Chapter 5

Identification of input-output LPV models

V. Laurain[†], R. Tóth[‡], M. Gilson[†] & H. Garnier[†]

[†] *Centre de Recherche en Automatique de Nancy*
Nancy-Université, CNRS
BP 239, F-54506 Vandoeuvre-lès-Nancy Cedex, France
vincent.laurain@cran.uhp-nancy.fr
marion.gilson@cran.uhp-nancy.fr
hugues.garnier@cran.uhp-nancy.fr

[‡] *Delft Center for Systems and Control*
Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
r.toth@tudelft.nl

This chapter presents an overview of the available methods for identifying input-output LPV models both in discrete time and continuous time with the main focus on noise modeling issues. First, a least-squares approach and an instrumental variable method are presented for dealing with LPV-ARX models. Then, a refined instrumental variable approach is discussed to address more sophisticated noise models like Box-Jenkins in the LPV context. This latter approach is also introduced in continuous time and efficient solutions are proposed for both the problem of time-derivative approximation and the issue of continuous-time modeling of the noise.

1. Introduction

The identification of LPV systems can be addressed with different representations based model structures: *Input-Output* (IO) [Bamieh and Giarré (2002); Abbas and Werner (2009); Butcher *et al.* (2008); Wei and Del Re (2006)], State-Space [Lovera and Mercère (2007); van Wingerden and Verhaegen (2009); Lopes dos Santos *et al.* (2007)] or Orthogonal Basis Functions based models [Tóth *et al.* (2009a)] (see [Tóth (2010)] for an overview of the existing methods). This Chapter focuses on the IO model representation which presents several benefits. Firstly, the recent *Prediction Error Minimization* (PEM) framework extension [Tóth (2010)] makes possible the stochastic analysis of identification methods. Moreover, this representation is well suited for the introduction of different noise models [Laurain *et al.* (2010)] depending on the application considered. Finally, the tremendous need for structure learning/selection in practical applications starts finding answers for both parametric and non-parametric models [Tóth *et al.* (2009b); Tóth *et al.* (2011b)].

Most of the methods developed for IO-LPV models are derived in *discrete-time* (DT) under a linear regression form with static dependence on the scheduling variable p. Moreover, the most widely assumed structure is the *Auto Regressive with eXogeneous Input* (ARX) [Bamieh and Giarré (2002); Wei and Del Re (2006)] model. The ARX noise assumption combined to a linear parametrization of the scheduling dependencies is mainly motivated by the possible direct extension of *Linear Time Invariant* (LTI) PEM framework using linear regression approaches. Though, in the LPV context, not only the ARX assumption means equivalent dynamic properties between the process and the noise, but also identical nonlinear dependencies on the scheduling variable. In other words, the statement that the ARX assumption is unrealistic in the LTI framework is even stronger in the LPV case. However, the extension of the LTI PEM framework to the LPV case for other types of noise models appears to be non-trivial, mostly due to the fact that multiplication with q is not commutative over the p-dependent coefficients. Consequently, it has been only recently understood how a PEM framework can be established for the estimation of general LPV models [Tóth (2010)]. Even if some recursive *Least Squares* (LS) and *Instrumental Variable* (IV) methods have been introduced [Giarré *et al.* (2003); Butcher *et al.* (2008)], it was shown in [Laurain *et al.* (2010)] that despite the apparent similarity to the LTI identification problem, the methods available could not lead to the PEM in the LPV case. Furthermore, it has been detailed how linear regression based methods can still be successfully applied to more realistic noise conditions via the reformulation of the LPV model. Another concern is that even though the identification methods are mainly derived in DT, most physical processes are expressed naturally in *Continuous-Time* (CT) differential equation form. Moreover, LPV controllers are commonly synthesized in CT as stability and performance requirements of the closed-loop behavior can be more conveniently expressed in CT. Therefore, the current design tools focus on CT-LPV controller synthesis requiring accurate and low order CT models of the system. Unfortunately, the DT identification methods cannot cope with this need as the transformation from DT-LPV models to CT-LPV models is in a very immature state. The direct identification of CT-LPV models is consequently an attractive approach but presents some intrinsic issues mainly linked to the sampled nature of the data and the mathematical difficulty to handle CT stochastic processes [Garnier and Wang (2008)]. Based on the recent advances of both CT-LTI and DT-LPV identification frameworks, an IV based method for direct CT-LPV identification has been proposed in [Laurain *et al.* (2011)].

Consequently, by retracing an historical background of the IO-LPV identification framework, this Chapter aims at addressing solutions leading to the PEM in the LPV framework for different noise assumptions, both in DT and CT settings. Section 2 details the structure of DT IO-LPV data-generating systems along with the widely used polynomial model associated. Section 3 focuses on the solutions offered for the estimation of LPV-ARX models. The *Output Error* (OE) and *Box-Jenkins* (BJ) structures are introduced in Section 4, where the importance of the noise assumption as well as the influence of the identification method chosen are illustrated on a representative example. Section 5 addresses the problematic of direct identification of CT systems, presenting some intrinsic difficulties generally involved in CT model identification from sampled data. Section 6 proposes some direct CT identification solution and the advantages over the DT identification scheme are illustrated using a representative example derived from a well-known CT benchmark [Rao and Garnier (2004)]. Finally, the different contributions are summarized and some conclusions are brought in Section 7.

2. Discrete-time LPV polynomials models

DT models suit well the sampled nature of the data and consequently, most models used in the literature are expressed in a DT setting. Moreover, they often assume a static dependence on the scheduling variable p (the system depends on instantaneous values of p): the relevance of this static assumption is not discussed here as the presented methods can easily include delayed version of p in the model for dynamic dependence. This section gives a definition of the DT LPV data-generating systems using polynomial input-output representations based model structures.

2.1. *LPV data generating system*

In DT, the LPV data-generating system $\mathscr{S}_\mathrm{o}$ is commonly formulated in the identification literature using the following polynomial IO form:

$$\mathscr{S}_\mathrm{o} \begin{cases} A_\mathrm{o}(p_k, q^{-1})\chi_\mathrm{o}(k) = B_\mathrm{o}(p_k, q^{-1})u(k) \\ y(k) = \chi(k) + v_\mathrm{o}(k) \end{cases} \tag{1}$$

where p_k is the value of the scheduling parameter p at sample time k, χ_o is the noise-free output, u is the input, v_o is the additive noise with bounded spectral density, y is the noisy output of the system and q is the time-shift operator, *i.e.* $q^{-i}u(k) = u(k-i)$. The study of these systems is here restricted to the *Single-Input-Single-Ouput* (SISO) case for clarity's sake, but the solutions proposed are all extensions of LTI identification methods which have been also successfully applied to the *Multi-Input-Multi-Ouput* (MIMO) case. Therefore, the signals u and y are considered to be bounded and quasi-stationnary signals on $\mathbb{R}$. $A_\mathrm{o}(p_k, q^{-1})$ and $B_\mathrm{o}(p_k, q^{-1})$ are polynomials in q^{-1} of degree n_a and n_b respectively:

$$A_\mathrm{o}(p_k, q^{-1}) = 1 + \sum_{i=1}^{n_\mathrm{a}} a_i^\mathrm{o}(p_k)q^{-i}, \quad \text{and} \quad B_\mathrm{o}(p_k, q^{-1}) = \sum_{j=0}^{n_\mathrm{b}} b_j^\mathrm{o}(p_k)q^{-j}, \tag{2}$$

where the coefficients a_i and b_j are real meromorphic functions ($f : \mathbb{R}^n \to \mathbb{R}$ is a real meromorphic function if $f = g/h$ with g, h analytic and $h \neq 0$) with static dependence on p. It is assumed that these coefficients are bounded on $\mathbb{P}$, thus the solutions of $\mathscr{S}_\mathrm{o}$ are well-defined and the process part is completely characterized by the coefficient functions $\{a_i^\mathrm{o}\}_{i=1}^{n_\mathrm{a}}$ and $\{b_j^\mathrm{o}\}_{j=0}^{n_\mathrm{b}}$. Like in any identification problem, it is assumed that a data sequence $\mathcal{D}_N = \{y(k), u(k), p(k)\}_{k=1}^N$ generated by $\mathscr{S}_\mathrm{o}$ is available. The data-generating system in Eq. (1) is then identified by defining a model set and some information criterion characterizing the quality of the given model based on $\mathcal{D}_N$. As a next step, we explore how such model structures and criteria can be chosen to efficiently solve the identification problem.

2.2. *Polynomial LPV model*

2.2.1. *Process model*

The studied process is fully characterized by the knowledge of functions $\{a_i^{\mathrm{o}}\}_{i=1}^{n_{\mathrm{a}}}$ and $\{b_j^{\mathrm{o}}\}_{j=0}^{n_{\mathrm{b}}}$, but in practical cases these functions are *a priori* unknown nonlinear functions. Even if some recent work gives some preliminary results oriented towards a model structure selection [Tóth *et al.* (2009b); Tóth *et al.* (2011b)], the usual solution to overcome this problem is to parameterize these functions using a sum of *a priori* known basis functions. Consequently, this Chapter focuses on the usual assumption used in the literature where the process model denoted by $\mathscr{G}_\rho$ is well defined in an IO-LPV representation form:

$$\mathscr{G}_\rho : \big(A(p_k, q^{-1}, \rho), B(p_k, q^{-1}, \rho)\big) = (\mathscr{A}_\rho, \mathscr{B}_\rho) \tag{3}$$

where the p-dependent polynomials A and B are parameterized as

$$\mathscr{A}_\rho \begin{cases} A(p_k, q^{-1}, \rho) = 1 + \displaystyle\sum_{i=1}^{n_{\mathrm{a}}} a_i(p_k) q^{-i}, \\ a_i(p_k) = \rho_{i,0} + \displaystyle\sum_{l=1}^{s_i} \rho_{i,l}\psi_{i,l}(p_k), \quad i = 1, \ldots, n_{\mathrm{a}} \end{cases}$$

$$\mathscr{B}_\rho \begin{cases} B(p_k, q^{-1}, \rho) = \displaystyle\sum_{j=0}^{n_{\mathrm{b}}} b_j(p_k) q^{-j}, \\ b_j(p_k) = \rho_{i,0} + \displaystyle\sum_{l=1}^{s_i} \rho_{i,l}\psi_{i,l}(p_k), \quad i = j + n_{\mathrm{a}} + 1, \quad j = 0, \ldots, n_{\mathrm{b}}. \end{cases}$$

In this parametrization, $\{\psi_{i,l}\}_{i=1,l=1}^{n_{\mathrm{a}}+n_{\mathrm{b}}+1,s_i}$ are *a priori* chosen meromorphic functions of p, with static dependence on p, allowing the identifiability of the model (linearly independent functions on $\mathbb{P}$ for example). The associated model parameters ρ are stacked columnwise:

$$\rho = \Big[\rho_{1,0} \cdots \rho_{n_{\mathrm{a}},0}, \rho_{1,1} \cdots \rho_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}, \rho_{n_{\mathrm{a}}+1,0} \cdots \rho_{n_{\mathrm{g}},s_{n_{\mathrm{g}}}}\Big]^\top \in \mathbb{R}^{n_\rho}, \tag{4}$$

with $n_{\mathrm{g}} = n_{\mathrm{a}} + n_{\mathrm{b}} + 1$ and $n_\rho = \sum_{i=1}^{n_{\mathrm{g}}} s_i + 1$. We introduce also $\mathcal{G} = \{\mathscr{G}_\rho \mid \rho \in \mathbb{R}^{n_\rho}\}$, as the collection of all process models in the form of Eq. (3).

