem for discrete-time LPV systems, based on the clustering of observed sample system poles.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

In general, many physical systems and control problems exhibit parameter variations due to non-stationary or nonlinear behavior or dependence on external variables, such as space coordinates, in particular found in servo-mechanical applications. These systems vary in size and complexity, but they share the common need for accurate and efficient control of the relevant process variables. However, accurate modeling of such systems is in general a complex and tedious task, involving the use of nonlinear differential equations, leading to models with many parameters and high computational complexity. For processes with mild nonlinearities or dependence on external variables, the theory of *Linear Parameter-Varying* (LPV) systems offers an attractive modeling framework (Rugh & Shamma, 2000). Discrete-time LPV systems are generally described in either a *State-Space* (SS) or an *Input/Output* (I/O) representation (Tóth, Felici, Heuberger, & Van den Hof, 2007), where the parameters are functions of a time-varying *scheduling signal* $p(k) : \mathbb{Z} \rightarrow \mathbb{P}$, that schedules between local *Linear Time Invariant* (LTI) behaviors of

the system. The compact set $\mathbb{P} \subset \mathbb{R}^{n_p}$ denotes the *scheduling space*. Practical use of the LPV framework is stimulated by the fact that control design for LPV systems is well worked out. For this class of systems, application of LTI control theory via *gain scheduling* (Rugh & Shamma, 2000) and LPV control synthesis techniques like μ -*synthesis* (Zhou & Doyle, 1998) or *Linear Matrix Inequality* (LMI)-based optimal control (Scherer, 1996) offer fast and reliable controller design, proved by a wide range of applied LPV control solutions from aerospace applications (Marcos & Balas, 2004) to CD players (Dettori & Scherer, 2001). However, it still remains a problem how to develop LPV models in a systematic fashion.

Recently several methods have been worked out, aiming at global identification of discrete-time LPV models from given measured data. This comprises methods based on multiple-model approaches (Murray-Smith & Johansen, 1997; Steinbuch, van de Molengraft, & van der Voort, 2003; Wassink, van de Wal, Scherer, & Bosgra, 2004), set-membership methods (Mazzaro, Movsichoff, & Pena, 1999; Milanese & Vicino, 1991), subspace techniques (dos Santos, Ramos, & de Carvalho, 2007; Felici, van Wingerden, & Verhaegen, 2006, 2007; Verdult & Verhaegen, 2002), basis functions (Tóth, Heuberger, & Van den Hof, 2007), LMI-based optimization (Sznaier, Mazzaro, & Inanc, 2000), simple *Least Mean Squares* (LMS) approaches (Giarré, Bauso, Falugi, & Bamieh, 2006; Wei & Del Re, 2006), and parameter estimation based gradient searches (Lee & Poola, 1996; Verdult, Ljung, & Verhaegen, 2002). Most of these approaches build on the fact that an LPV system \mathcal{S} can always be viewed as a collection of “local” behaviors and p -dependent weighting functions, i.e. *scheduling functions* that

[☆] The material in this paper was partially presented at 45th IEEE Conf. on Decision and Control, 2006 Dec., San Diego, California, USA. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor Torsten Söderström.

^{*} Corresponding author. Tel.: +31 (0) 631382635; fax: +31 (0)15 2786679.

E-mail addresses: tothrola@ieee.org, r.toth@tudelft.nl (R. Tóth), p.s.c.heuberger@tudelft.nl (P.S.C. Heuberger), p.m.j.vandenhof@tudelft.nl (P.M.J. Van den Hof).

schedule between them (Rugh & Shamma, 2000; Tóth, Heuberger et al., 2007). For any constant scheduling signal: $p(k) = \bar{p}$ for all $k \in \mathbb{Z}$ where $\bar{p} \in \mathbb{P}$, the LPV system \mathcal{S} is identical to an LTI system $\mathcal{F}_{\bar{p}}$. Thus, the set of local behaviors of \mathcal{S} is given as $\mathfrak{F}_{\mathbb{P}} = \{\mathcal{F}_{\bar{p}}\}_{\bar{p} \in \mathbb{P}}$. The p -dependent scheduling function set, that schedules on $\mathfrak{F}_{\mathbb{P}}$, is denoted by $\mathfrak{H}_{\mathbb{P}} = \{h_{\bar{p}}(\cdot)\}_{\bar{p} \in \mathbb{P}}$.