2.2.2. *Noise model*

Another term which needs to be modeled is the process noise $v_{\rm o}$. In the PEM framework, the noise model implicitly defines the error criterion which assesses the quality of a model when minimized. In most practical applications, information about the noise structure or properties is unavailable. It is a critical issue considering that the quality of the estimated model will highly depend on the noise assumption. Consequently, it is important to introduce a noise structure able to model a large set of behaviors. A rather general description in the IO-LPV context is a noise model denoted by $\mathscr{H}_\eta$ and defined by the following representation:

$$\mathscr{H}_\eta : D(p_k, q^{-1}, \eta)v(k) = C(p_k, q^{-1}, \eta)e(k), \tag{5}$$

where $e(k)$ is assumed to be a white noise process with a normal distribution, and the p-dependent polynomials are defined as

$$\begin{aligned} C(p_k, q^{-1}, \eta) &= 1 + \sum_{i=1}^{n_{\rm a}} c_i(p_k)q^{-i}, i = 1, \ldots, n_c \\ D(p_k, q^{-1}, \eta) &= 1 + \sum_{j=1}^{n_{\rm b}} d_j(p_k)q^{-j}, j = 1, \ldots, n_d. \end{aligned} \tag{6}$$

The focus has recently been given to simpler structures in the literature and the next sections present estimation methods relying on different simplifications of this noise model.

3. Estimating LPV-ARX models in DT

The first model structure which appeared in the IO-LPV identification framework is the LPV-ARX model [Bamieh and Giarré (2002); Butcher *et al.* (2008)]. This very specific assumption on the noise is motivated by a substantial simplification of the model which renders possible the direct application of the LTI PEM framework using regression based methods.

3.1. *LPV-ARX models*

This model is characterized by the fact that the noise model is assumed to have the same dynamics and nonlinearities as the process model. More specifically, the noise model is denoted by $\mathscr{H}_\rho$ and defined by the following

LPV-IO representation:

$$\mathscr{H}_\rho : A(p_k, q^{-1}, \rho)v(k) = e(k), \tag{7}$$

where $e(k)$ is assumed to be a white noise process. Under this noise assumption and using the process model given in Eq. (3), the full LPV-ARX model denoted $\mathscr{M}_\rho$ can be written in the linear regression form:

$$\mathscr{M}_\rho : y(k) = \varphi^\top(k)\rho + e(k) \tag{8}$$

with ρ as defined in Eq. (4), and

$$\begin{aligned}\varphi(k) = \big[-y(k-1) \quad \dots \quad &- y(k-n_\mathrm{a}), -y(k-1)\psi_{1,1}(p_k) \quad \dots \\ &- y(k-n_\mathrm{a})\psi_{n_\mathrm{a},s_{n_\mathrm{a}}}(p_k), u(k) \quad \dots \quad u(k-n_\mathrm{b}), \\ &u(k)\psi_{n_\mathrm{a}+1,1}(p_k) \quad \dots \quad u(k-n_\mathrm{b})\psi_{n_\mathrm{g},s_{n_\mathrm{g}}}(p_k)\big].\end{aligned} \tag{9}$$

It must be pointed out that under this polynomial LPV-ARX model assumption, the time-varying nature of the system is transfered to the n_ρ signals composing the regressor $\varphi(k)$. Moreover, the parameter vector ρ is time-invariant and the error equation $e(k)$ is a white noise with a normal distribution: consequently, Eq. (8) defines an LTI model.

3.2. *The LS solution*

By acknowledging that the LPV-ARX model leads to a MIMO LTI formulation of the LPV problem, the LTI PEM framework can be directly used: the LS solution dedicated to IO-LPV models was introduced by Giarré [Bamieh and Giarré (2002)] and is described here. Let $\mathcal{M} = \{\mathscr{M}_\rho \mid \rho \in \mathbb{R}^{n_\rho}\}$ be the collection of all models in the form of Eq. (8). If the studied system is well described by the ARX structure ($\mathrm{S_o} \in \mathcal{M}$), then $\mathscr{S}_\mathrm{o}$ can also be written as a linear regression:

$$y(k) = \rho_\mathrm{o}^\top \varphi(k) + e_\mathrm{o}(k), \tag{10}$$

where $\rho_\mathrm{o} \in \mathbb{R}^{n_\rho}$ is the *true* parameter vector ($\mathscr{M}_{\rho_\mathrm{o}} = \mathscr{S}_\mathrm{o}$) and $\varphi(k) \in \mathbb{R}^{n_\rho}$ is the regressor vector (9). In order to estimate ρ_o, the quality of the model fit is formulated in terms of a cost function $\mathscr{V}(\rho, \mathcal{D}_N)$. The minimization of $\mathscr{V}(\rho, \mathcal{D}_N)$ corresponds to the estimation of ρ. In case of the so-called LS estimation, the cost function is the *squared equation error*:

$$\mathscr{V}(\rho, \mathcal{D}_N) = \frac{1}{N}\sum_{k=1}^{N} e^2(k) = \frac{1}{N}\|e(k)\|_{\ell_2}^2, \tag{11}$$

where e is given by Eq. (8). $\mathcal{M}$ represents the set of models in which we are searching for the "best" $\mathscr{M}_\rho$ that describes $\mathscr{S}_\mathrm{o}$ given a dataset $\mathcal{D}_N = \{y(k), u(k), p(k)\}_{k=1}^N$ generated by $\mathscr{S}_\mathrm{o}$. As Eq. (11) is quadratic in ρ and under this LTI representation, the LS solution is analytic and given by:

$$\hat{\rho}_{\mathrm{LS}} = \left[\sum_{k=1}^{N} \varphi(k)\varphi^\top(k)\right]^{-1} \sum_{k=1}^{N} \varphi(k)y(k). \tag{12}$$

If the data generating the system is actually in the model set defined, Eq. (12) is a statistically optimal estimator (minimum variance and unbiased). Nonetheless, it must be pointed out that this model assumption is very limited in practice. Even though it might be a fair assumption to consider that the process is well parameterized using Eq. (3), the probability that $v_o(t_k)$ is correctly described by Eq. (7) is very low. Indeed, in most cases, there is no reasonable explanation to justify why the noise v_o and the process part of $\mathscr{S}_\mathrm{o}$ should contain the same dynamics and nonlinearities. In terms of estimation, it means that using the LS method in practical applications will most often lead to biased estimates. Consequently, some methods have been developed in order to cope with the error induced by this invalid assumption on the noise. A solution introduced in [Butcher *et al.* (2008)] and relying on the IV method is described in the next section.

3.3. *Non-white noise: An IV solution*

In the LTI context, the PEM methods have evolved in order to deal with general and realistic types of noise [Ljung (1999)]. Their efficiency and convergence have been proven for generic types of models such as: ARX, OE or BJ for example. The most common assumption is the inclusion of the system studied in the model set defined, both for the process and the noise. Nonetheless, the physical insight of the noise process is very unrealistic in practice and some methods have especially been designed to consistently estimate models when the noise assumption is incorrect: a set of these approaches are the IVs methods [Young and Jakeman (1980); Söderström and Stoica (1983)].

3.3.1. *Instrumental variable method in the LTI case*

It is well-known from the LTI theory that when considering a regression form $y(k) = \varphi(k)\rho + \tilde{v}(k)$, the LS estimate from Eq. (12) is also:

$$\hat{\rho}_{\rm LS} = \rho_{\rm o} + \left[\sum_{k=1}^{N} \varphi(k)\varphi^\top(k)\right]^{-1} \sum_{k=1}^{N} \varphi(k)\tilde{v}(k). \quad (13)$$

This equation induces that $\hat{\rho}_{LS}$ is a consistent estimate of $\rho_{\rm o}$ (unbiased for finite data) if both the following conditions are respected [a]:

CLS1 $\bar{\mathbb{E}}\{\varphi(k)\varphi^\top(k)\}$ is full column rank.

CLS2 $\bar{\mathbb{E}}\{\varphi(k)\tilde{v}(k)\} = 0$.

While CLS1 is usually respected if the input is exciting enough, CLS2 only holds when $\tilde{v}(k)$ is a white noise. Unfortunately, this assumption is only verified for the ARX case [Ljung (1999)]. Hence, as previously pointed out, using the LS estimate is not the most pertinent choice in most practical applications. The original aim of IV methods is to cope with the fact that in most cases, $\tilde{v}(k)$ is a colored process. The idea is to introduce an instrument ζ which is used to produce a consistent estimate independently on the noise model taken. The IV estimate is given as:

$$\hat{\rho}_{\rm IV} = \left[\sum_{k=1}^{N} \zeta(k)\varphi^\top(k)\right]^{-1} \sum_{k=1}^{N} \zeta(k)y(k), \quad (14)$$

and using this definition, it can be seen that

$$\hat{\rho}_{\rm IV} = \rho_{\rm o} + \left[\sum_{k=1}^{N} \zeta(k)\varphi^\top(k)\right]^{-1} \sum_{k=1}^{N} \zeta(k)\tilde{v}(k). \quad (15)$$

Therefore, and similarly to the LS solution $\hat{\rho}_{\rm IV}$ is a consistent estimate of $\rho_{\rm o}$ if

C1 $\bar{\mathbb{E}}\{\zeta(k)\varphi^\top(k)\}$ is full column rank.

C2 $\bar{\mathbb{E}}\{\zeta(k)\tilde{v}(k)\} = 0$.

There is a considerable amount of freedom in the choice of an instrument respecting simultaneously C1 and C2. Nonetheless, even if any of these instruments leads to a consistent estimate, the amount of correlation between the instrument and the regressor will have a serious impact on the variance of the estimated parameters.

[a]The notation $\bar{\mathbb{E}}\{.\} = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \mathbb{E}\{.\}$ is adopted from the prediction error framework of [Ljung (1999)].

3.3.2. *The optimal instrument for LPV-ARX models*

In the LTI context, the choice of the instrument has been widely discussed and most advanced IV methods offer similar performance as extended LS methods or other PEM methods (see [Rao and Garnier (2004); Ljung (2009)]) while providing consistent results even for an imperfect noise structure. It can be proven that under the ARX model assumption, the variance of the IV estimate is minimal if the instrument is chosen as the noise-free version of the regressor. In other words, when directly applying the IV theory to the LPV-ARX model from Eq. (8) (the LPV-ARX model is an LTI model), the optimal IV estimate is given by:

$$\hat{\rho}_{\mathrm{IV}}^{\mathrm{opt}} = \left[\sum_{k=1}^{N} \zeta_{\mathrm{opt}}(k)\varphi^{\top}(k)\right]^{-1} \sum_{k=1}^{N} \zeta_{\mathrm{opt}}(k) y(k), \tag{16}$$

where the optimal instrument is defined as [Söderström and Stoica (1983)]:

$$\begin{aligned} \zeta_{\mathrm{opt}}(k) = \big[-\chi_{\mathrm{o}}(k-1) \quad \dots \quad -\chi_{\mathrm{o}}(k-n_{\mathrm{a}}), -\chi_{\mathrm{o}}(k-1)\psi_{1,1}(p_k) \quad \dots \\ -\chi_{\mathrm{o}}(k-n_{\mathrm{a}})\psi_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}(p_k), u(k) \quad \dots \quad u(k-n_{\mathrm{b}}), \\ u(k)\psi_{n_{\mathrm{a}}+1,1}(p_k) \quad \dots \quad u(k-n_{\mathrm{b}})\psi_{n_{\mathrm{g}},s_{n_{\mathrm{g}}}}(p_k)\big] \end{aligned} \tag{17}$$

It can be noticed that in the LPV-ARX assumption both the IV solution from Eq. (16) and the LS solution from Eq. (12) are statistically optimal estimators. Nonetheless, the deterministic output terms χ_{o} contained in ζ_{opt} are *a priori* unknown in practice. Consequently, some estimates of χ_{o} commonly compose the instrument, requiring some iterative process and implying some difficulties for the convergence analysis. It could consequently be argued that the LS estimate is simpler and hence, better suited to solve the identification problem: this statement is true as long as the true system respects the ARX structure. Nonetheless, this is even more unlikely in the LPV context than in the LTI context: the ARX assumption implies that not only the dynamics of the noise and process are the same but that the measuring devices are affected by the external scheduling variable in the exact same nonlinear way as the system studied. In the probable situation where the noise assumption is violated, the IV method is a consistent estimate while the LS is inevitably biased.

3.3.3. *The LPV-IV4 method*

As the knowledge of the deterministic output χ_{o} is unavailable, an algorithm is needed to approximate the deterministic output terms and to build the optimal instrument ζ_{opt}. Therefore, the use of an IV method similar to the IV4 method ([Ljung (1999)]) is proposed in [Butcher *et al.* (2008)], where the instrument is built using the simulated data generated from an estimated auxiliary ARX model. This method is detailed as:

Algorithm 1 (LPV-IV4 method).

Step 1 *Estimate an ARX model by the LS method (minimizing Eq. (11)) using the extended regressor* (9)*.*

Step 2 *Generate an estimate $\hat{\chi}(k)$ of $\chi(k)$ based on the resulting ARX model of the previous step. Build an instrument based on $\hat{\chi}(k)$ and then estimate ρ using the IV method.*

In [Butcher *et al.* (2008)], it was shown that according to the theory, in case $\mathscr{S}_{\text{o}}$ corresponds to an LPV-OE model ($v_{\text{o}} = e_{\text{o}}$), Algorithm 1 leads to an unbiased estimate. Moreover, like in the LTI case, the noise modeling error results in a bias for the LS estimates while only the variance is affected when using this IV method. Nevertheless, the bigger the difference between the true noise process and the noise model assumed is, the higher the resulting variance in the IV estimates will be. Depending on the size of the dataset, the increasing variance of the IV estimates can lead to less relevant results than the LS method (for which the variance is known to remain low). Consequently, it is highly important to assume a noise model as realistic as possible in the first place. In the LTI case, many IV methods are dedicated to more general noise models such as OE of BJ [Young (2008)]. The next section describes some available methods for LPV-OE and LPV-BJ which were introduced in [Laurain *et al.* (2010)].

4. Addressing estimation with general noise models

In order to assess more efficiently the quality of a given model, more realistic noise models have to be taken into consideration. Hence, this section introduces noise structures which have been widely used in the LTI framework and which have not been dealt with in the LPV case until recently.

4.1. *LPV-BJ model*

The general LPV-BJ model structure from Eq. (5) has never been dealt with in the literature so far. Nevertheless, as a preliminary step towards this p-dependent noise, it can be assumed that the rational spectral density is not dependent on p. In the case of high-tech applications where the major source of noise is due to measurement error (e.g. wafer-scanners), this assumption often holds true and is in general far more realistic than the LPV-ARX assumption.

4.1.1. *Noise model*

In the defined LPV-BJ structure, the noise model denoted by $\mathscr{H}$ is defined as a DT *autoregressive moving average* (ARMA) model

$$\mathscr{H}_\eta : (H(q,\eta)), \tag{18}$$

where H is a monic rational function given in the form of

$$H(q,\eta) = \frac{C(q^{-1},\eta)}{D(q^{-1},\eta)} = \frac{1 + c_1 q^{-1} + \ldots + c_{n_\mathrm{c}} q^{-n_\mathrm{c}}}{1 + d_1 q^{-1} + \ldots + d_{n_\mathrm{d}} q^{-n_\mathrm{d}}} \tag{19}$$

and all roots of $z^{n_\mathrm{d}} D(z^{-1})$ and $z^{n_\mathrm{c}} C(z^{-1})$ are inside the unit disc. It can be noticed that in case $C(q^{-1}) = D(q^{-1}) = 1$, Eq. (19) defines an OE model. The associated model parameters η are stacked columnwise in the parameter vector,

$$\eta = \begin{bmatrix} c_1 \ldots c_{n_\mathrm{c}} \; d_1 \ldots d_{n_\mathrm{d}} \end{bmatrix}^\top \in \mathbb{R}^{n_\eta}, \tag{20}$$

where $n_\eta = n_\mathrm{c} + n_\mathrm{d}$. Additionally, denote $\mathcal{H} = \{\mathscr{H}_\eta \mid \eta \in \mathbb{R}^{n_\eta}\}$, the collection of all noise models in the form of Eq. (18).

4.1.2. *Whole model*

With respect to a given process and noise part $(\mathscr{G}_\rho, \mathscr{H}_\eta)$, the parameters can be collected as $\theta = [\,\rho^\top \; \eta^\top\,]$ and the signal relations of the LPV-BJ model $\mathscr{M}_\theta$, are defined as:

$$\mathscr{M}_\theta \begin{cases} A(p_k, q^{-1}, \rho)\chi(k) = B(p_k, q^{-1}, \rho)u(k) \\ v(k) = \dfrac{C(q^{-1},\eta)}{D(q^{-1},\eta)} e(k) \\ y(k) = \chi(k) + v(k) \end{cases} \tag{21}$$

Based on this model structure, the model set, denoted as $\mathcal{M}$, with process ($\mathscr{G}_\rho$) and noise ($\mathscr{H}_\eta$) models parameterized independently, takes the form

$$\mathcal{M} = \left\{ (\mathscr{G}_\rho, \mathscr{H}_\eta) \;\mid\; [\rho^\top, \eta^\top]^\top = \theta \in \mathbb{R}^{n_\rho + n_\eta} \right\}. \tag{22}$$

4.1.3. *The LPV-BJ model issue*

If the system belongs to the model set defined in Eq. (22), then $y(k)$ can be written in the linear regression form:

$$y(k) = \varphi^\top(k)\rho + \tilde{v}(k) \tag{23}$$

with ρ and $\varphi(k)$ as defined in Eq. (4) and Eq. (9), respectively. Under the LPV-BJ assumption, it can be seen from Eq. (21) that

$$\tilde{v}(k) = A(p_k, q^{-1}, \rho)\frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)}e(k) = Q(p_k, q^{-1}, \rho, \eta)e(k). \tag{24}$$

Hence, using Eq. (12) to estimate ρ corresponds to the minimization of the LS criterion on $\tilde{v}(k)$, which is clearly different from the PEM (the minimization of the LS criterion on $e(k)$) [Laurain *et al.* (2010)]. It must be pointed out that despite the apparent similarity between Eq. (8) and Eq. (23), the estimation problem of the LPV-BJ model cannot be considered as an LTI problem like it is under the LPV-ARX assumption. This statement is motivated by the presence of the filter Q in Eq. (24) which is obviously p-dependent. In an LTI context, Q is p-independent and most estimation methods use the inverse filtering of Q to solve the estimation problem in an optimal way. Nonetheless, in this LPV case, it was shown in [Laurain *et al.* (2010)] that the non-commutativity of LPV filters as well as the impossibility to compute the exact inverse of an LPV filter in practice renders the PEM impossible under the regression form (23). A possible solution to overcome this problem is to reformulate the model from Eq. (21) in order to use directly the LTI theory. This reformulation is detailed in the next section.

4.1.4. *Reformulation of the model equations*

In order to introduce a method which provides a solution to the identification problem of LPV-BJ models, rewrite the signal relations of Eq. (21)

as

$$\mathscr{M}_\theta \begin{cases} \underbrace{\chi(k) + \sum_{i=1}^{n_\mathrm{a}} a_{i,0}\chi(k-i)}_{F(q^{-1},\rho)\chi(k)} + \sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{s_i} a_{i,l}\underbrace{\psi_{i,l}(p_k)\chi(k-i)}_{\chi_{i,l}(k)} \\ \qquad\qquad = \sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{s_{\tilde{j}}} b_{j,l}\underbrace{\psi_{\tilde{j},l}(p_k)u(k-j)}_{u_{j,l}(k)} \\ v(k) = \dfrac{C(q^{-1},\eta)}{D(q^{-1},\eta)}e(k) \\ y(k) = \chi(k) + v(k), \end{cases} \tag{25}$$

where $F(q^{-1},\rho) = 1 + \sum_{i=1}^{n_\mathrm{a}} a_{i,0}q^{-i}$, $\tilde{j} = j + n_\mathrm{a} + 1$ and $\psi_{\tilde{j},0}(k) = 1$. Note that in this way, the LPV-BJ model is rewritten as a Multiple Input Single Output (MISO) system with $\sum_{i=1}^{n_\mathrm{g}} s_i + 1$ inputs $\{\chi_{i,l}\}_{i=1,l=1}^{n_\mathrm{a},s_i}$ and $\{u_{j,l}\}_{j=0,l=0}^{n_\mathrm{b},s_{\tilde{j}}}$ as represented in Fig. 1. Given the fact that the polynomial operator commutes in this representation ($F(q^{-1},\rho)$ does not depend on p_k), Eq. (25) can be rewritten as

$$y(k) = -\sum_{i=1}^{n_\mathrm{a}}\sum_{l=1}^{s_i} \frac{a_{i,l}}{F(q^{-1},\rho)}\chi_{i,l}(k) + \sum_{j=0}^{n_\mathrm{b}}\sum_{l=0}^{s_{\tilde{j}}} \frac{b_{j,l}}{F(q^{-1},\rho)}u_{j,l}(k) + H(q,\eta)e(k), \tag{26}$$

which is an LTI representation. Consequently, it becomes possible to apply the LTI theory in order to solve the estimation problem.

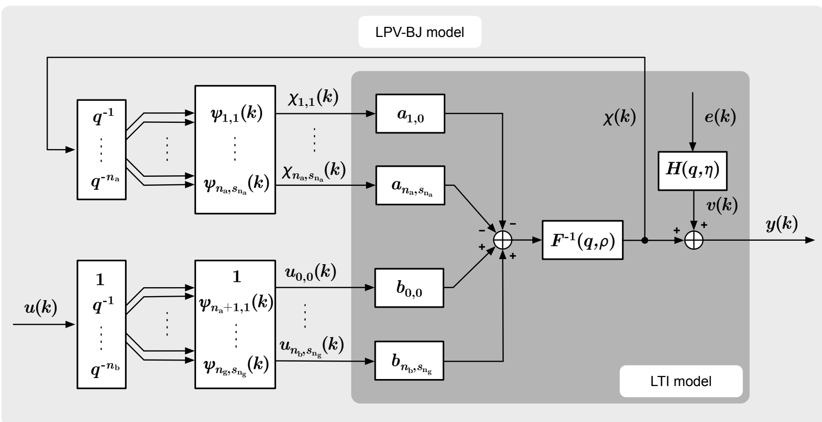

Fig. 1. MISO LTI interpretation of the LPV-BJ model.

4.2. *A refined instrumental approach*

Based on the MISO-LTI formulation Eq. (26), it becomes possible in theory to achieve optimal PEM using linear regression. The only method available for LPV-BJ models in the literature so far is an extended version of the *Refined IV* (RIV) approach of the LTI identification framework [Laurain *et al.* (2010)] which provides an efficient way of identifying LPV-BJ models. It derives directly from the extended IV scheme developed for LTI-BJ models which is described in the next section.

4.2.1. *Instrumental variable for LTI-BJ models*

Based on Eq. (26), it is possible to introduce some LS-based methods to solve the identification problem. However, it is important to realize that in experimental conditions, even if a realistic noise assumption will lead to more accurate estimates, the chances that the true noise process is not perfectly described are high. In this case, similarly to the LPV-ARX case, an LS based method will produce biased results. Therefore, the need to cope with noise modeling errors remains. As a refinement of the IV scheme presented in Section 3.3, IV methods have been developed to cope with more general structures such as the BJ case, which can be expressed in the linear regression form $y(k) = \varphi^\top(k)\rho + Q(q)e(k)$ ($Q(q)$ is an LTI transfer function, with $Q^{-1}(q)$ stable and $e(k)$ a white noise). Based on this form, the extended-IV estimate can be given as [Söderström and Stoica (1983)]:

$$\hat{\rho}_{\mathrm{XIV}}(N) = \arg \min_{\rho \in \mathbb{R}^{n_\rho}} \left\| \left[\frac{1}{N} \sum_{k=1}^{N} L(q)\zeta(k)L(q)\varphi^\top(k) \right] \rho - \left[\frac{1}{N} \sum_{t=1}^{N} L(q)\zeta(k)L(q)y(k) \right] \right\|_W^2, \quad (27)$$

where $\zeta(k)$ is the instrument, $\|x\|_W^2 = x^T W x$, with W a positive definite weighting matrix and $L(q)$ is a stable prefilter. The conditions for consistency now read:

CX1 $\bar{\mathbb{E}}\{L(q)\zeta(k)L(q)\varphi^\top(k)\}$ is full column rank.