Identification of $\mathfrak{F}_{\mathbb{P}}$ is commonly accomplished in a sampled sense by LTI identification of \mathcal{S} for a set of constant scheduling signals, associated with (for instance equidistant) points in the scheduling space \mathbb{P} . Then, assuming that the scheduling functions $\{h_{\bar{p}}\}$ have a particular structure of dependence, like polynomial, an interpolation problem is formulated on \mathbb{P} to give a global approximation of \mathcal{S} . Recently it was exposed that this approach should be handled with care for several reasons (Tóth, Felici et al., 2007; Tóth, Heuberger et al., 2007). In Tóth, Felici et al. (2007) it was shown that for general discrete-time LPV systems each $h_{\bar{p}}$ is a function of time-shifted versions of p (dynamic dependence). Then, if the particular interpolation structure of $\{h_{\bar{p}}\}$ is chosen to be too simple (dependence only on $p(k)$ (static dependence), linear dependence, etc.) the interpolation based on state-space or I/O model parametrization can result in significantly different models (Tóth, Felici et al., 2007). An additional concern of interpolation is that the McMillan degree of the local systems $\{\mathcal{F}_{\bar{p}}\}$ may vary for different values of $\bar{p} \in \mathbb{P}$. This shows that the choice of an easily interpolatable model structure which can incorporate aspects of dynamical dependence and local order changes is a crucial point of this identification approach.

The *Orthonormal Basis Function* (OBF)-based model representation offers such a structure with a well worked-out theory in the context of LTI system approximation and identification (Heuberger, Van den Hof, & Wahlberg, 2005). The basis functions, that provide bases for the system space \mathcal{H}_2 (Hilbert space of complex functions that are squared integrable on the unit circle), are generated by a cascaded network of stable all-pass filters, whose pole locations represent the prior knowledge about the system at hand. This approach characterizes the transfer function of a strictly proper LTI system as

$$F(z) = \sum_{i=1}^{\infty} w_i \phi_i(z), \quad (1)$$

where $\{w_i\}_{i=1}^{\infty}$ is the set of coefficients and $\Phi_{\infty} = \{\phi_i\}_{i=1}^{\infty}$ represents the sequence of OBFs. This implies that every $\mathcal{F}_{\bar{p}} \in \mathfrak{F}_{\mathbb{P}}$ can be represented as a linear combination of a given Φ_{∞} , i. e. $\mathfrak{F}_{\mathbb{P}} \subset \text{span}\{\Phi_{\infty}\}$. In practice, only a finite number of terms is used in (1), like in *Finite Impulse Response* (FIR) models. In contrast with FIR structures, the OBF parametrization can achieve almost zero modeling error with a relatively small number of parameters, due to the infinite impulse response characteristics of the basis functions. In this way, it is always possible to find a finite $\Phi_n \subset \Phi_{\infty}$, with a relatively small number of functions $n \in \mathbb{N}$, such that the representation error for all $\mathcal{F}_{\bar{p}}$ is negligible. Using this idea in the time-domain (substitution of z with the forward time-shift operator q), it is possible to prove that LPV systems also have a series expansion representation in terms of LTI basis functions, but with coefficients $\{w_i\}_{i=1}^{\infty}$ dependent on p . Thus in terms of a finite OBF set $\Phi_n \subset \Phi_{\infty}$, the following approximation of the I/O map of \mathcal{S} can be introduced:

$$y \approx \sum_{i=1}^n w_i(p) \phi_i(q) u, \quad (2)$$

where $\{w_i\}_{i=1}^n$ is a set of *coefficient functions*, with dynamic dependence on p . Note that in this structure, Φ_n gives the basis set used to approximate each element of $\mathfrak{F}_{\mathbb{P}}$ while $\{w_i\}_{i=1}^n$ describes the scheduling functions $\mathfrak{H}_{\mathbb{P}}$. Thus for a given $\Phi_n = \{\phi_i\}$, identification of the LPV system based on (2) simplifies to the identification of

the scheduling functions. Assuming static dependence of $\{w_i\}_{i=1}^n$, such a task can be accomplished via two approaches:

- **Local approach:** Identify some $\mathcal{F}_{\bar{p}} \in \mathfrak{F}_{\mathbb{P}}$ for constant $p(t) = \bar{p}$ with the LTI OBF model structure

$$\hat{y} = \sum_{i=1}^n r_{\bar{p},i} \phi_i(q) u. \quad (3)$$

Based on a chosen functional dependence, e.g. polynomial, interpolate the resulting $\{r_{\bar{p},i}\}$ for an estimate of $\{w_i\}_{i=1}^n$ in (2), such that $w_i(\bar{p}) = r_{\bar{p},i}$.

- **Global approach:** Parametrize the functional dependence of $\{w_i\}_{i=1}^n$ linearly (e.g. polynomial). Then for a data record with varying p , the estimation of the parameters of $\{w_i\}_{i=1}^n$ reduces to linear regression based on (2) in a least-squares prediction error setting.

There are many beneficial properties of the structure (2). For instance, the obtained model simplifies control design (see Section 6) and this parametrization is not affected by local order changes. The problem that remains to be solved with the proposed OBF-based identification approaches is to choose the set of OBFs Φ_n , “sufficiently rich” to describe $\mathfrak{F}_{\mathbb{P}}$ with a predefined number of functions. Seeking the solution for this problem is the purpose of the present paper.

Even in the case of LTI systems, the choice of OBFs to approximate a given system \mathcal{F} in an “optimal” sense (based on some error measure) is a highly non-trivial task (Heuberger et al., 2005). For the LTI case, already quite some effort has been put into tackling the basis function selection problem resulting in methods of nonlinear optimization (Heuberger et al., 2005) and iterative search (Bodin, Villemoes, & Wahlberg, 1997). One of the concepts used for this purpose is the *Kolmogorov n-width* (KnW) theory for OBFs (Oliveira e Silva, 1996), which establishes optimality in the sense of the worst-case modeling error for any LTI system with pole locations in a given region of the complex plane. Denote by

$$\Omega_{\mathbb{P}} = \{\lambda \in \mathbb{C} \mid \lambda \text{ is a pole of } \mathcal{F}_{\bar{p}} \in \mathfrak{F}_{\mathbb{P}} \text{ for } \bar{p} \in \mathbb{P}\},$$

the collection of all pole locations belonging to the local behaviors of the LPV system \mathcal{S} . Then, based on $\Omega_{\mathbb{P}}$, the KnW theory can be evidently applied (e.g. by the approach of Heuberger et al. (2005)) to solve the optimal selection of OBFs with respect to $\mathfrak{F}_{\mathbb{P}}$. However, this approach is not applicable if $\mathfrak{F}_{\mathbb{P}}$ is unknown. This underlines the need for a mechanism that guarantees optimality of the OBF selection (selection of Φ_n) based on the available information.

In this paper, we assume as a starting point that a collection of pole locations, some samples of $\Omega_{\mathbb{P}}$, is available that are obtained from local linear behaviors of the LPV system \mathcal{S} . This set of pole samples $\bar{\Omega} \subset \Omega_{\mathbb{P}}$ can result – but not necessarily – from identification of the related local linear models. Based on $\bar{\Omega}$, we aim at the derivation of a basis function selection mechanism, that is capable of accomplishing the following objectives:

- Reconstruction of $\Omega_{\mathbb{P}}$ from $\bar{\Omega}$.
- Determination of the set of OBFs, which has the least possible worst-case modeling error for any LTI system with pole locations in $\Omega_{\mathbb{P}}$, therefore for all $\mathcal{F}_{\bar{p}} \in \mathfrak{F}_{\mathbb{P}}$.

This choice of model structure leads to the local and global identification methods. The proposed method is the joint application of the KnW theory and *Fuzzy c-Means* (FcM) clustering (Jain & Dubes, 1988). The contribution of this method is to provide a practical model structure selection tool for the local and global LPV identification methods based on globally fixed OBFs. Earlier work along this line is proposed in Tóth, Heuberger, and Van den Hof (2006a), Tóth, Heuberger, and Van den Hof (2006b) and Vergeer (2005).

The paper is organized as follows: Section 2 introduces the description and properties of OBFs while Section 3 describes the n -width result with respect to these functions; in Section 4, the mechanism of the KnW-based FcM pole clustering is given that solves simultaneously the determination of $\Omega_{\mathbb{P}}$ from sampled poles and the selection of optimal OBFs with respect to $\Omega_{\mathbb{P}}$; in Section 5, the OBF-based LPV system identification scheme is specified to provide a brief description how the selected basis functions are used in an identification scenario; in Section 6, the applicability of the introduced method is shown through an example; and finally, in Section 7, the main results of the paper are discussed.