CX2 $\bar{\mathbb{E}}\{L(q)\zeta(k)L(q)\tilde{v}(k)\} = 0$.

Again, there is a considerable amount of freedom in the choice of the instruments. Nonetheless, if the studied system belongs to the model set defined, *i.e.* if there exists ρ_{o} and Q_{o} (Q_{o}^{-1} being stable) such that

$$\mathscr{S}_{\mathrm{o}} : y(k) = \varphi^{\top}(k)\rho_{\mathrm{o}} + Q_{\mathrm{o}}(q)e_{\mathrm{o}}(k), \tag{28}$$

with $e_{\mathrm{o}}(k)$ a white noise, it has been shown in [Söderström and Stoica (1983)] and [Young (1984)] that the minimum variance estimator can be achieved in the BJ case if:

CX3 $W = I$.

CX4 ζ is chosen as the noise-free version of the extended regressor φ.

CX5 $L(q)$ is chosen as $Q_{\mathrm{o}}^{-1}(q)$.

In case of noise modeling error, the extended IV method is consistent and the variance of the estimates should be significantly decreased with respect to the IV4 method: even if the noise process is not in the noise model set defined, it is more likely to be better described by the BJ model than by the ARX model.

4.2.2. *The optimal instrument for LPV-BJ systems*

Using Eq. (26), $y(k)$ can be written in the regression form:

$$y(k) = \varphi^{\top}(k)\rho + \tilde{v}(k), \tag{29}$$

where

$$\begin{aligned}
\varphi(k) = \Big[&-y(k-1) \ \dots \ -y(k-n_{\mathrm{a}}) \ -\chi_{1,1}(k) \ \dots \\
&-\chi_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}(k) \ u_{0,0}(k) \ \dots \ u_{n_{\mathrm{b}},s_{n_{\mathrm{g}}}}(k) \Big]^{\top}, \\
\tilde{v}(k) = &F(q^{-1},\rho)v(k), \\
= &F(q^{-1},\rho)\frac{C(q^{-1},\eta)}{D(q^{-1},\eta)}e(k) = Q(q,\rho,\eta)e(k).
\end{aligned} \tag{30}$$

It is important to notice that Eq. (29) and Eq. (23) are not equivalent as the extended regressor in Eq. (29) contains the noise-free output terms $\{\chi_{i,l}\}$. Therefore, by momentary assuming that $\{\chi_{i,l}(k)\}_{i=1,l=0}^{n_{\mathrm{a}},s_i}$ are known *a priori* and that the data-generating system $\mathscr{S}_{\mathrm{o}}$ is in the model set defined ($v_{\mathrm{o}}(k) = C_{\mathrm{o}}(q^{-1},\eta)/D_{\mathrm{o}}(q^{-1},\eta)e_{\mathrm{o}}(k)$ and $e_{\mathrm{o}}(k)$ is a white noise), then the conditions for optimal estimates expressed in [CX3-CX5] lead to the choice of optimal instrument given in this LPV-BJ case as [Laurain *et al.* (2010)]:

$$\zeta_{\mathrm{opt}}(k) = \big[-\chi_{\mathrm{o}}(k-1) \quad \ldots \quad -\chi_{\mathrm{o}}(k-n_{\mathrm{a}}) \quad -\chi^{\mathrm{o}}_{1,1}(k) \\ -\chi^{\mathrm{o}}_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}(k) \quad u_{0,0}(k) \quad \ldots \quad u_{n_{\mathrm{b}},s_{n_{\mathrm{g}}}}(k)\big]^{\top} \tag{31}$$

while the optimal filter is given as:

$$L_{\mathrm{opt}}(q) = \frac{D_{\mathrm{o}}(q^{-1})}{C_{\mathrm{o}}(q^{-1})F_{\mathrm{o}}(q^{-1})}. \tag{32}$$

4.2.3. *Iterative LPV-RIV algorithm for BJ models*

In a practical situation, the optimal instrument (31) and filter (32) are unknown *a priori*. Therefore, the RIV estimation normally involves an iterative (or relaxation) algorithm in which, at each iteration, an 'auxiliary model' is used to generate the instrument (which guarantees CX2), as well as the associated filters. This auxiliary model is based on the parameter estimates obtained at the previous iteration. Consequently, if convergence occurs, CX4 and CX5 are fulfilled. Based on the previous considerations, the iterative scheme of the RIV algorithm dedicated to the LPV case is given as:

Algorithm 2 (LPV-RIV).

Step 1 *Generate an initial estimate of the process model parameter* $\hat{\rho}^{(0)}$ *(e.g. using the LS method). Set* $C(q^{-1},\hat{\eta}^{(0)}) = D(q^{-1},\hat{\eta}^{(0)}) = 1$*. Set* $\tau = 0$.

Step 2 *Compute an estimate of* $\chi(k)$ *via*

$$A(p_k, q^{-1}; \hat{\rho}^{(\tau)})\hat{\chi}(k) = B(p_k, q^{-1}; \hat{\rho}^{(\tau)})u(k),$$

where $\hat{\rho}^{(\tau)}$ *is the estimated obtained at the previous iteration. Based on* $\mathscr{M}_{\hat{\theta}^{(\tau)}}$*, deduce* $\{\hat{\chi}_{i,l}(k)\}_{i=1,l=0}^{n_{\mathrm{a}},s_i}$ *as given in Eq. (25).*

Step 3 *Compute the filter:*

$$L(q^{-1},\hat{\theta}^{(\tau)}) = \frac{D(q^{-1},\hat{\eta}^{(\tau)})}{C(q^{-1},\hat{\eta}^{(\tau)})F(q^{-1},\hat{\rho}^{(\tau)})}$$

and the associated filtered signals $\{u^{\mathrm{f}}_{j,l}(k)\}_{j=0,l=0}^{n_{\mathrm{b}},s_{\bar{j}}}$, $y_{\mathrm{f}}(k)$ *and* $\{\chi^{\mathrm{f}}_{i,l}(k)\}_{i=1,l=0}^{n_{\mathrm{a}},s_i}$.

Step 4 *Build the filtered estimated regressor $\hat{\varphi}_{\mathrm{f}}(k)$ and in terms of CX4 the filtered instrument $\hat{\zeta}_{\mathrm{f}}(k)$ as:*

$$\hat{\varphi}_{\mathrm{f}}(k) = \Big[-y_{\mathrm{f}}(k-1) \ldots -y_{\mathrm{f}}(k-n_{\mathrm{a}}) \; -\hat{\chi}^{\mathrm{f}}_{1,1}(k) \ldots -\hat{\chi}^{\mathrm{f}}_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}(k) \; u^{\mathrm{f}}_{0,0}(k) \ldots u^{\mathrm{f}}_{n_{\mathrm{b}},s_{n_{\mathrm{g}}}}(k) \Big]^{\top}$$

$$\hat{\zeta}_{\mathrm{f}}(k) = \Big[-\hat{\chi}_{\mathrm{f}}(k-1) \ldots -\hat{\chi}_{\mathrm{f}}(k-n_{\mathrm{a}}) \; -\hat{\chi}^{\mathrm{f}}_{1,1}(k) \ldots -\hat{\chi}^{\mathrm{f}}_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}(k) \; u^{\mathrm{f}}_{0,0}(k) \ldots u^{\mathrm{f}}_{n_{\mathrm{b}},s_{n_{\mathrm{g}}}}(k) \Big]^{\top}$$

Step 5 *The IV solution is obtained as*

$$\hat{\rho}^{(\tau+1)}(N) = \left[\sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(k) \hat{\varphi}_{\mathrm{f}}^{\top}(k) \right]^{-1} \sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(k) y_{\mathrm{f}}(k).$$

The resulting $\hat{\rho}^{(\tau+1)}(N)$ is the IV estimate of the process model associated parameter vector at iteration $\tau + 1$ based on the prefiltered input/output data.

Step 6 *An estimate of the noise signal v is obtained as*

$$\hat{v}(k) = y(k) - \hat{\chi}(k, \hat{\rho}^{(\tau)}). \tag{33}$$

Based on $\hat{v}$, the estimation of the noise model parameter vector $\hat{\eta}^{(\tau+1)}$ follows, using in this case the ARMA estimation algorithm of the MATLAB identification toolbox (an IV approach can also be used for this purpose, see [Young (2008)]).

Step 7 *If $\theta^{(\tau+1)}$ has converged or the maximum number of iterations is reached, then stop, else increase τ by 1 and go to Step 2.*

4.2.4. *LPV-SRIV algorithm for OE models*

Based on a similar concept, the so-called *simplified* LPV-RIV (LPV-SRIV) method, can also be developed for the estimation of LPV-OE models. This method is based on a model structure Eq. (21) with $C(q^{-1}, \eta) = D(q^{-1}, \eta) = 1$ and consequently, Step 6 of Algorithm 2 can be skipped. In practical cases, it is a fair assumption to consider that the noise model assumed is incorrect for both LPV-OE and LPV-BJ models. In this case, the LPV-SRIV algorithm might perform as well as the LPV-RIV algorithm: the BJ assumption might be more realistic, but this is compensated by the reduced number of parameters to be estimated under the OE assumption.

4.3. *Examples*

This section presents the performance of the presented algorithm on a representative set of examples.

4.3.1. *Data-generating system*

The system taken into consideration is inspired by the example in [Butcher *et al.* (2008)] and is mathematically described as

$$\mathscr{S}_{\rm o}\begin{cases} A_{\rm o}(q,p_k) = 1 + a_1^{\rm o}(p_k)q^{-1} + a_2^{\rm o}(p_k)q^{-2} \\ B_{\rm o}(q,p_k) = b_0^{\rm o}(p_k)q^{-1} + b_1^{\rm o}(p_k)q^{-2} \\ v_{\rm o}(k) = \dfrac{1}{1-q^{-1}+0.2q^{-2}}e_{\rm o}(k), \end{cases} \tag{34}$$

where $v(k) = H_{\rm o}(q)e(k)$ and

$$a_1^{\rm o}(p_k) = 1 - 0.5p_k - 0.1p_k^2,\ a_2^{\rm o}(p_k) = 0.5 - 0.7p_k - 0.1p_k^2, \tag{35a}$$
$$b_0^{\rm o}(p_k) = 0.5 - 0.4p_k + 0.01p_k^2,\ b_1^{\rm o}(p_k) = 0.2 - 0.3p_k - 0.02p_k^2. \tag{35b}$$

4.3.2. *Model assumptions*

In order to identify the presented data generating system by using the different methods described, three different model structures are defined: LPV-ARX, LPV-OE and LPV-BJ. In the example presented, all the methods assume a correct process model by using the representation:

$$\mathscr{G}_\rho\begin{cases} A(p_k,q^{-1},\rho) = 1 + a_1(p_k)q^{-1} + a_2(p_k)q^{-2} \\ B(p_k,q^{-1},\rho) = b_0(p_k)q^{-1} + b_1(p_k)q^{-2}, \end{cases}$$

where

$$a_1(p_k) = a_{1,0} + a_{1,1}p_k + a_{1,2}p_k^2,\quad a_2(p_k) = a_{2,0} + a_{2,1}p_k + a_{2,2}p_k^2 \tag{36a}$$
$$b_0(p_k) = b_{0,0} + b_{0,1}p_k + b_{0,2}p_k^2,\quad b_1(p_k) = b_{1,0} + b_{1,1}p_k + b_{1,2}p_k^2 \tag{36b}$$

Both the LS and LPV-IV4 methods [Bamieh and Giarré (2002); Butcher *et al.* (2008)] presented in Section 3 assume an LPV-ARX model structure and therefore the following noise model:

$$\mathscr{H}_\rho^{\rm LS,IV4}\Big\{A(p_k,q^{-1},\rho)v(t_k) = e(t_k).$$

The LPV Simplified RIV (LPV-SRIV) assumes a more realistic type of noise using the OE hypothesis:

$$\mathscr{H}_{\rho}^{\mathrm{LPV-SRIV}} \Big\{ v(t_k) = e(t_k).$$

Finally, the LPV *Refined Instrumental Variable* method (LPV-RIV) represents the situation where both the process and noise are correcty modeled using the following LPV-BJ representation:

$$\mathscr{H}_{\eta}^{\mathrm{LPV-RIV}} \Big\{ v(t_k) = \frac{1}{1 + d_1 q^{-1} + d_2 q^{-2}} e(t_k).$$

The robustness of the proposed and existing algorithms are investigated with respect to different *signal-to-noise ratios* SNR $= 10 \log \frac{P_{\chi_o}}{P_{e_o}}$, where P_{χ_o} and P_{e_o} are the average power of signals χ_o and e_o, respectively.

4.3.3. *Results*

In the upcoming examples, the scheduling signal p is considered as a periodic function of time: $p_k = 0.5 \sin(0.35\pi k) + 0.5$. The input $u(k)$ is taken as a white noise with a uniform distribution $\mathscr{U}(-1, 1)$ and with length $N = 4000$ to generate data sets $\mathcal{D}_N$ of $\mathscr{S}_o$. To provide representative results, a Monte Carlo simulation is realized based on $N_{\mathrm{run}} = 100$ random realization, with the Gaussian white noise e_o being selected randomly for each realization at different noise levels: 15dB, 10dB, 5dB and 0dB. With respect to the considered methods, Table 1 shows the norm of the bias (BN) and standard deviation (SN) of the estimated parameter vector with

$$\begin{cases} \mathrm{BN} = ||\rho_o - \bar{\mathbb{E}}(\hat{\rho})||_{\ell_2}, \\ \mathrm{SN} = ||\bar{\mathbb{E}}(\hat{\rho} - \bar{\mathbb{E}}(\hat{\rho}))||_{\ell_2}, \end{cases} \tag{37}$$

where $\bar{\mathbb{E}}$ is the mean operator over the Monte Carlo simulation. The table also displays the mean number of iterations (Nit) the algorithms needed to converge to the estimated parameter vector.

It can be seen from Table 1 that the IV methods are unbiased according to the theoretical results. It might not appear clearly for the LPV-IV4 method when using SNR under 10dB but considering the variances induced, the bias is only due to the relatively low number of simulation runs. The LPV-SRIV method offers satisfying results, considering that the noise model assumption is also not correct. The better behavior of the LPV-SRIV method with respect to the LPV-IV4 method is justified by two facts. On the one hand, the LPV-IV4 method suffers from the ARX assumption which

Table 1. Estimator bias and variance norm at different SNR.

Method		15dB	10dB	5dB	0dB
LPV-LS	BN	2.9107	3.2897	3.0007	2.8050
	SN	0.0074	0.0151	0.0215	0.0326
LPV-IV4	BN	0.1961	1.8265	6.9337	10.85
	SN	1.3353	179.42	590.78	11782
LPV-SRIV	BN	0.0072	0.0426	0.1775	0.2988
	SN	0.0149	0.0537	0.4425	0.4781
	Nit	22	22	25	30
LPV -RIV	BN	0.0068	0.0184	0.0408	0.1649
	SN	0.0063	0.0219	0.0696	0.2214
	Nit	31	30	30	32

is not as realistic as the OE assumption for the considered BJ system. On the other hand, the LPV-SRIV method profits from an iterative process which leads to a close estimate of the optimal instrument and filter. The consequences are a variance in close range to the LS method for the LPV-SRIV algorithm while the LPV-IV4 method can be considered as irrelevant for a SNR under 10dB. Finally, and for SNR down to 5dB, the LPV-RIV method produces variance in the estimated parameters which is very close to the LPV-LS method, not mentioning that the bias has been completely suppressed. This latter result shows on this example, that the LPV-RIV algorithm dramatically improves the accuracy of the estimates.

4.4. *LPV noise system example*

The assumption of a p-independent noise model assumption is a restriction in comparison to the general case presented in Section 2.2.1. In practice, the true noise might actually be of a p-dependent nature, in which case none of the presented algorithms is suited for statistically optimal estimation. Any IV-based method is unbiased as long as the instrument is not correlated to noise and the noise modeling error will mainly have impact on the variance. Furthermore, the variance of the estimators will depend on how close the noise model and the true noise process are. The example chosen is derived from the system (34) in which the noise model is turned into an LPV model and the studied system is consequently expressed by:

$$\mathscr{S}_\mathrm{o}\begin{cases} A_\mathrm{o}(q,p_k) = 1 + a_1^\mathrm{o}(p_k)q^{-1} + a_2^\mathrm{o}(p_k)q^{-2}, \\ B_\mathrm{o}(q,p_k) = b_0^\mathrm{o}(p_k)q^{-1} + b_1^\mathrm{o}(p_k)q^{-2}, \\ \left(1 + d_1^\mathrm{o}(p_k)q^{-1} + d_2^\mathrm{o}(p_k)q^{-2}\right) v_\mathrm{o}(k) = e_\mathrm{o}(k), \end{cases} \tag{38}$$

where a_i^o and b_j^o are chosen as in Eq. (36a-b) and

$$d_1^\mathrm{o}(p_k) = -1 + 0.2p_k - 0.1p_k^2, \quad d_2^\mathrm{o}(p_k) = 0.2 - 0.1p_k + 0.05p_k^2. \tag{39}$$

The excitation conditions are the same as previously, and the results of a Monte Carlo simulation are exposed in Table 2 with a new realization of e_o for each run. The compared methods and their respective associated models are the same as the ones considered in Section 4.3.2. The Monte Carlo simulation results are based on $N_\mathrm{run} = 100$ using a SNR = 10dB and a number of samples $N = 4000$. Again, the bias and standard deviation norms of the resulting estimates are analyzed and given in Table 2. From this example, it can be seen that the methods LPV-LS, LPV-IV4, LPV-SRIV give similar performance as in the previous example. This can be easily explained considering that in both examples, the noise assumption is false for these methods. Consequently, the bias of the LPV-LS method is identical, the LPV-IV4 method displays a very large parameter estimate variance, while the LPV-SRIV technique keeps a low variance and removes the bias. The variance for the LPV-RIV parameter estimates is larger than in the p-independent noise case: it can be explained by the wrong noise model assumption in this example. Nonetheless, the standard deviation remains lower than the one obtained with LPV-SRIV. In practice, for more general type of noise (e.g. if it cannot be expressed in a stochastic way), it might be better to use LPV-SRIV as the number of parameters to be estimated is lower.

Table 2. Results for a LPV noise model.

	LPV-LS	LPV-IV4	LPV-SRIV	LPV-RIV
BN	3.4943	4.2293	0.0335	0.0257
SN	0.0148	1377	0.0466	0.0318

5. Direct estimation of continuous-time LPV systems

Commonly LPV controllers are synthesized in CT as stability and performance requirements of the closed loop behavior can be more conveniently expressed in CT, like in a mixed-sensitivity setting [Zhou and Doyle (1998)].

Furthermore, developing CT-LPV models based on first principle laws is a very costly and time consuming process, often resulting in a high order model unsuitable for control design. On the other hand, identification methods are almost exclusively expressed in DT, this being mainly motivated by an easier estimation of the parameter-varying dynamics. Unfortunately, transformation of DT-LPV models to CT-LPV models is more complicated than in the LTI case and despite recent advances in LPV discretization theory (see [Tóth *et al.* (2010)]), the theory of CT realization of DT models is still in an immature state. Thus, it is often difficult in practice to obtain an adequate CT realization from an identified DT model. In order to illustrate the underlying problems, consider the following simple CT-LPV model

$$\frac{\mathrm{d}}{\mathrm{d}t}y(t) + p(t)y(t) = bu(t), \tag{40}$$

where p, y and u are the scheduling, output and input signals of the system respectively and $b \in \mathbb{R}$ is a constant parameter. When approximating the derivative in DT by, for example, using the backward Euler approximation: $\frac{\mathrm{d}}{\mathrm{d}t}y(k) = \frac{y(k)-y(k-1)}{T_\mathrm{s}}$ with $T_\mathrm{s} > 0$ the sampling period, Eq. (40) transforms into:

$$y(k) = \frac{1}{1+T_\mathrm{s}p(k)}y(k-1) + \frac{bT_\mathrm{s}}{1+T_\mathrm{s}p(k)}u(k). \tag{41}$$

This discretized model has now two p-dependent coefficients to be estimated instead of the one single constant parameter in Eq. (40). Moreover, the dependence of the coefficients on p is not linear anymore but rational with a singularity whenever $p(k) = -\frac{1}{T_\mathrm{s}}$. An alternative way to approximate derivatives in DT is to apply a forward Euler approximation: $\frac{\mathrm{d}}{\mathrm{d}t}y(k) = \frac{y(k+1)-y(k)}{T_\mathrm{s}}$, which turns the CT model into:

$$y(k) = (1 - T_\mathrm{s}p(k-1))y(k-1) + bT_\mathrm{s}u(k-1). \tag{42}$$

This discretized model has only one p-dependent coefficient and the linearity of the dependence is preserved, however the model equation depends now on $p(k-1)$ instead of $p(k)$. This so-called dynamic dependence (dependence of the model coefficients on time-shifted versions of p) is a common result of model transformations in the LPV case and rises problems in LPV system identification and control alike (see [Tóth *et al.* (2008)]). Furthermore, it is well-known in numerical analysis that the forward Euler approximation is more sensitive for the choice of T_s in terms of numerical stability than the backward Euler approximation [Atkinson (1989)]. This means that Eq. (42) requires much faster sampling rate than Eq. (41) to give a stable approximation of the system and it is more sensitive to parameter uncertainties which rises problems if Eq. (42) is used for estimation. Consequently, it can be seen that even for a very simple CT-LPV model, estimation of a DT model with the purpose of obtaining afterwards a CT realization is a tedious task with many underlying problems for which there are no general theoretical solutions available: there is a large gap between the possibilities offered in terms of CT-LPV system modeling and the need of the control framework. A possible solution to the CT system identification is to use a direct approach in which the parameters of the CT model are directly identified [Garnier and Wang (2008)]. However, in practice, CT systems can only be identified based on sampled measured data records which results in several intrinsic issues for which some solutions are presented in this section.

5.1. *System description*

Consider the data generating CT-LPV system with static dependence on p described by the following equations

$$\mathscr{S}_\mathrm{o}\begin{cases} A_\mathrm{o}(p_t,d)\chi_\mathrm{o}(t)=B_\mathrm{o}(p_t,d)u(t)\\ y(t)=\chi_\mathrm{o}(t)+v_\mathrm{o}(t),\end{cases} \tag{43}$$

where d denotes the differentiation operator with respect to time, i.e., $d=\frac{\mathrm{d}}{\mathrm{d}t}$, $p:\mathbb{R}\to\mathbb{P}$ is the scheduling variable with $p_t=p(t)$, χ_o is the noise-free output, v_o is a quasi stationary noise process with bounded spectral density and uncorrelated to p. A_o and B_o are polynomials in d with coefficients a_i^o and b_i^o that are meromorphic functions of p

$$A_\mathrm{o}(p_t,d)=d^{n_\mathrm{a}}+\sum_{i=1}^{n_\mathrm{a}}a_i^\mathrm{o}(p_t)d^{n_\mathrm{a}-i},\quad\text{and}\quad B_\mathrm{o}(p_t,d)=\sum_{j=0}^{n_\mathrm{b}}b_j^\mathrm{o}(p_t)d^{n_\mathrm{b}-j}. \tag{44}$$

In terms of identification we can assume that sampled measurements of (y,p,u) are available using a sampling period $T_\mathrm{s}>0$. Hence, we will denote the DT samples of these signals as $u(k)=u(kT_\mathrm{s})$, where $k\in\mathbb{Z}$.