2. Orthonormal basis functions

We consider only the case of real rational (finite-dimensional) discrete-time, SISO transfer functions. For details see Heuberger et al. (2005), Heuberger, Van den Hof, and Bosgra (1995) and Ninness and Gustafsson (1997). Let $G_0 \equiv 1$ and $\{G_i\}_{i=1}^{\infty}$ be a sequence of inner functions (i.e. stable transfer functions with $G_i(z)G_i(\frac{1}{z}) = 1$), and let $\{A_i, B_i, C_i, D_i\}$ be minimal balanced SS representations of G_i . Let $\{\xi_1, \xi_2, \dots\}$ denote the collection of all poles of the inner functions G_1, G_2, \dots . Under the (completeness) condition that $\sum_{i=1}^{\infty} (1 - |\xi_i|) = \infty$, the scalar elements of the sequence of vector functions

$$V_n(z) = (zI - A_n)^{-1} B_n \prod_{i=0}^{n-1} G_i(z), \quad (4)$$

constitute a basis for $\mathcal{H}_{2-}(\mathbb{E})$, the *Hardy space* of functions, which are 0 for $z = \infty$, analytic on \mathbb{E} , the exterior of the unit disk \mathbb{D} , and squared integrable on the unit circle \mathbb{T} with norm $\|\cdot\|_{\mathcal{H}_2}$. In this way $\mathcal{H}_{2-}(\mathbb{E})$ is the space of all stable strictly proper transfer functions. These functions (4) are often referred to as the *Takenaka-Malmquist functions*. The special cases when all G_i are equal, i.e. $G_i(z) = G_b(z)$, $\forall i > 0$, where G_b has McMillan degree $n_b > 0$, are known as *Hambo functions* or *generalized orthonormal basis functions* (GOBFs) for arbitrary n_b , *2-parameter Kautz functions* for $n_b = 2$, and as *Laguerre functions* for $n_b = 1$. Note that for these cases the completeness condition is always fulfilled. In the remainder we will only consider the set of Hambo functions. Let G_b be an inner function with McMillan degree $n_b > 0$ and minimal balanced SS representation $\{A_b, B_b, C_b, D_b\}$. Define $V_1(z) = (zI - A_b)^{-1} B_b$ and $\phi_j = [V_1]_j$, $j \in \mathbb{I}_1^{n_b}$, where $\mathbb{I}_1^{s_2} = \{s_1, s_1 + 1, \dots, s_2\} \subset \mathbb{Z}$ is the index set. The Hambo basis then consists of the functions $\Phi_{n_b}^{\infty} = \{\phi_j G_b^i\}_{j=1, \dots, n_b}^{i=0, \dots, \infty}$. An important aspect of these bases is that the inner function G_b is, modulo the sign, completely determined by its poles $\mathcal{E}_{n_b} := \{\xi_1, \dots, \xi_{n_b}\}$:

$$G_b(z) = \pm \prod_{j=1}^{n_b} \frac{1 - z\xi_j^*}{z - \xi_j}, \quad (5)$$

where $*$ denotes complex conjugation, and it is immediate that the function V_1 has the same poles. Any $F \in \mathcal{H}_{2-}(\mathbb{E})$ can be decomposed as

$$F(z) = \sum_{i=0}^{\infty} \sum_{j=1}^{n_b} w_{ij} \phi_j(z) G_b^i(z), \quad (6)$$

and it can be shown that the rate of convergence of this series expansion is bounded by $\rho = \max_k |G_b(\lambda_k^{-1})|$, called the *decay rate*, where $\{\lambda_k\}$ are the poles of $F(z)$. In the “best” case, where the poles of F are the same (with multiplicity) as the poles of G_b , only the terms with $i = 0$ in (6) are non-zero. The I/O relation of the OBF parametrization (6) is illustrated in Fig. 1.