5.2. *Handling the time derivatives*

In comparison with the DT counterpart, direct CT model identification raises several technical issues. The first is related to implementation: the differential equation model is not a linear combination of the sampled process input and output signals but contains time-derivatives of the signals.

A suitable model fitting the system description assumes that these time-derivatives are available, and therefore, the parameter estimation procedure can be directly applied on them. However, these input and output time-derivatives are not available as measured data in most practical cases. A standard approach used in CT model identification is to introduce a low-pass stable filter $\mathscr{F}$, i.e., define

$$x_f(t) = f(d)x(t), \tag{45}$$

where the subscript f is used to denote the prefiltered forms of the associated variable x. The filtered time-derivatives can then be obtained by applying an array of filters in the form $f(d)p^i$:

$$x_f^{(i)}(t) = (f(d)d^i x)(t), \tag{46}$$

where $x^{(i)}(k)$ denotes the value of the time-derivative of signal $x(t)$. Various types of CT filters have been devised to deal with the need to reconstruct the time-derivatives [Garnier *et al.* (2003)]. Nonetheless, a critical issue is that these filters depend on design parameters, that need to be adequately chosen with respect to the underlying system (which is actually unknown in practice) to achieve reasonable performance of the estimation.

5.3. *Hybrid models*

Another technical issue is that direct CT model estimation, even in the LTI case, implies dealing with CT random noise sources, requiring a complicated mathematical machinery and introducing many problematic issues regarding estimation [Pintelon *et al.* (2000)]. Therefore, the basic idea to solve the noisy CT modeling problem is to assume that the CT noise process can be written at the sampling instances as a DT noise process filtered by a DT transfer function [Johansson (1994)]. This avoids the rather difficult mathematical problem of treating sampled CT random process and their equivalent in terms of a filtered piece-wise constant CT noise source (see [Pintelon *et al.* (2000); Johansson (1994)]). In terms of Eq. (43), $y(k)$ can be written as

$$y(k) = \chi_\mathrm{o}(k) + v_\mathrm{o}(k), \tag{47}$$

which corresponds to a so called hybrid Box-Jenkins system concept already used in CT identification of LTI systems (see [Pintelon *et al.* (2000); Johansson (1994); Laurain *et al.* (2008)]).

5.4. *Hybrid LPV-BJ polynomial models*

This section aims at presenting some possible tools in order to handle the estimation of CT LPV models. In this case, the p-dependent polynomials A and B are now polynomials in d:

$$A(p_t, d, \rho) = d^{n_a} + \sum_{i=1}^{n_a} a_i(p_t) d^{n_a - i}, \quad \text{and } B(p_t, d, \rho) = \sum_{j=0}^{n_b} b_j(p_t) d^{n_b - j}.$$

Their parametrization is defined using the set of functions $\psi_{i,l}$ as given in Section 4.1 and the coefficients are gathered into ρ in a very similar way as in the DT case. Moreover, in order to avoid the complexity of CT noise processes, the noise process is expressed in DT using the exact same model as in Eq. (19) and the parameters are stored in the vector η. Consequently, the full hybrid LPV polynomial model is expressed as:

$$\mathscr{M}_\theta \begin{cases} \underbrace{\chi(t) + \sum_{i=1}^{n_a} a_{i,0} \chi^{(n_a - i)}(t)}_{F(d,\rho)\chi(t)} + \sum_{i=1}^{n_a} \sum_{l=1}^{s_i} a_{i,l} \underbrace{\psi_{i,l}(p(t)) \chi^{(n_a - i)}(t)}_{\chi_{i,l}(t)} \\ \qquad = \sum_{j=0}^{n_b} \sum_{l=0}^{s_{\tilde{j}}} b_{j,l} \underbrace{\psi_{\tilde{j},l}(p(t)) u^{(n_b - j)}(t)}_{u_{j,l}(t)} \\ v(k) = \dfrac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} e(k) \\ y(k) = \chi(k) + v(k) \end{cases} \tag{48}$$

where $\tilde{j} = j + n_a + 1$ $F(d, \rho) = 1 + \sum_{i=1}^{n_a} a_{i,0} d^{n_a - i}$ and $\psi(i, 0)(t) = 1$. In this way, just like as in the DT case, the hybrid LPV-BJ model is rewritten as a MISO system with $\sum_{i=1}^{n_a+n_b+1} s_i + 1$ inputs $\{\chi_{i,l}\}_{i=1,l=1}^{n_a,s_i}$ and $\{u_{j,l}\}_{j=0,l=0}^{n_b,s_{\tilde{j}}}$:

$$\chi_{i,l}(t) = \psi_{i,l}(p(t)) \chi^{(n_a - i)}(t) \qquad \{i, l\} \in \{1 \ldots n_a, 1 \ldots s_i\}, \tag{49}$$

$$u_{j,l}(t) = \psi_{\tilde{j},l}(p(t)) u^{(n_b - j)}(t) \qquad \{j, l\} \in \{1 \ldots n_b, 1 \ldots s_{\tilde{j}}\}. \tag{50}$$

The hybrid representation of the model in Eq. (48) is a combined operation of CT and DT filtering. In order to clearly define the coexistence of DT and CT filtering, a detailed investigation and discussion about the assumptions and the structure of the model are needed. As previously pointed out, the practically feasible situation such that only sampled measurements of the CT signals (y, p, u) are available is considered here. In order to apply a CT filter on sampled data one can either interpolate the samples to obtain a CT signal and apply the CT filter on this reconstructed signal or use a numerical approximation, i.e., DT approximation of the considered system. This is a common problem for simulation of CT systems. For simulation purposes, DT approximation of the system can efficiently be dealt with by using powerful numerical algorithms available [Atkinson (1989)]. Note that to derive an accurate DT approximation of the system, it is often sufficient in terms of the classical discretization theory to assume that the sampled free CT signals of the system are restricted to be constant in the sampling period [Goodwin *et al.* (2000)], which has also been shown in case of LPV systems with static dependence [Tóth (2010)]. This provides the hypothesis, also used in [Pintelon *et al.* (2000); Johansson (1994)], that if CT (p, u) are piecewise constant between two samples, then the trajectory of y is completely determined by its observations at the sample period $T_{\mathrm{s}}k$. Therefore, under these inter-sampling conditions, the following operation is well-defined [Garnier and Wang (2008)]:

$$\left(F(d)y\right)(t_k) = F(d)y(t_k), \tag{51}$$

Under this assumption, and considering that a CT filter can only be applied to sampled data through numerical approximation, the usual filter properties such as commutativity holds between a DT filter and the numerical approximation of a CT filter. Nevertheless, it is important to notice that the numerical approximation method used for the evaluation of a CT filter does not have any impact on the coefficients to be estimated which remain, the coefficients of the parsimonious CT model. Consequently, $\mathscr{M}_\theta$ can be rewritten as

$$y(k) = -\left(\sum_{i=1}^{n_{\mathrm{a}}}\sum_{l=1}^{s_i}\frac{a_{i,l}}{F(d,\rho)}\chi_{i,l}\right)(k) + \left(\sum_{j=0}^{n_{\mathrm{b}}}\sum_{l=0}^{s_{\bar{j}}}\frac{b_{j,l}}{F(d,\rho)}u_{k,j}\right)(k) + H(q,\eta)e(k). \tag{52}$$

which is an LTI representation.

Based on this model structure, the model set, denoted as $\mathcal{M}$, with process takes the form

$$\mathcal{M} = \left\{\mathscr{M}_\theta \;\middle|\; \left[\rho^\top, \eta^\top\right]^\top = \theta \in \mathbb{R}^{n_\rho + n_\eta}\right\}. \tag{53}$$

6. Instrumental variable approach in continuous-time

Based on the MISO-LTI formulation Eq. (52), the LTI theory can be extended to achieve optimal PEM using linear regression. Nonetheless, the inherent problem of the time-derivative estimation still holds and a solution based on the RIV method for CT models (RIVC) proposed in [Laurain *et al.* (2011)] is detailed in this section.

6.1. *The optimal instrument for hybrid LPV-BJ systems*

Using the LTI model Eq. (48), $y(k)$ can be written in this CT case under the regression form:

$$y^{(n_\mathrm{a})}(k) = \varphi^\top(k)\rho + \tilde{v}(k), \tag{54}$$

where

$$\varphi(k)=[-y^{(n_\mathrm{a}-1)}(k)\dots -y(k)-\chi_{1,1}(k)\dots-\chi_{n_\mathrm{a},s_{n_\mathrm{a}}}(k)u_{0,0}(k)\dots u_{n_\mathrm{b},s_{n_\mathrm{g}}}(k)\,]^\top$$

$$\tilde{v}(k) = F(d,\rho)v(k) = F(d,\rho)\frac{C(q^{-1},\eta)}{D(q^{-1},\eta)}e(k).$$

The definition of the optimal filter and instrument can be defined in a very similar way as in the DT case [Laurain *et al.* (2011)]. The optimal instrument is given as:

$$\zeta_{\mathrm{opt}}(k) = \Big[-\chi_\mathrm{o}^{(n_\mathrm{a}-1)}(k-1) \quad \dots \quad -\chi_\mathrm{o}(k) - \chi_{1,1}^\mathrm{o}(k) \quad \dots \\ -\chi_{n_\mathrm{a},s_{n_\mathrm{a}}}^\mathrm{o}(k) \quad u_{0,0}(k) \quad \dots \quad u_{n_\mathrm{b},s_{n_\mathrm{g}}}(k)\Big]^\top \tag{55}$$

while the optimal filter is given as [Young *et al.* (2007)]

$$y_\mathrm{f}(k) = \frac{D_\mathrm{o}(q^{-1})}{C_\mathrm{o}(q^{-1})}\left[\left(\frac{1}{F_\mathrm{o}(d)}y\right)(k)\right]. \tag{56}$$

As pointed out in Section 5.2, it is important to note that the regressor and instrument contain time-derivatives of y and u which, in the assumed framework considering sampled data, can only be approximated and therefore need some low pass filtering. However, it can be seen from Eq. (27), that the IV optimization problem does not require the knowledge of the instrument and regressor, but of their filtered version. It can be moreover seen from the filter expression (56) that $F_\mathrm{o}(d)$ (or its estimated version $F(d,\rho)$ in practice) achieves this stable low-pass filtering directly. Therefore, it is a particular strength of the presented reformulation that the estimated filter $F(d,\rho)$ is not only used for the minimization of the prediction error but it also provides an efficient solution to the signal time-derivatives approximation problem.

6.1.1. *The LPV-RIVC algorithm for BJ models*

Just as in the DT case, the signals $\{\chi_{i,l}(k)\}_{i=1,l=0}^{n_\mathrm{a},s_i}$ and the knowledge of the different filters involved in Eq. (56) are unknown *a priori*. Therefore, similarly to the DT case, an iterative procedure is proposed to refine these unknowns and to aim at the PEM. The iterative RIV algorithm for hybrid LPV-BJ models reads as follows.

Algorithm 3 (LPV-RIVC).

Step 1 *Generate an initial estimate $\hat{\rho}^{(0)}$ using the SRIVC algorithm (by forcing an LTI structure). Set $C(d,\hat{\eta}^{(0)}) = D(d,\hat{\eta}^{(0)}) = 1$. Set $\tau = 0$.*

Step 2 *Compute an estimate of $\chi(k)$ via numerical approximation of*

$$A(p_t, d, \hat{\rho}^{(\tau)})\hat{\chi}(t) = B(p_t, d, \hat{\rho}^{(\tau)})u(t).$$

Step 3 *Compute the estimated CT filter using:*

$$L_\mathrm{c}(d,\hat{\rho}^{(\tau)}) = \frac{1}{F(d,\hat{\rho}^{(\tau)})} \tag{57}$$

where $F(d,\hat{\rho}^{(\tau)})$ is as given in Eq. (48).

Step 4 *Use the CT filter $L_\mathrm{c}(d,\hat{\rho}^{(\tau)})$ as well as $\hat{\chi}(k)$ in order to generate the estimates of the derivatives which are needed:*

$$\{L_\mathrm{c}(d,\hat{\rho}^{(\tau)})\hat{\chi}^{(i)}(k)\}_{i=0}^{n_\mathrm{a}-1}, \qquad \{L_\mathrm{c}(d,\hat{\rho}^{(\tau)})\hat{y}^{(i)}(k)\}_{i=0}^{n_\mathrm{a}-1}$$

$$\{L_\mathrm{c}(d,\hat{\rho}^{(\tau)})u_{j,l}(k)\}_{j=0,l=0}^{n_\mathrm{b},s_{\bar{j}}}, \qquad \{L_\mathrm{c}(d,\hat{\rho}^{(\tau)})\hat{\chi}_{i,l}(k)\}_{i=1,l=0}^{n_\mathrm{a},s_i}$$

Step 5 *Compute the estimated DT filter:*

$$L_\mathrm{d}(q,\hat{\eta}^{(\tau)}) = \frac{D(q^{-1},\hat{\eta}^{(\tau)})}{C(q^{-1},\hat{\eta}^{(\tau)})}.$$

Step 6 *The filtered signals $\{u^\mathrm{f}_{j,l}(k)\}_{j=0,l=0}^{n_\mathrm{b},s_{\bar{j}}}$, $y_\mathrm{f}(k)$ and $\{\chi^\mathrm{f}_{i,l}(k)\}_{i=1,l=0}^{n_\mathrm{a},s_i}$ are computed using the DT filter on the estimated derivatives obtained from Step 4.*

Step 7 *Build the filtered regressor $\hat{\varphi}_\mathrm{f}(k)$ and, in terms of C4, the filtered instrument $\hat{\zeta}_\mathrm{f}(k)$ as:*

$$\hat{\varphi}_\mathrm{f}(k) = \Big[-y_\mathrm{f}^{(n_\mathrm{a}-1)}(k) \ldots -y_\mathrm{f}(k) \\ -\hat{\chi}^\mathrm{f}_{1,1}(k) \ldots -\hat{\chi}^\mathrm{f}_{n_\mathrm{a},s_{n_\mathrm{a}}}(k)\ u^\mathrm{f}_{0,0}(k) \ldots u^\mathrm{f}_{n_\mathrm{b},s_{n_\mathrm{g}}}(k) \Big]^\top$$

$$\hat{\zeta}_{\mathrm{f}}(k) = \Big[-\hat{\chi}_{\mathrm{f}}^{(n_{\mathrm{a}}-1)}(k-1) \ldots -\hat{\chi}_{\mathrm{f}}(k) \\ -\hat{\chi}_{1,1}^{\mathrm{f}}(k) \ldots -\hat{\chi}_{n_{\mathrm{a}},s_{n_{\mathrm{a}}}}^{\mathrm{f}}(k)\; u_{0,0}^{\mathrm{f}}(k) \ldots u_{n_{\mathrm{b}},s_{n_{\mathrm{g}}}}^{\mathrm{f}}(k) \Big]^{\top}.$$

Step 8 *The IV solution is obtained as*

$$\hat{\rho}^{(\tau+1)}(N) = \left[\sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(k) \hat{\varphi}_{\mathrm{f}}^{\top}(k) \right]^{-1} \sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(k) y_{\mathrm{f}}^{(n_{\mathrm{a}})}(k).$$

The resulting $\hat{\rho}^{(\tau+1)}(N)$ *is the IV estimate of the process model associated parameter vector at iteration* $\tau + 1$ *based on the prefiltered input/output data.*

Step 9 *An estimate of the noise signal* v *is obtained as*

$$\hat{v}(k) = y(k) - \hat{\chi}(k, \hat{\rho}^{(\tau)}). \tag{58}$$

Based on $\hat{v}$*, the estimation of the noise model parameter vector* $\hat{\eta}^{(\tau+1)}$ *follows, using in this case the ARMA estimation algorithm of the MATLAB identification toolbox (an IV approach can also be used for this purpose, see [Young (2008)]).*

Step 10 *If* $\theta^{(\tau+1)}$ *has converged or the maximum number of iterations is reached, then stop, else increase* τ *by* 1 *and go to Step 2.*

6.1.2. *The LPV-SRIVC algorithm for OE models*

Based on a similar concept and like in the DT framework, the so-called *simplified* LPV-RIVC (LPV-SRIVC) method, can also be developed for the estimation of CT-LPV-OE models. Therefore LPV-SRIVC uses the same algorithm as LPV-RIVC except for Step 5, 6 and 10 of Algorithm 3 which are skipped. It can be noticed, as well, that CT-LPV-OE models do not involve any DT filtering and, consequently, their structure is fully CT unlike the hybrid BJ LPV model.

6.2. *Examples*

As a next step, the performance of the described algorithms are presented on a representative simulation example. The system taken into consideration is inspired by a benchmark example proposed by Rao and Garnier in

[Rao and Garnier (2004)]. Since then, it has been widely used to demonstrate the performance of direct CT identification methods [Ljung (2003); Campi *et al.* (2008); Laurain *et al.* (2008); Rao and Garnier (2004)]. In order to create a CT-LPV system on which the strength of direct CT identification can be demonstrated, a "moving pole" is considered. A particular feature of LPV systems is that they have an LTI representation for every constant trajectory of p. Such an LTI representation describes the so-called frozen behavior of the system and can be expressed in a transfer function form. In terms of the frozen concept, the "moving pole" means that a particular pole of these frozen transfer functions of $\mathscr{S}_\mathrm{o}$ is a function of p. This phenomenom often occurs in mechatronic applications such as for instance, wafer scanners [Wassink (2002)]. In our case, the Rao-Garnier benchmark inspired "moving pole" LPV system is a fourth order system with non-minimum phase frozen dynamics and a p-dependent complex pole pair. It is defined as follows:

$$\mathscr{S}_\mathrm{o}\begin{cases} A_\mathrm{o}(d,p) = \omega_2^2(p)\omega_1^2 + \left(2\zeta_2\omega_2(p)\omega_1^2 + 2\zeta_1\omega_2^2(p)\omega_1\right)d + \\ \quad \left(\omega_1^2 + \omega_2^2(p) + 4\zeta_2\zeta_1\omega_2(p)\omega_1\right)d^2 + \left(2\zeta_2\omega_2(p) + 2\zeta_1\omega_1\right)d^3 + d^4 \\ B_\mathrm{o}(d,p) = -K\omega_2^2(p)\omega_1^2 d + \omega_2^2(p)\omega_1^2 \\ H_\mathrm{o}(q) = \dfrac{1}{1-q^{-1}+0.2q^{-2}}, \end{cases} \tag{59}$$

where $K = 4$ [sec], $\omega_1 = 20$ [rad/s], $\zeta_1 = 0.1$, $\zeta_2 = 0.5$. The slow frozen mode ω_2 is p-dependant and chosen as: $\omega_2 = 2{+}0.5p$. Notice that the frozen behavior (p is fixed to a constant trajectory) of $\mathscr{S}_\mathrm{o}$ for $p = 0$ corresponds exactly to the Rao-Garnier benchmark defined as

$$G_\mathrm{RG}(d) = \frac{-Kd+1}{(\frac{d^2}{\omega_1^2} + 2\zeta_1\frac{d}{\omega_1} + 1)(\frac{d^2}{\omega_2^2(0)} + 2\zeta_2\frac{d}{\omega_2(0)} + 1)}. \tag{60}$$

$\mathscr{S}_\mathrm{o}$ takes therefore the following form

$$\mathscr{S}_\mathrm{o}\begin{cases} A_\mathrm{o}(d,p) = d^4 + a_1^\mathrm{o}(p)d^3 + a_2^\mathrm{o}(p)d^2 + a_3^\mathrm{o}(p)d + a_4^\mathrm{o}(p) \\ B_\mathrm{o}(d,p) = b_0^\mathrm{o}(p)d + b_1^\mathrm{o}(p) \\ H_\mathrm{o}(q) = \dfrac{1}{1-q^{-1}+0.2q^{-2}}, \end{cases} \tag{61}$$

where

$$a_1^\mathrm{o}(p) = 5 + 0.25p, \qquad a_2^\mathrm{o}(p) = 408 + 3p + 0.25p^2, \tag{62a}$$
$$a_3^\mathrm{o}(p) = 416 + 108p + p^2, \qquad a_4^\mathrm{o}(p) = 1600 + 800p + 100p^2, \tag{62b}$$
$$b_0^\mathrm{o}(p) = -6400 - 3200p - 400p^2, \quad b_1^\mathrm{o}(p) = 1600 + 800p + 100p^2. \tag{62c}$$

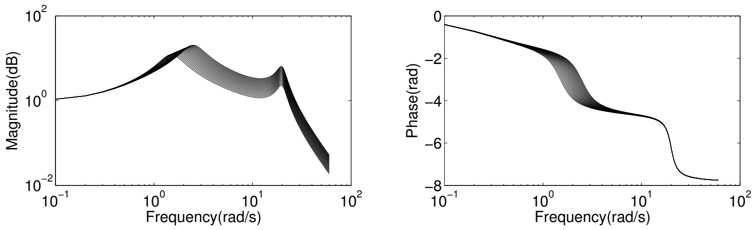

Fig. 2. Bode plot of the frozen behaviors of the true LPV system for 20 values of the scheduling variable p (between -1 to 1).

The Bode plot of 20 frozen behaviors of $\mathscr{S}_\text{o}$ is depicted in Fig. 2 for 20 fixed values of the scheduling variable p equally distributed from -1 to 1 where the consequence of the moving low frequency mode can be clearly observed. To obtain data records for identification purposes, the input signal u is chosen as a uniformly distributed sequence $\mathscr{U}(-1,1)$, with a scheduling trajectory $p(t) = \sin(\pi t)$. Furthermore, the sampling period is chosen as $T_\text{s} = 1\text{ms}$, and the simulation time is $T_\text{max} = 10\text{s}$ which, considering the currently available acquisition possibilities, is a fair assumption. In the sequel, the LPV-RIVC algorithm is studied for the identification of the data generating system $\mathscr{S}_\text{o}$. The proposed LPV-RIVC method is applicable to hybrid LPV-BJ model and assumes the following model structure:

$$\mathscr{M}_{\text{LPV}-\text{RIVC}} \begin{cases} A(d,p) = d^4 + a_1(p)d^3 + a_2(p)d^2 + a_3(p)d + a_4(p) \\ B(d,p) = b_0(p)d + b_1(p) \\ H(q) = \dfrac{1}{1 + d_1 q^{-1} + d_2 q^{-2}}, \end{cases}$$

where

$$a_1(p) = a_{1,0} + a_{1,1}p, \qquad a_2(p) = a_{2,0} + a_{2,1}p + a_{2,2}p^2, \tag{63a}$$

$$a_3(p) = a_{3,0} + a_{3,1}p + a_{3,2}p^2, \qquad a_4(p) = a_{4,0} + a_{4,1}p + a_{4,2}p^2, \tag{63b}$$

$$b_0(p) = b_{0,0} + b_{0,1}p + b_{0,2}p^2, \qquad b_1(p) = b_{1,0} + b_{1,1}p + b_{1,2}p^2. \tag{63c}$$