In practice, only a finite number of terms $\Phi_{n_b}^{n_e} = \{\phi_j(z) G_b^i\}_{j=1, \dots, n_b}^{i=0, \dots, n_e}$ with $n_e \geq 0$ is used in (6), like in *Finite Impulse Response* (FIR) models. In contrast with FIR structures, which are

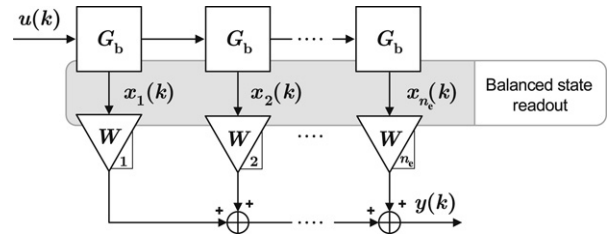


Fig. 1. I/O signal flow graph of the OBF model structure described by (6) for a finite n_e number of extensions of G_b and with $W_i = [w_{i1}, \dots, w_{in_b}]$.

described by (6) as a finite linear combination of *pulse basis functions* $\phi_1(z) = z^{-1}$ with $n_b = 1$ and $G_b^i(z) = z^{-i}$, the OBF parametrization uses a broad class of basis functions with infinite impulse responses. Therefore, OBF parametrization can achieve almost zero modeling error with a relatively small number of parameters due to the faster convergence of the series representation than in the FIR case. Moreover, the reduced number of parameters in the structure results in decreased variance of the final model estimate.

Identification of any $F \in \mathcal{H}_{2-}(\mathbb{E})$ based on a predefined set of OBFs $\Phi_{n_b}^{n_e}$ consisting of $n = (n_e + 1)n_b$ basis functions, is performed as a linear regression with respect to the basis coefficients $W_{n_b}^{n_e} = [w_{ij}]_{j=1, \dots, n_b}^{i=0, \dots, n_e}$ due to the linear parametrization of (6). The OBF-based identification has valuable properties. Non-asymptotic variance bounds of the estimates are computable through reproducing kernels and the identified models are unbiased if the input signal is uncorrelated to the noise. This is explained by the *Output Error* (OE) like structure of the OBF parametrization (Heuberger et al., 2005). However, selection of the basis function set has a major impact on the outcome of the identification process as the distance between basis poles and the original system poles determines the convergence rate of the coefficients, meaning that with a “better” basis function set a better approximation can be achieved.

As discussed, OBF-based parametrization can be effectively used for LTI system representation and in this way to describe each $\mathcal{F}_{\mathbb{P}} \in \mathfrak{F}_{\mathbb{P}}$ of an LPV system \mathcal{S} . However, if the same OBFs are used to compose each $\mathcal{F}_{\mathbb{P}}$, then it is required that the basis function set is “well chosen” with respect to the entire $\mathfrak{F}_{\mathbb{P}}$. In the next section, the concept of optimality of an OBF set with respect to $\mathfrak{F}_{\mathbb{P}}$ is established, giving the key theorem to solve the basis function selection problem of the proposed identification scheme.

3. Kolmogorov n -width for OBFs

In the proposed LPV identification approach, it is crucial to find an appropriate model set, i.e. set of basis functions $\Phi_{n_b}^{n_e}$ for the local behaviors $\mathfrak{F}_{\mathbb{P}}$, in the sense that $\Phi_{n_b}^{n_e}$ is sufficiently rich to describe the systems belonging to $\mathfrak{F}_{\mathbb{P}}$, with a relatively small number of statistically meaningful parameters. In LTI system identification, one approach to find appropriate model sets is based on the n -width concept (Pinkus, 1985), which was shown to result in appropriate model sets for robust modeling of linear systems (Mäkilä & Partington, 1993). Using this concept, Oliveira e Silva (1996), (Heuberger et al., 2005, Ch. 11) showed that OBF model structures are optimal in the n -width sense for specific subsets of systems. In the following, the basic ingredients of this theory for discrete-time, stable, SISO systems are described.

Let \mathfrak{F} denote a set of systems with transfer functions $\{F\} = \mathfrak{F} \subseteq \mathcal{H}_{2-}(\mathbb{E})$, that we want to approximate with the linear combination of n elements of $\mathcal{H}_{2-}(\mathbb{E})$. Let $\Phi_n = \{\phi_i\}_{i=1}^n$ be a sequence of n linearly independent elements of $\mathcal{H}_{2-}(\mathbb{E})$, and let $\Psi_n = \text{span}(\Phi_n)$. The distance $d_{\mathcal{H}_{2-}}(F, \Psi_n)$ between $F \in \mathcal{H}_{2-}(\mathbb{E})$ and Ψ_n is defined as

$$d_{\mathcal{H}_{2-}}(F, \Psi_n) = \inf_{G \in \Psi_n} \|F - G\|_{\mathcal{H}_2}. \quad (7)$$