In terms of identification, the model $\mathscr{M}_{\texttt{LPV-RIVC}}$ corresponds to the case $\mathscr{S}_\text{o} \in \mathcal{M}$ and 19 parameters are to be estimated. The results of a Monte Carlo simulation are presented, using $N_\text{run} = 200$ random realizations with a SNR of 20dB. In order to assess the performance of the presented method, the Bode plot of each estimated model ($N_\text{MC} = 200$) at 20 frozen values of p from -1 to 1 are depicted in Fig. 3. It can be clearly observed from the Bode plot that there is a lack of excitation in the low frequencies area. To analyze the distribution of responses at low frequencies, the densities of curves

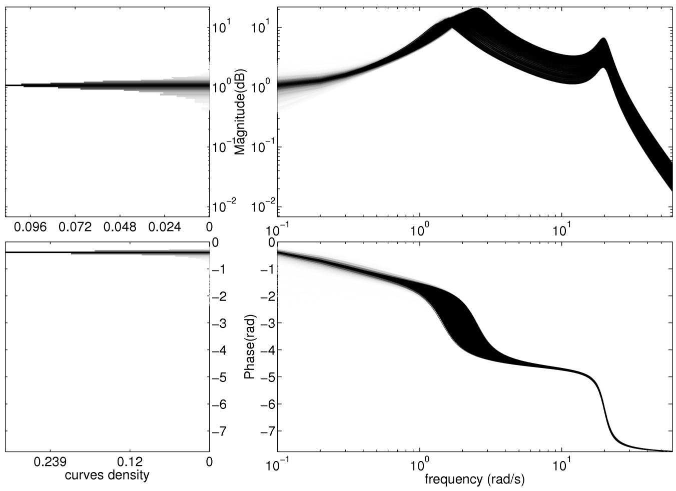

Fig. 3. Bode plots of the estimated models.

at frequency $\omega = 0.1$ rad/s are displayed on the left hand-side of Fig. 3 for both magnitude and phase diagrams. Furthermore, this distribution density is used as color coding for drawing the frequency responses in the Bode plot: the darker a given line, the higher the number of estimated models having the same response at $\omega = 0.1$ rad/s. It becomes clear that the estimated models are normally distributed and centered on the true model. Finally, the most interesting advantage of the direct CT estimation in LPV framework is shown in Fig. 4 where the poles of the 200 estimated models at 20 fixed values of p between -1 and 1 are plotted along with the true model poles. This figure also displays in the top part the density of the pole real parts. Using the same idea as in Fig. 3, the intensity of each displayed pole is related to the number of estimated poles which have the same real part. Using this representation, it can be clearly seen that the pole distribution around the "fixed pole" (around -2 ± 20i) has a Gaussian distribution while there is a sparse repartition around the "moving pole" (around $(-0.5 \pm 0.16) \pm (2 + 0.5p)$i). Furthermore, the imaginary part of all estimated poles is close to the true imaginary part of the "fixed pole". This means that by using a parsimonious CT model, the estimated models perfectly capture the time-varying nature of $\mathscr{S}_{\mathrm{o}}$ as well as the transient behavior from one operating condition to the other, which is a relevant achievement considering the complexity of the studied system and the level of additive noise. It must be noticed that trying to identify a DT model

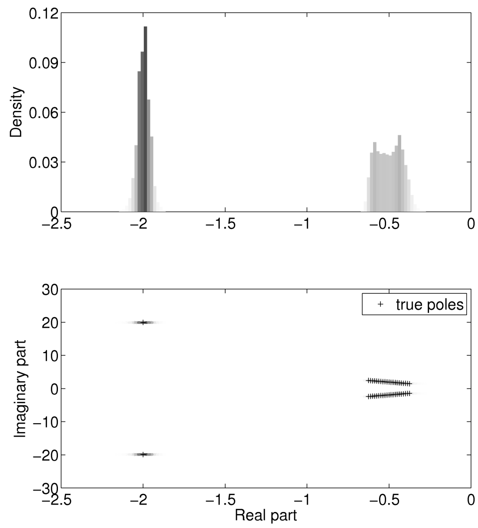

Fig. 4. Poles of the estimated models and of the true system.

for such a system is quite a tedious task. As pointed out in Section 5, the discretization of the system Eq. (59) would result in i) the augmentation of the number of parameters to be estimated, ii) the non-linear-in-the-parameter dependence on p iii) some dynamic dependence on p (appearance of $p(k-1), p(k-2)\dots$ terms). Moreover, the time-varying property of the LPV system (in this case, a "moving pole") would not anymore explicitly appear in the DT model coefficients and might not be clearly identified.

7. Conclusion

Input-output LPV models are widely used in real world applications but the dedicated identification methods have not focused on realistic noise conditions until recently. This chapter aimed at presenting some recently developed methods dedicated to the identification of IO models under different noise assumptions. It was pointed out that the most common assumption of an LPV-ARX model is attractive in the sense that it allows the direct extension of well-known LTI methods (LS and IV4). Unfortunately, this noise assumption is very restrictive and is very unlikely to be verified in practice. In order to cope with more general types of noise, a reformulation of the LPV model into a MISO LTI representation has been proposed along with some RIV based methods to solve the estimation problem. The importance of the noise model assumption and of the associated identification methods have been demonstrated through a bias and variance analysis on representative examples. Whereas most methods are developed in DT settings, very few solutions have been proposed in the literature to deal with the direct identification of CT-LPV models. Therefore, there is a large gap between the control needs and the identification solutions proposed. The inherent difficulties of direct CT identification have been listed and some solutions have been proposed relying on a hybrid representation of CT systems. An RIVC based method has been introduced, based on the same model reformulation as the one presented in the DT framework. In the CT framework however, the iterative filtering used for aiming at the PEM allows also an accurate time-derivative approximation. The method was analyzed on a fourth order moving pole system with colored noise and the advantages of direct approaches such as the possibility of capturing the true dynamical behavior of the system have been addressed. It can be finally pointed out that most of the methods presented here can be extended to the closed-loop case [Boonto and Werner (2008); Abbas and Werner (2009); Butcher *et al.* (2008); Tóth *et al.* (2011a)] for which the added difficulty lies in the noise corrupted input.

References

Abbas, H. and Werner, H. (2009). An instrumental variable technique for open-loop and closed-loop identification of input-output LPV models, in *Proceedings of the European Control Conference* (Budapest, Hungary).

Atkinson, K. E. (1989). *An introduction to numerical analysis* (John Wiley & Sons, Singapore).

Bamieh, B. and Giarré, L. (2002). Identification of linear parameter varying models, *International Journal of Robust and Nonlinear Control* **12**, 9, pp. 841–853.

Boonto, S. and Werner, H. (2008). Closed-loop system identification of LPV input-output models - application to an arm-driven pendulum, in *Proceedings of the IEEE Conference on Decision and Control* (Cancún, Mexico).

Butcher, M., Karimi, A. and Longchamp, R. (2008). On the consistency of certain identification methods for linear parameter varying systems, in *Proceedings of the IFAC World Congress* (Seoul, Korea).

Campi, M., Sugie, T. and Sakai, F. (2008). An iterative identification method for linear continuous-time systems, *IEEE Transactions on Automatic Control* **53**, 7, pp. 1661–1669.

Garnier, H., Mensler, M. and Richard, A. (2003). Continuous-time model identification from sampled data: implementation issues and performance evaluation, *International Journal of Control* **76**, 13, pp. 1337–1357.

Garnier, H. and Wang, L. e. (2008). *Identification of continuous-time models from sampled data* (Springer-Verlag, London, UK).

Giarré, L., Bauso, D., Falugi, P. and Bamieh, B. (2003). Identification for gain scheduling control: an application to rotating stall and surge control problem, *Control Engineering Practice* **14**, 4, pp. 351–361.

Goodwin, G., Graebe, S. and Salgado, M. (2000). *Control system design* (Prentice-Hall, Englewood Cliffs, NJ, USA).

Johansson, R. (1994). Identification of continuous–time models, *IEEE Transactions on Signal Processing* **42**, 4, pp. 887–897.

Laurain, V., Gilson, M., Garnier, H. and Young, P. C. (2008). Refined instrumental variable methods for identification of Hammerstein continuous time Box-Jenkins models, in *Proceedings of the IEEE Conference on Decision and Control* (Cancún, Mexico).

Laurain, V., Gilson, M., Tóth, R. and Garnier, H. (2010). Refined instrumental variable methods for identification of LPV output-error and Box-Jenkins models, *Automatica* **46**, 6, pp. 959–967.

Laurain, V., Tóth, R., Gilson, M. and Garnier, H. (2011). Direct identification of continuous-time LPV input/ouput models, *IET Control Theory & Applications* **5**, 7, pp. 878–888.

Ljung, L. (1999). *System identification: theory for the user - Second edition* (Prentice-Hall, Upper Saddle River, NJ, USA).

Ljung, L. (2003). Initialisation aspects for subspace and output-error identification methods, in *Proceedings of the European Control Conference* (Cambridge, U.K.).

Ljung, L. (2009). Experiments with identification of continuous time models, in *Proceedings of the IFAC Symposium on System Identification* (Saint-Malo, France).

Lopes dos Santos, P., Ramos, J. A. and Martins de Carvalho, J. L. (2007). Identification of linear parameter varying systems using an iterative deterministic-stochastic subspace approach, in *Proceedings of the European Control Conference*.

Lovera, M. and Mercère, G. (2007). Identification for gain-scheduling: a balanced subspace approach, in *Proceedings of the American Control Conference* (New York City, NY, USA).

Pintelon, R., Schoukens, J. and Rolain, Y. (2000). Box–Jenkins continuous–time modeling, *Automatica* **36**, 7, pp. 983–991.

Rao, G. P. and Garnier, H. (2004). Identification of continuous-time systems: direct or indirect ? *Systems Science* **30**, 3, pp. 25–50.

Söderström, T. and Stoica, P. (1983). *Instrumental variable methods for system identification* (Springer-Verlag, New York, NY, USA).

Tóth, R. (2010). *Modeling and identification of linear parameter-varying systems*, *Lecture Notes in Control and Information Sciences*, Vol. 403 (Springer-Verlag, Berlin/Heidelberg, Germany).

Tóth, R., Heuberger, P. S. C. and Van den Hof, P. M. J. (2008). Flexible model structures for LPV identification with static scheduling dependency, in *Proceedings of the IEEE Conference on Decision and Control* (Cancún, Mexico).

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., Heuberger, P. S. C. and Van den Hof, P. M. J. (2010). On the discretization of LPV state-space representations, *IET Control Theory & Applications* **4**, 4, pp. 2082–2096.

Tóth, R., Laurain, V., Gilson, M. and Garnier, H. (2011a). On the closed loop identification of LPV models using instrumental variables, in *Proceedings of the the IFAC World Congress* (Milan, Italy).

Tóth, R., Laurain, V., Zheng, W. and Poolla, K. J. (2011b). Model structure learning: a support vector machine approach for LPV linear-regression models, in *submitted to the 50th 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).

van Wingerden, J. W. and Verhaegen, M. (2009). Subspace identification of bilinear and LPV systems for open- and closed-loop data, *Automatica* **45**, 2, pp. 372–381.

Wassink, M. G. (2002). *LPV control for wafer stages*, Msc thesis, CTB595-02-2097, Philips CFT, Eindhoven, The Netherlands.

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).

Young, P. (2008). The refined instrumental variable method: Unified estimation of discrete and continuous-time transfer function models, *Journal Européen des Systèmes Automatisés* **42**, pp. 149–179.

Young, P., Garnier, H. and Gilson, M. (2007). *Identification of continuous-time models from sampled data*, chap. Refined instrumental variable identification of continous-time hybrid Box-Jenkins models (Springer-Verlag, London, UK), pp. 91–131.

Young, P. C. (1984). *Recursive estimation and time-series analysis* (Springer-Verlag, Berlin, Germany).

Young, P. C. and Jakeman, A. (1980). Refined instrumental variable methods of recursive time-series analysis —part III. extensions, *International Journal of Control* **31**, pp. 441–764.

Zhou, K. and Doyle, J. C. (1998). *Essentials of robust control* (Prentice-Hall, Upper Saddle River, NJ, USA).